5)

Parameter Estimation in Nonlinearly Parameterized Systems

by

Aleksandar M. Kojić

Submitted to the Department of Mechanical Engineering in partial fulfillment of the requirements for the degree of

Master of Science in Mechanical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1998

Author . . . .
Department of Mechanical Engineering
May 26, 1998

Certified by . . . .
Anuradha M. Annaswamy
Associate Professor
Thesis Supervisor

Accepted by . . . .
Ain A. Sonin
Chairman, Departmental Committee on Graduate Students

Parameter Estimation in Nonlinearly Parameterized Systems

by

Aleksandar M. Kojić

Submitted to the Department of Mechanical Engineering
on May 26, 1998, in partial fulfillment of the
requirements for the degree of
Master of Science in Mechanical Engineering

Abstract

This thesis presents a treatise on some topics of the problem of parameter identification in systems whose behavior is governed by nonlinearly parameterized functions. The emphasis of the thesis is on the analytical treatment of stability and parameter convergence issues. The thesis examines two cases of nonlinear parameterization: convex/concave and monotonic, and their corresponding estimation algorithms. In the case of the convex/concave parameterization, the conditions for parameter convergence are derived for a recently developed min-max estimator. In linearly parameterized systems, parameter convergence conditions impose requirements solely on the outside input to the system of interest. However, it is shown that with the convex/concave functions and the min-max algorithm the system governing function must satisfy certain prerequisites if the unknown parameter value is to be identified precisely. Several examples of functions and corresponding inputs which both satisfy and do not satisfy the prescribed conditions are given. In the case of monotonic parameterizations, two types of systems are considered: one where only the filtered function output is available for measurement, and one where that output is directly measurable. In the first case, since the function output is not known exactly at each instant of time, it is shown that instability can result with the gradient parameter update law. The second case deals with nonlinear parameterization by a specific kind of monotonic function, the sigmoidal function. The sigmoidal function is often utilized in neural network applications. For a reduced order neural network, it is shown how parameter convergence can be analytically guaranteed with the local gradient update law.

Thesis Supervisor: Anuradha M. Annaswamy
Title: Associate Professor

Мојим родитељима

Acknowledgments

It has been often said that what may seem as individual achievements are only the reflection of underlying complex and intense interactions and experiences with other individuals. Most certainly, that has been the case in the development of this thesis. Hence, I would like to express my gratitude to a number of persons without whom all of this would not have been possible.

I would like to thank my advisor, *Professor Anuradha Annaswamy*, for giving me the opportunity to work with her on a number of very interesting and challenging analytical problems, and for her invaluable guidance, support and patience throughout the research process.

Outside of the world of research, I greatly acknowledge the help and support received from *Mijailović* and *Mikić* families. Surely, their assistance was of fundamental importance in enabling me to concentrate on the required academic tasks.

I am very grateful to *Zoran Spasojević* and *Srdjan Divac* for their aid in dealing with complex mathematical problems. I also thank *Ssu-Hsin Yu*, *Fredrik Skantze* and *Ashish Krupadanam* from the MIT Adaptive Control Laboratory for their help and many insightful discussions on a variety of topics. Many thanks also to the other students of the Adaptive Control Laboratory, *Rienk Bakker*, *Steingrimur Karason*, *Kathy Misovec*, *Chayakorn Thanomsat*, *Jennifer Rumsey*, and *Jean-Pierre Hathout* for creating a pleasant atmosphere in the lab.

I would also like to thank my friends for their support during my stay in the United States. Rest assured, your contribution is noted and appreciated. Also, thanks are in order for *Kate Melvin* and *Leslie Regan* for their help in administrative issues.

Last, but by no means least, many greatful thanks go to my family. Needless to say, their support and encouragement is a pillar on which stand all of my current personal achievements.

Contents

Chapter 1

Introduction

1.1 Motivation and Previous Work

Through the careful use of conservation laws, constitutive relations and geometric compatibility constraints, algebraic models can be derived which accurately capture and predict future behavior of the observed system. Regularly, such models are generic in the sense that they are applicable to a whole class of the systems similar to the one observed. What sets the particular observed system apart from the rest of its family is the value of the certain constant quantities. These constant quantities are called the system parameters. In modeling complex systems, it is often found that these parameters enter the system nonlinearly. Examples of such nonlinearly parameterized systems are models for describing low-velocity friction [3], magnetic bearings [2], chemical reactors [5], combustion models [7, 24], and various Hammerstein-Uryson representations [6]. Complex nonlinearly parameterized models, like neural networks [20], can represent many systems for which only the input and output quantities are available with no knowledge of the underlying physical interactions.

The values of the physical system parameters are a very important characteristic of the system. They often portray the state of the system, and provide invaluable information for tasks such as fault detection and diagnosis. Many powerful techniques for control of nonlinear systems exist [13, 12] if the values of the parameters are known.

In the case that the values of the parameters are unknown, there are available control techniques if the system depends linearly on its parameters [19, 14, 15]. The control strategy in this case is often coupled with a methodology for estimating the values of the unknown parameters. The so-called "persistent excitation" requirements [19] state under what conditions the values of the linear parameters can be obtained precisely. In the case that the parameters enter the system nonlinearly, brute force methods for control exist [27] if the bounds on the values of the parameters are known.

In the case that the nonlinear parameters are unknown, few analytical results for obtaining their values are available. In [29] a strategy was proposed that uses neural networks for estimating the values of the nonlinear parameters in dynamical systems of interest. However, neural networks require that their internal nonlinear parameters be determined for the task at hand. The presented strategy, like all other neural network training methodologies, relies on the numerical calculation of the neural network parameters . The approach used in [8] suffices for stabilization in a limited class of nonlinear parameterizations, but is inadequate for tracking. In the presence of general nonlinear parameterization a novel approach through a min-max strategy for designing an adaptive controller was presented in [16]. However, this approach does not address the issues of parameter convergence.

1.2 Contribution of the thesis

Arguably, due to the plethora of nonlinear functions which are encountered in different systems, no single design scheme can attempt to solve the problem of precise parameter estimation for all nonlinear parameterizations. Rather, a case by case approach seems more feasible. This thesis takes that approach as it examines two different cases of nonlinear parameterization: the convex/concave and monotonic parameterization.

The thesis builds on the results found in [1, 16, 26]. These results pertain to the problem of adaptively controlling nonlinear systems with convex/concave or general nonlinear parameterization. In the first part of the thesis, parameter convergence conditions of the proposed controllers for convex/concave parameterization are derived.

Unlike their linear counterparts, it is shown that these conditions impose requirements not only on the outside input, but also on the type of convex/concave nonlinearity. Parameter convergence results are displayed for nonlinearities that satisfy the derived conditions. A type of nonlinearity which does not satisfy these conditions is also presented.

In the second part of the thesis, monotonic parameterization is addressed. The use of the local gradients in parameter estimation is investigated. It is shown how for some type of systems, this method can lead to instability and divergence. For the case of neural network models, a new methodology for examining system behavior is introduced. The demand that a Lyapunov-like energy function be always nonincreasing is relaxed. Rather, the asymptotic behavior of the system is examined. With this approach, it is shown how local gradients can lead to global stability by examining a low order system.

1.3 Notation

The thesis follows the notation used in many of the similar works on adaptive control. The unknown parameters are denoted by θ. Depending on the number of unknown parameters present in the system, θ can be a scalar or a vector, in which case it will consist of N components θ_i, $i = 1, ..., N$. In the case θ is a vector, it will represent a point in a N-dimensional space with coordinate axes specified by unit vectors $\mathbf{i}_i$, $i = 1, ..., N$. An estimate of θ at a time instant t is denoted by $\widehat{\theta}(t)$. Accordingly, the value of a function $f(t, \theta)$ is estimated as $\widehat{f}(t, \widehat{\theta})$, and the error between the estimate and the actual value is denoted as $\widetilde{f}\left(t, \widetilde{\theta}(t)\right) = \widehat{f} - f$. Assuming a given constant θ, the function $\widetilde{f}$ is a function of two time-varying variables, $\phi(t)$ and $\widetilde{\theta}(t)$. Hence, $\widetilde{f}\,(t_1, t_2)$ represents $\widetilde{f}$ evaluated at the point $\phi(t_1)$ and $\widetilde{\theta}(t_2)$. $\widetilde{f}\,(t_1)$ is used as a shorthand form of $\widetilde{f}\,(t_1, t_1)$.

1.4 Organization of the thesis

The organization of the thesis is as follows. Chapter 2 summarizes the min-max controller for convex/concave functions and derives conditions on the system input and the present nonlinearity for guaranteeing parameter convergence. Chapter 3 discusses the possible instability which can occur with the gradient algorithm in monotonically parameterized systems where the function output is not immediately available for measurement. In Chapter 4 the applicability of the gradient algorithm on certain classes of systems parameterized by sigmoidal nonlinearities is investigated. A necessary and a sufficient condition for parameter convergence in a low order neural network are given. Concluding remarks are offered in Chapter 5.

Chapter 2

Convex/Concave Parameterization

2.1 Introduction

Based on observation and physical laws, for many systems of interest the general form of the function which can adequately represent observed behavior is known. However, for a specific case, the known general function can depend on one or several constant parameters, whose exact values cannot be determined precisely. The question then arises how such classes of systems can be controlled to behave in a desired fashion, and whether in doing so, it is possible to gain an accurate estimate of the values of the underlying unknown parameters. The field of adaptive control and estimation has addressed these issues. Currently, many powerful techniques have been developed for the aforementioned problems (for example, see [19, 9]). In all of these results, the common feature is a fundamental assumption that the unknown parameters in the system occur linearly. Furthermore, this assumption is required to hold for both linear and nonlinear systems (see [19, 27, 14]).

The requirement for linear parametrization constrains the applicability of adaptive control, since many of the dynamical systems in nature exhibit such behavior which can only be accurately captured and represented by nonlinearly parametrized models. These nonlinear models can, perhaps, be converted to linearly parametrized ones by a suitable transformation. However, deriving such a transformation can be a nontrivial task, and may introduce further inaccuracies into the model. Hence, in order to accurately model complex systems, nonlinear parametrization seems unavoidable.

This chapter examines a recently developed algorithm [1, 16, 26] for control of a class of nonlinearly parametrized systems. The chapter is restricted only to the discussion of systems where the parameterization is convex or concave, and a single unknown parameter is present. An outline of the algorithm is given and proof of global stability summarized in section 2.2.2 The proof establishes the global convergence of the tracking error. Since the convergence of the tracking error does not imply the convergence of the parameter error, a separate analysis of the behavior of the algorithm with respect to the parameter error is warranted. The main results of the chapter are presented in sections 2.3 and sec 2.4. The former carries out a phase plane analysis of the min-max algorithm, while the latter one states the necessary conditions for accurate parameter estimation, similar to the persistent excitation conditions for linearly parameterized systems. Finally, the discussion and obtained results are illustrated by several numerical simulations in section 2.5.

2.2 The min-max adaptive algorithm

2.2.1 Properties of convex/concave functions

Many of the methods in adaptive control have their roots in the area of functional minimization. Namely, the adaptive control task of obtaining the correct parameter values can be viewed as a problem of finding the minimum of a certain cost function. The cost function is defined in terms of the difference of the observed system behavior and the behavior predicted from the current estimates of the unknown parameters. Therefore, the cost function is dependent upon the parameter estimates, and the process of its minimization is a search for such values of the parameter estimates which best fit the observed behavior. If the system is linearly parametrized, then the cost function is quadratic in the parameter error, and hence one is guaranteed of reaching the global minimum by calculating the gradient of the cost function with respect to the parameter error at our current parameter estimate, and then moving in the negative direction of the gradient. If the system is nonlinearly parametrized, the cost function no longer is quadratic in the parameters, and hence it is questionable whether the use of the local gradient will result in parameter convergence and overall closed-loop stability of the system and the adaptive controller. In this section, a summary of the recently developed min-max adaptive strategy is presented. Only the case of convex/concave nonlinear scalar parametrization is given here. The restriction to the scalar case is made for the sake of simplicity in analyzing the behavior of the parameter error of the min-max adaptive controller. For the controller in the case of a general nonlinearity see [16].

The min-max algorithm has two distinctive properties. Realizing that the parameter update along the local gradient can no longer guarantee global stability in nonlinear systems, the algorithm introduces a sensitivity function for determining parameter updates. This sensitivity function is equal to the local gradient only at certain instances. The second distinct feature of the algorithm is the use of a tuning function in the adaptive control input. These two functions are computed on-line by the algorithm based on available system measurements and a-priori knowledge of the convexity or concavity of the system parametrization. The algorithm ensures overall global stabilization and tracking to within a desired precision ϵ.

Since convexity/concavity properties of a function play an important role in the min-max adaptive controller, they are stated in the following definition.

Definition 2.1 *Given a set Θ, a function $f(\theta)$ is (i) convex on Θ if it satisfies the inequality*

$$f(\lambda\theta_1 + (1-\lambda)\theta_2) \leq \lambda f(\theta_1) + (1-\lambda)f(\theta_2) \qquad \forall \theta_1, \theta_2 \in \Theta \tag{2.1}$$

and (ii) concave on Θ if it satisfies the inequality

$$f(\lambda\theta_1 + (1-\lambda)\theta_2) \geq \lambda f(\theta_1) + (1-\lambda)f(\theta_2) \qquad \forall \theta_1, \theta_2 \in \Theta \tag{2.2}$$

where $0 \le \lambda \le 1$.

Convex/concave functions have specific geometric properties, which can be easily derived from the above definitions. Letting $\theta_3 = \lambda\theta_1 + (1-\lambda)\theta_2 = \theta_2 - \lambda(\theta_2 - \theta_1)$, it follows that $\theta_3 \in [\theta_1, \theta_2]$, since $0 \le \lambda \le 1$. Thus, the left hand side of the above inequalities represents the value of the function $f(\theta)$ at some point θ_3 on the interval bounded by θ_1 and θ_2. In the same manner, the right hand side of the above inequalities can be rewritten as $f(\theta_2) - \lambda(f(\theta_2) - f(\theta_1))$. Clearly, this represents the value of a linear function $g(\theta)$ defined by points $(\theta_1, f(\theta_1))$ and $(\theta_2, f(\theta_2))$ at at $\theta = \theta_3$. Thus, a function f is convex on a bounded set if on the set it lies below the line which connects the value of the function at the set end-points. Conversely, a function f is convex if it lies above the line connecting its values at the set end-points.

Another important property of these functions is their relation with respect to the gradient. When $f(\theta)$ is convex on Θ, then it can be shown that

$$f(\theta) - f(\theta_0) \ge \nabla f_{\theta_0}(\theta - \theta_0) \qquad \forall \theta, \theta_0 \in \Theta \tag{2.3}$$

and when $f(\theta)$ is concave on Θ, then

$$f(\theta) - f(\theta_0) \le \nabla f_{\theta_0}(\theta - \theta_0) \qquad \forall \theta, \theta_0 \in \Theta \tag{2.4}$$

where $\nabla f_{\theta_0} = \frac{\partial f}{\partial \theta}|_{\theta_0}$.

Before the min-max controller is stated, the following definition and lemma are required

Definition 2.2 *The saturation function, sat(.), is defined as*

$$sat(y) = \begin{cases} 1 & y \ge 1 \\ y & |y| < 1 \\ -1 & y \le -1 \end{cases}$$

Lemma 2.1 *Let Θ be a compact set specified by $\Theta = \left[\underline{\theta}, \overline{\theta}\right]$. For a given $\widehat{\theta} \in \Theta$, let*

$$J(\omega, \theta) = \beta\left[f(\phi, \widehat{\theta}) - f(\phi, \theta) - \omega(\widehat{\theta} - \theta)\right] \tag{2.5}$$

$$a_0 = \min_{\omega \in \mathbb{R}} \max_{\theta \in \Theta} J(\omega, \theta) \tag{2.6}$$

$$\omega_0 = arg \min_{\omega \in \mathbb{R}} \max_{\theta \in \Theta} J(\omega, \theta) \tag{2.7}$$

where β and ϕ are known quantities independent of θ, and f is either convex or concave in θ. Then

$$a_0 = \begin{cases} 0 & \text{if } \beta f \text{ is convex on } \Theta \\ \beta\left[\widehat{f} - f_{\min} - \dfrac{f_{\max} - f_{\min}}{\overline{\theta} - \underline{\theta}}(\widehat{\theta} - \underline{\theta})\right] & \text{if } \beta f \text{ is concave on } \Theta \end{cases} \tag{2.8}$$

$$\omega_0 = \begin{cases} \nabla f_{\widehat{\theta}} & \text{if } \beta f \text{ is convex on } \Theta \\ \dfrac{f_{\max} - f_{\min}}{\overline{\theta} - \underline{\theta}} & \text{if } \beta f \text{ is concave on } \Theta \end{cases} \tag{2.9}$$

where $\widehat{f} = f(\phi, \widehat{\theta})$, $f_{\max} = f(\phi, \overline{\theta})$, and $f_{\min} = f(\phi, \underline{\theta})$.

Proof : See [1]. ■

Based on Definition 2.1 and eq. (2.6), the range of possible values of a_0 is examined. In particular, the possible values of a_0 in the case when the product βf is a concave function is investigated, since in the converse case of convex βf a_0 is set to zero. First, the possible combinations of properties of β and f which give a concave βf are noted. They are: (a) positive β and concave f and (b) negative β and convex f. Second, the expression in parenthesis is rewritten as: $\widehat{f} - [f_{min} - \lambda(f_{min} - f_{max})]$, with $\lambda = 1 - \dfrac{\widehat{\theta} - \underline{\theta}}{\overline{\theta} - \underline{\theta}}$. Since $\widehat{\theta} \in [\underline{\theta}, \overline{\theta}]$, then $0 \leq \lambda \leq 1$. In case (a) f is concave, and it follows from Definition 2.1 that the expression in the parenthesis is positive. In this case β is positive, and thus a_0 is positive as well. In case (b) because f is convex, the expression in the parenthesis is negative. Since in this case β is also negative, a_0 again obtains a positive value. Thus it can be concluded that, whether f is convex/concave,

$$a_0 \geq 0 \qquad\qquad \forall\, \beta. \tag{2.10}$$

Another crucial relationship between the values of f, a_0 and ω_0 is stated in the following lemma.

Lemma 2.2 *Let α, ϵ be arbitrary positive quantities, and let $\alpha_{max} \geq \alpha$. For a given $\hat{\theta} \in \Theta$, and all $\theta \in \Theta$, let a_0 and ω_0 be chosen as in eqs. (2.6)- (2.7) with $\beta = \alpha_{max} sign(x)$. The following is then true, whether f is concave or convex:*

$$x\left\{\alpha\left[\hat{f} - f - (\hat{\theta} - \theta)\omega_0\right] - a_0 sat\left(\frac{x}{\epsilon}\right)\right\} \leq 0 \qquad \forall |x| \geq \epsilon \tag{2.11}$$

Proof: From eq. 2.10 we have that $a_0 \geq 0$. Then, from eq. (2.6) and from the choice of β, it follows that

$$a_0 \geq \alpha_{max} sign(x)\left[\hat{f} - f - \omega_0(\hat{\theta} - \theta)\right] \geq \alpha sign(x)\left[\hat{f} - f - \omega_0(\hat{\theta} - \theta)\right] \tag{2.12}$$

Therefore, for $x > \epsilon$, by multiplying both sides of eq. 2.12 by $sat\left(\frac{x}{\epsilon}\right) = sign(x) = 1$ it follows that

$$a_0 sat\left(\frac{x}{\epsilon}\right) \geq \alpha\left[\hat{f} - f - \omega_0(\hat{\theta} - \theta)\right] \tag{2.13}$$

and hence eq. (2.11) holds. For $x < -\epsilon$, $sat\left(\frac{x}{\epsilon}\right) = sign(x) = -1$. Thus, in this case

$$a_0 sat\left(\frac{x}{\epsilon}\right) \leq \alpha\left[\hat{f} - f - \omega_0(\hat{\theta} - \theta)\right] \tag{2.14}$$

and hence, again, eq. (2.11) holds. •

2.2.2 The min-max globally stable estimator

Having defined in the previous section the necessary mathematical preliminaries, this section gives a summary of the recently developed min-max estimator strategy (see

[1, 25]). The min-max estimator is developed for the following class of dynamical system models:

$$\dot{y} = f(\phi(t), \theta) \tag{2.15}$$

where y is the measurable system output, $\theta \in \mathbb{R}$ is an unknown parameter, ϕ is a scalar function of time and f is a scalar function that is nonlinear both in ϕ and θ.

Further developments are based on the following assumptions about the above system:

(A1) $\theta \in \Theta$, where $\Theta = [\underline{\theta}, \overline{\theta}]$ is a known compact set defined by its lower ($\underline{\theta}$) and upper ($\overline{\theta}$) bounds.

(A2) $\phi(t)$ is a measurable and bounded function of time.

(A3) For any $\phi(t)$, only one of the following is true

(i) f is concave for all $\theta \in \Theta$

(ii) f is convex for all $\theta \in \Theta$

(A5) f is a known smooth and bounded function of its arguments.

The goal of the estimator is to closely track the output of the system and, in doing so, provide estimates of the value of the unknown parameter θ. To accomplish that task, the following estimator has been proposed:

$$\dot{\hat{y}} = -ke_{\varepsilon} + f(\phi, \hat{\theta}) - a\,sat\left(\frac{e_c}{\epsilon}\right) \tag{2.16}$$

$$\dot{\hat{\theta}} = -e_{\varepsilon}\omega \tag{2.17}$$

$$a = \min_{\omega \in \mathbb{R}} \max_{\theta \in \Theta} sign(e_c)\left[f(\phi(t), \hat{\theta}) - f(\phi(t), \theta) - \omega(\hat{\theta} - \theta)\right] \tag{2.18}$$

$$\omega = arg \min_{\omega \in \mathbb{R}} \max_{\theta \in \Theta} sign(e_c)\left[f(\phi(t), \hat{\theta}) - f(\phi(t), \theta) - \omega(\hat{\theta} - \theta)\right] \tag{2.19}$$

$$e_{\varepsilon} = e_c - \epsilon\,sat\left(\frac{e_c}{\epsilon}\right) \tag{2.20}$$

where e_c is the tracking error defined as $e_c = \widehat{y} - y$, ϵ is an arbitrary positive constant and $\widehat{\theta}$ represents an estimate of θ. The stability feature of this estimator is stated in the following theorem.

Theorem 2.1 *For the system in eq. (2.15), under the assumptions (A1)- (A4), the estimator given in eqs. (2.16)- (2.19), assures that all signals in the closed-loop system are globally bounded and that* $|e_c(t)| \to \epsilon$ *as* $t \to \infty$, *provided that* $\widehat{\theta}(t) \in \Theta$ *for* $t \geq t_0$.

Proof: By subtracting eq. (2.15) from eq. (2.16), the following closed-loop dynamics are obtained:

$$\dot{e}_c = -ke_\varepsilon + f(\phi, \widehat{\theta}) - f(\phi, \theta) - asat\left(\frac{e_c}{\epsilon}\right) \tag{2.21}$$

$$\dot{\widetilde{\theta}} = -e_\varepsilon \omega \tag{2.22}$$

where $\widetilde{\theta} = \widehat{\theta} - \theta$. Adding and subtracting the term $\widetilde{\theta}\omega$ to the right- hand side of eq. (2.21) yields the overall adaptive system dynamics as:

$$\dot{e}_c = -ke_\varepsilon + \widetilde{\theta}\omega + \left[f(\phi, \widehat{\theta}) - f(\phi, \theta) - asat\left(\frac{e_c}{\epsilon}\right) - \widetilde{\theta}\omega\right] \tag{2.23}$$

$$\dot{\widetilde{\theta}} = -e_\varepsilon \omega \tag{2.24}$$

This structure is reminiscent of the structure of linear adaptive systems, with the only difference being the presence of $\left[f(\phi, \widehat{\theta}) - f(\phi, \theta) - asat\left(\frac{e_c}{\epsilon}\right) - \widetilde{\theta}\omega\right]$ nonlinear term in eq. (2.23). If this term was absent, the resulting system would be identical to a linear adaptive system with an adaptation deadzone, for which stability and convergence results are well established. Thus, it must be shown that the extra term does not have a destabilizing effect on the whole system. In particular, since the nonlinear term appears only in the first equation of the system, its influence on the dynamics of e_c needs to be examined. This is done by noting the choice of a and ω and applying Lemma 2.2 to obtain that

$$e_c\left[f(\phi, \widehat{\theta}) - f(\phi, \theta) - asat\left(\frac{e_c}{\epsilon}\right) - \widetilde{\theta}\omega\right] \leq 0.$$

Thus, the nonlinear term has a stabilizing effect on e_c, and hence the overall adaptive system is stable. This conclusion can be further verified by choosing a Lyapunov function as $V = \frac{1}{2}(e_\varepsilon^2 + \tilde{\theta}^2)$. In a straightforward manner it can be confirmed that the time derivative of V along the system trajectories is indeed nonpositive, indicating system stability. This implies that all of the system signals are bounded, and hence Barbalat's Lemma can be invoked to show that $|e_c| \to \epsilon$ as $t \to \infty$. •

The stability of the proposed estimator is based on the assumption that both θ and $\widehat{\theta}(t)$ are in the set Θ during the course of estimation. In order to ensure that $\widehat{\theta}(t) \in \Theta$ for all $t \geq t_0$, $\widehat{\theta}(t_0)$ can be chosen to be in Θ, and the update of $\widehat{\theta}$ can be turned off whenever it leaves Θ. Alternatively, the following projection strategy suggested in [4] can be coupled with the presented parameter estimation laws:

$$\begin{aligned} \dot{\theta}_c &= -e_\varepsilon \omega - \sigma(\theta_c - \widehat{\theta}) \qquad \sigma > 0 \\ \widehat{\theta} &= \begin{cases} \theta_c & \theta_c \in \Theta \\ \overline{\theta} & \theta_c > \overline{\theta} \\ \underline{\theta} & \theta_c < \underline{\theta}. \end{cases} \end{aligned} \tag{2.25}$$

By replacing eq. (2.17) with eq. (2.25) global boundedness is ensured if $\widehat{\theta}(t_0) \in \Theta$.

2.3 Phase-plane analysis

In the previous section, the stability and convergence of the tracking error was established. In the linear scalar parameter case, this was sufficient to establish the convergence of the parameter error as well. Based on demonstrated similarities between the linear adaptive estimator and the min-max estimator, initially it can be hypothesized that the same sufficiency holds for the min-max case with nonlinear parametrization. This section examines the validity of such a hypothesis by examining the behavior of the min-max estimator in the phase plane.

For the sake of simplicity, the following dynamical system is considered:

$$\dot{y} = e^{-\phi(t)\theta} \tag{2.26}$$

Without loss of generality, we assume that the set $\Theta = [\underline{\theta}, \overline{\theta}]$, $\theta \in \Theta$, is a subset of the positive real axis.

Definition 2.1 states that the function $f(\phi(t), \theta) = e^{-\phi(t)\theta}$ is a convex function in θ for all $\phi(t)$. This implies that the required structure of the min-max estimator is as follows:

$$\dot{\hat{y}} = -ke_\varepsilon + f(\phi, \hat{\theta}) - asat\left(\frac{e_c}{\epsilon}\right) \tag{2.27}$$

$$\dot{\hat{\theta}} = -e_\varepsilon \omega \tag{2.28}$$

$$a = \begin{cases} 0 & e_c \geq 0 \\ -\left[\hat{f} - f_{min} - \omega_s(\hat{\theta} - \underline{\theta})\right] & e_c < 0 \end{cases} \tag{2.29}$$

$$\omega = \begin{cases} -u(t)e^{-u(t)\hat{\theta}} & e_c \geq 0 \\ \omega_s & e_c < 0 \end{cases} \tag{2.30}$$

where ω_s is the global slope defined by

$$\omega_s = \frac{f_{max} - f_{min}}{\overline{\theta} - \underline{\theta}}. \tag{2.31}$$

From eqs. (2.26)- (2.30), the closed-loop model and estimator error dynamics are derived as:

$$\begin{aligned} \dot{e}_c &= -ke_\varepsilon + \tilde{f} \\ \dot{\tilde{\theta}} &= -e_\varepsilon\left(-\phi(t)e^{-\phi(t)\hat{\theta}}\right) \end{aligned} \qquad e_c > 0 \tag{2.32}$$

$$\begin{aligned} \dot{e}_c &= -\tfrac{e_c}{\epsilon}\left[f_{min} + \omega_s(\hat{\theta} - \underline{\theta}) - \hat{f}\right] + \tilde{f} \\ \dot{\tilde{\theta}} &= 0 \end{aligned} \qquad 0 \geq e_c \geq -\epsilon \tag{2.33}$$

$$\begin{aligned} \dot{e_c} &= -ke_\varepsilon + f_{min} + \omega_s(\widehat{\theta} - \underline{\theta}) - f \\ \dot{\tilde{\theta}} &= -e_\varepsilon \omega_s \end{aligned} \qquad e_c < -\epsilon \tag{2.34}$$

Before proceeding to analyze the behavior of the above closed loop system in the phase plane defined by the states e_c and $\tilde{\theta}$, another characteristic of the nonlinearity $f = e^{-\phi(t)\theta}$ is observed. As stated above, f retains its convexity with respect to θ for all values of ϕ. However, the sign of ϕ determines another important property of f. In the case that ϕ is positive, f is a decreasing convex function with respect to $\theta \in \Theta$. In the case that ϕ is negative, f is an increasing convex function. This characteristic is important in the system analysis, since it correlates the sign of $\tilde{f} = \widehat{f} - f$ and the sign of $\tilde{\theta}$. In the case of an increasing function, the signs of the two terms are the same. In the case of a decreasing function, the signs are reversed. For that reason, the system analysis is partitioned into two separate cases depending on the sign of ϕ. The first case to be considered is the case when $\phi(t) > 0$.

The behavior of the system in eqs. (2.32)- (2.34), is now examined in three distinct regions of the phase-plane $(e_c, \tilde{\theta})$ where different type of system motion can occur. These regions are defined as: (a) $e_c \geq \epsilon$, (b) $|e_c| < \epsilon$, and (c) $|e_c| \leq -\epsilon$. The reason for choosing these three regions is that the first and the last differ in the way of calculating a, while in the second region differs from the other two because $|e_c| < \epsilon$ implies that $e_\varepsilon = 0$. For the choice of f as in eq. (2.26), the following observations and definitions about the system motion are noted:

case (a) $e_c \geq \epsilon$

(a-i) The gradient is negative, and thus $\omega < 0$. Since e_c is positive, in this case $\dot{\tilde{\theta}} > 0$.

(a-ii) Let $\tilde{\theta}_1 = -\frac{1}{u} \log\left(ke_\varepsilon e^{\phi\theta} + 1\right)$, and let the curve L_1 be the set

$$L_1 = \{(e_c, \tilde{\theta}) \mid e_c \geq \epsilon,\ \tilde{\theta} = \tilde{\theta}_1(e_c)\}. \tag{2.35}$$

From the choice of f and definition of $\dot{e_c}$ in eq. (2.32), it follows that for any point above the curve L_1, $\dot{e_c} < 0$. Conversely, $\dot{e_c} \geq 0$ for any point below the curve, with the equal sign applying to the points on the curve.

case (b) $e_c \leq -\epsilon$

(b-i) The global slope ω_s and e_c are negative, implying that $\dot{\tilde{\theta}} < 0$.

(b-ii) Let $\tilde{\theta}' = \frac{1}{\omega_s}(f - f_{min}) + \underline{\theta} - \theta$, and define $\tilde{\theta}_2 = \frac{k}{\omega_s} e_\varepsilon + \tilde{\theta}'$. Let the curve L_2 be the set

$$L_2 = \{(e_c, \tilde{\theta}) \mid e_c \leq -\epsilon,\ \tilde{\theta} = \tilde{\theta}_2(e_c)\}. \tag{2.36}$$

It then follows that for any point above the curve L_2, $\dot{e_c} < 0$. Conversely, $\dot{e_c} \geq 0$ for any point below the curve, with the equal sign applying to the points on the curve.

case (c) $|e_c| < \epsilon$

(c-i) $e_\varepsilon = 0$ and thus $\dot{\tilde{\theta}} = 0$. Hence, all system equilibrium points must lie in the set $|e_c| < \epsilon$.

(c-ii) Let

$$L_3 = \{(e_c, \tilde{\theta}) \mid 0 \leq e_c < \epsilon,\ \tilde{\theta} = 0)\}. \tag{2.37}$$

On this set, $\tilde{\theta} = 0$ implies $\tilde{f} = 0$. Thus, it can be concluded that the L_3 represents a set of equilibrium points.

(c-iii) Let $e_{c3} = \epsilon \dfrac{\tilde{f}}{f_{min} + \omega_s(\tilde{\theta} + \theta - \underline{\theta}) - \hat{f}}$. From eq. (2.33), it follows that $\dot{e_c} = 0$ when $e_c = e_{c3}$. For given values of $u(t)$ and θ, e_{c3} is a function of $\tilde{\theta}$. Hence, in order to find the set of possible equilibrium points, the range of values of $\tilde{\theta}$ for which $-\epsilon \leq e_{c3}(\tilde{\theta}) \leq 0$ must be calculated. From the definition of e_{c3}, it follows that

$$-\epsilon \leq e_{c3}(\tilde{\theta}) \leq 0 \Leftrightarrow -1 \leq \frac{\tilde{f}}{f_{min} + \omega_s(\tilde{\theta} + \theta - \underline{\theta}) - \hat{f}} \leq 0. \tag{2.38}$$

Since the quantity $f_{min} + \omega_s(\tilde{\theta} + \theta - \underline{\theta}) - \hat{f}$ is positive due to convexity of f, the inequality in eq. (2.38) can be rewritten as:

$$\hat{f} - f_{min} - \omega_s(\tilde{\theta} + \theta - \underline{\theta}) \leq \tilde{f} \leq 0. \tag{2.39}$$

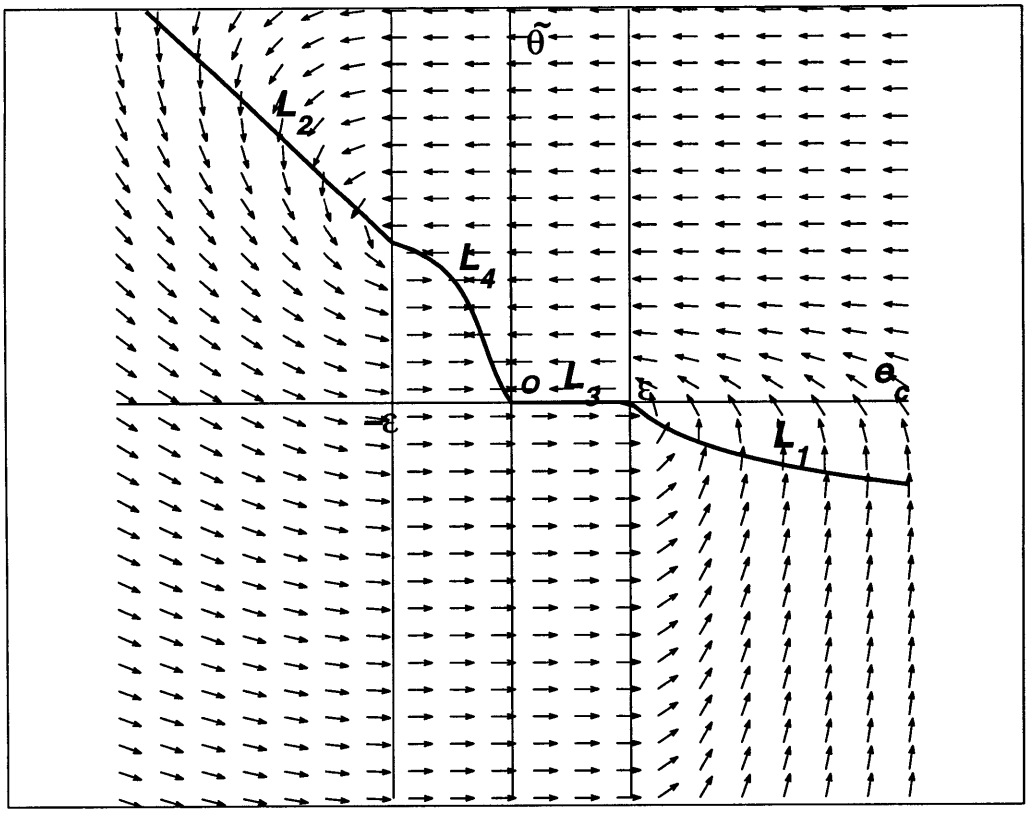

Figure 2-1: Phase-plane adaptive system motion for the case when $\phi > 0$. The arrows represent the direction of the velocity vector, and the thick line is used to represent sets L_1-L_4, respectively.

It can be verified that this inequality holds when $0 \leq \tilde{\theta} \leq \tilde{\theta}'$. Hence the set L_4

$$L_4 = \{(e_c, \tilde{\theta}) \mid e_c = e_{c3},\ 0 \leq \tilde{\theta} \leq \tilde{\theta}'\} \tag{2.40}$$

is a set of equilibrium points as well.

Based on these observations and definitions, the motion of the system in the phase-plane is shown graphically in Fig. (2-1).

The stability result of Theorem 2.1 establishes that the system in eqs. (2.32)-(2.34) tends toward the adaptation dead-zone defined by $|e_c| \leq \epsilon$. From Fig. 2-1 and above calculations it is observed that the set of equilibrium points L_e in the dead-zone consists of the sets L_3 and L_4, that is $L_e = L_3 \bigcup L_4$. On the set L_3, $\tilde{\theta}$ is zero, and thus the parameter estimate $\widehat{\theta}$ converges to the true value of the unknown parameter θ. However, on the set L_4, $\tilde{\theta} \neq 0$. This is due to a couple of factors. First, once the system enters the dead-zone, the adjustment of $\widehat{\theta}$ is turned off. Therefore,

$\dot{e}_c = 0$ suffices for the adaptive system to have an equilibrium point in the dead-zone even if $\tilde{\theta}$ is different from zero. Second, it is noted that eq. (2.33) represents a first order filter on the quantity $\tilde{f}$. From classic linear systems theory, it is known that first-order systems attain equilibrium values for non-zero inputs. Hence, for a range of values of $\tilde{f}$ different from zero the system will settle to an equilibrium point. The maximum value of $|\tilde{\theta}|$ for which the system will achieve an equilibrium is $\tilde{\theta}'$.

The case when $\phi(t) < 0$ is now considered. The only difference between this and the prior case is that the sign of the gradient is changed. The change of the gradient sign implies the reversal of the sign of $\dot{\tilde{\theta}}$. The change in the gradient further implies the change in the relation between the sign of $\tilde{f}$ and $\tilde{\theta}$. Practically, this has the same effect as changing the orientation of the $\tilde{\theta}$-axis in Fig. 2-1, as can be evidenced from Fig. 2-1 which depicts the phase-plane behavior of the system in this case. Curves which correspond to the sets L_1, L_2, L_3, and L_4 of the previous case are denoted by M_1, M_2, M_3, and M_4, respectively. Similarily, the equilibrium set M_e consists of $M_e = M_3 \bigcup M_4$.

By comparing Figs. (2-1) and (2-2), it can be observed that $L_3 = M_3$. More importantly, the equilibrium set E, when u switches between positive and negative values consists of the intersection of equilibrium sets for the two general cases, respectively:

$$E = L_e \bigcap M_e = (L_3 \bigcup L_4) \bigcap (M_3 \bigcup M_4) = L_3 \bigcup M_3 = L_3.$$

On the set L_3, $\tilde{\theta} = 0$, and thus parameter convergence can be obtained if u switches between positive and negative values. The requirements on the nonlinearity f and on the values of u which ensure parameter convergence is achieved are further investigated in the following section.

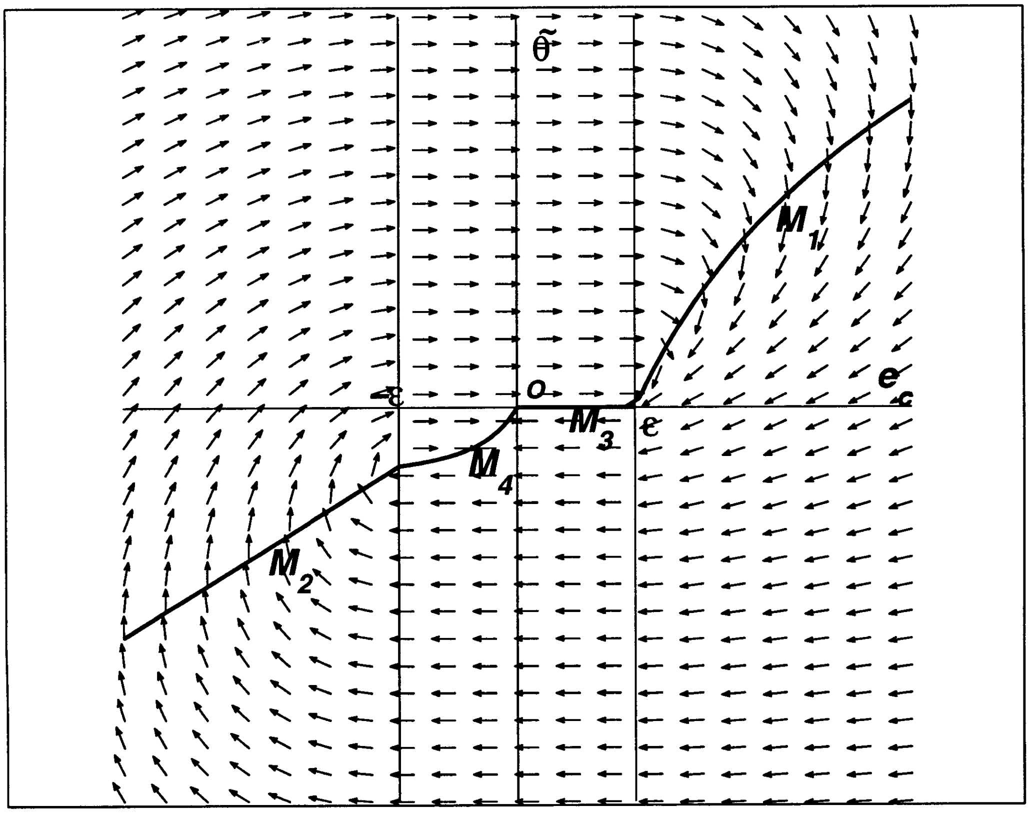

Figure 2-2: Phase-plane adaptive system motion for the case when $\phi < 0$. The arrows represent the direction of the velocity vector, and the thick line is used to represent sets L_1-L_4, respectively.

2.4 Parameter Convergence

The stability of the proposed adaptive algorithm was established in Section 2.2.2. As is known from linear adaptive control, stability and tracking do not necessarily imply the convergence of unknown parameters to their true values. This section investigates the problem of parameter convergence for the min-max adaptive algorithm. A sufficient condition for ensuring parameter convergence is derived.

2.4.1 Preliminaries

For the sake of clarity, we restate the adaptive system dynamics below:

$$\begin{aligned} \dot{e}_c &= -ke_\varepsilon - a sat\left(\frac{e_c}{\epsilon}\right) + \tilde{f} \\ \dot{\tilde{\theta}} &= e_\varepsilon \, \omega \end{aligned} \quad (2.41)$$

where $\tilde{f} = f(\phi(t), \widehat{\theta}) - f(\phi(t), \theta)$, and the quantities e_ε, a, ω were defined in eqs. (2.18)-(2.20). By inspecting the above system, it follows that all its equilibrium points are contained within the region where $e_\varepsilon = 0$. In this region, $|e_c| \leq \epsilon$. The system dynamics in this region are:

$$\begin{aligned} \dot{e}_c &= -a\,\frac{e_c}{\epsilon} + \tilde{f} \\ \dot{\tilde{\theta}} &= 0 \end{aligned} \tag{2.42}$$

When $a = 0$, the only equilibrium set is reduced to the case when $\tilde{f} = 0$, implying satisfactory parameter convergence. In the case when $a \neq 0$, eq. (2.42) implies the existence of equilibrium points where $\tilde{f} \neq 0$, so that parameter convergence is not achieved. From eq. (2.18) and Lemma 2.1 $a \neq 0$ when the function $\beta\, f = sign(e_c)\, f$ is concave. For a given function f, let β^* denote the value of $sign(e_c)$ when $a \neq 0$. Thus, the allowed values for β^* are *(i)* $\beta^* = -1$ if f is convex, and *(ii)* $\beta^* = 1$ when f is concave. The solutions of the differential equation in eq. (2.42) are such that e_c will always tend towards $e_{c_{ss}}$, where $e_{c_{ss}}$ is given by:

$$e_{c_{ss}} = \epsilon\,\frac{\tilde{f}}{a} \tag{2.43}$$

Eq. (2.42) is only valid while both $sign(e_c) = \beta^*$ and $|e_c| \leq \epsilon$ hold. The first condition implies that $\beta^*\, sign(e_c) > 0$. It has been established that $a > 0$, and since $\epsilon > 0$, the first condition translates into:

$$\beta^*\, \tilde{f} > 0 \tag{2.44}$$

Meanwhile, the second condition is satisfied if $|\tilde{f}|/|a| \leq 1$. Using eq. (2.44) it is obtained that

$$1 \geq \frac{|\,\tilde{f}\,|}{|\,a\,|} = \frac{|\,\beta^*\, \tilde{f}\,|}{a} = \frac{\beta^*\, \tilde{f}}{a} \tag{2.45}$$

The two conditions can be combined as:

$$0 < \beta^* \tilde{f} \leq a \tag{2.46}$$

Hence, as long as eq. (2.46) holds, the system has an equilibrium point for which $\tilde{f} \neq 0$.

To assist with further study of parameter convergence, several new concepts and definitions are introduced. In standard adaptive control literature (see [19]), the concept of attractiveness of a single point, usually the origin, is defined. The same idea of attractiveness, when applied to a set which contains at least one point leads to the inauguration of the concept of uniform attractiveness. That concept is stated in the following definition.

Definition 2.3 *A set $B_\rho(x) = \{x \mid \|x\| < \rho\}$ is uniformly attractive (u.a.t.t.) if for some $\epsilon_1 > 0$ and every $\epsilon_2 > \rho$ there exists a $T(\epsilon_1, \epsilon_2) > 0$ such that if $\|x(t_0)\| < \epsilon_1$, then $\|x(t; x_0, t_0)\| < \epsilon_2$ for all $t \geq t_0 + T$.*

Several other quantities are now specified. When referring to a time interval, it is understood that the time interval in question is contiguous. That is, let T be a time interval specified by lower and upper bounds, t_{min} and t_{max}, respectively. The notation $T = [t_{min}, t_{max}]$ will be used for representing this. The length of an interval is given by $t_{max} - t_{min}$. If for a specific t, $t_{min} \leq t \leq t_{max}$ is true, then $t \in T$. Similarily, an unbounded interval is one which has a lower bound only. For example, let T_u be an unbounded interval specified by the lower bound t_{min}. Then all t such that $t \geq t_{min}$ are elements of T_u.

For the motion of the system in Eq. (2.41), define the set Ω_D as the set of all time such that:

$$\Omega_D = \{ \; t \mid \|e_c\| \; \leq \epsilon \}. \tag{2.47}$$

Thus, Ω_D represents the time that the systems spends in the "dead-zone" in which the parameter update signal is turned off. The set Ω_D can be empty, or can consist of several distinct intervals. Hence, Ω_D can be expressed as $\Omega_D = \bigcup_i T_{D_i}$ where

$T_{D_i} = [\underline{t_{D_i}}, \overline{t_{D_i}}]$ is a sequence of intervals, with $T_{D_i} \neq T_{D_j}$ if $i \neq j$. Each of the intervals T_{D_i} is such that *(a)* if $\underline{t_{D_i}} < t < \overline{t_{D_i}}$ then $\|e_c(t)\| < \epsilon$ and *(b)* $\|e_c(t)\| = \epsilon$ at $t = \underline{t_{D_i}}, \overline{t_{D_i}}$.

The set Ω_D can consist of two subsets Ω_D^L and Ω_D^R defined as:

$$\Omega_D^L = \{\ t \mid |e_c| \leq \epsilon \text{ and } sign(e_c)\, f(t) \text{ is concave}\} \tag{2.48}$$

$$\Omega_D^R = \{\ t \mid |e_c| \leq \epsilon \text{ and } sign(e_c)\, f(t) \text{ is convex }\} \tag{2.49}$$

In an attempt to justify the notation in the previous equations, it is noted that for the case of a convex f, if $t \in \Omega_D^L$ then $e_c < 0$, indicating that the system is in the left half of the phase plane. Likewise, if $t \in \Omega_D^R$ then the system is in the right hand plane.

Similarly, each T_{D_i} can consist of two subsets $T_{D_i}^L$ and $T_{D_i}^R$:

$$T_{D_i}^L = T_{D_i} \bigcap \Omega_D^L \ \ ; \ \ T_{D_i}^R = T_{D_i} \bigcap \Omega_D^R \tag{2.50}$$

Next, let the set Ω_P represent the time for which the system is outside the dead-zone. Thus, $\Omega_P = \Omega_D^C$. In an analogous manner to T_{D_i}, the intervals T_{P_i} can be defined as: $T_{P_i} = (\underline{t_{P_i}}, \overline{t_{P_i}})$ where $\|e_c(t)\| > \epsilon \ \ \forall t \in T_{P_i}$. Hence, $\Omega_P = \bigcup_i T_{P_i}$.

Two properties of the function $f(\phi(t), \theta)$ that are useful in establishing parameter convergence are now defined. In stating these properties it is useful to note that for a specific function $f(\phi, \theta)$, β^* as defined above is a function of the values of ϕ, since there exist functions which can switch between convexity and concavity in θ depending on the value of ϕ. Therefore, for two distinct $\phi_1 \neq \phi_2$ let the corresponding β^* be denoted by β_1^* and β_2^*, respectively. Forthwith, the properties can be stated.

(P1) The set of values of ϕ for which, given $\epsilon_\theta > 0$, there doesn't exist an ϵ_F such that if $\|\tilde{\theta}\| > \epsilon_\theta$ then $\|\tilde{f}(\phi(t), \tilde{\theta})\| > \epsilon_F$ is finite. The set of all possible values of $\phi \in \mathbb{R}$ is infinite.

(P2) A function $f(\phi(t), \theta)$ satisfies the property (P1) if, for a given $\phi_1 \in \mathbb{R}$, any

$\theta_1 \in \Theta$ and all $\theta_2 \in \Theta$, $\theta_2 \neq \theta_1$, such that

$$0 \leq \beta_1^* \left(f(\phi_1, \theta_2) - f(\phi_1, \theta_1)\right)$$

there exists $\phi_2 \neq \phi_1$ and $\epsilon_F > 0$ such that

$$\beta_2^* \left(f(\phi_1, \theta_2) - f(\phi_1, \theta_1)\right) \leq -\epsilon_F$$

Property (P1) is an identifiability property which states that f and ϕ must be such that a difference in θ arguments of f must produce an observable difference in the function values. If (P1) is not satisfied, then the unknown parameter cannot be identified since other parameter values provide the same function output. Specifically, (P1) states that the number of possible values of ϕ when the parameter is unidentifiable is small compared to all the possible values of ϕ. In many practical functions of interest, the set referred to by (P1) contains only one element. The property (P1) is a fundamental property, and it can be seen that property (P2) implicitly implies that (P1) is satisfied.

Property (P2) states the conditions on f so that an $\phi(t)$ can be chosen which will enable the system to exit from the adaptation deadzone. Since the system will stay in the deadzone as long as eq. (2.46) holds. Examining that inequality, it can be stipulated that the only degree of freedom is to attempt to change the quantity $\beta^* \tilde{f}$, since a is an inherent feature of the min-max algorithm and the present nonlinearity. There are four possible courses of action to take in order to change the value of $\beta^* \tilde{f}$. Using the notation of (P2), and letting $\tilde{f}(\phi) = f(\phi, \theta_2) - f(\phi, \theta_1)$, the four options are to choose ϕ_2 such that

(C1) the sign of $\tilde{f}$ is reversed, while preserving the convexity/concavity of f. Thus,

$$sign\left(\tilde{f}(\phi_2)\right) \neq sign\left(\tilde{f}(\phi_1)\right) \text{ and } \beta_1^* = \beta_2^*.$$

(C2) the convexity/concavity is reversed, while preserving the sign of $\tilde{f}$.

$$sign\left(\tilde{f}(\phi_2)\right) = sign\left(\tilde{f}(\phi_1)\right) \text{ and } \beta_1^* \neq \beta_2^*.$$

(C3) both convexity/concavity and the sign of $\tilde{f}$ are retained, and $\tilde{f}$ is increased.

$$sign\left(\tilde{f}(\phi_2)\right) = sign\left(\tilde{f}(\phi_1)\right) \text{ and } \beta_1^* = \beta_2^*.$$

(C4) both convexity/concavity and the sign of $\tilde{f}$ are reversed.

$$sign\left(\tilde{f}(\phi_2)\right) \neq sign\left(\tilde{f}(\phi_1)\right) \text{ and } \beta_1^* \neq \beta_2^*.$$

From the aspect of the problem of parameter convergence, the worthwhile choices are only *(C1)* and *(C2)*. It is these two choices that are allowed when the property (P2) is satisfied. Since, by assumption, f and $\phi(t)$ are bounded, there is a limit as to how much $\tilde{f}$ can be increased. By the mean value theorem (see [23]), it can be written that $\tilde{f} = B(\theta_2 - \theta_1)$, where B is the value of gradient of f with respect to θ at some point between θ_2 and θ_1. Since B is bounded, there is a lower limit on $\theta_2 - \theta_1$ which will produce an $\tilde{f}$ such that eq. (2.46) holds. The existence of this lower limit is undesirable in establishing parameter convergence. It is easy to verify that choice *(C4)* reduces to *(C3)*, and thus is also undesirable. Whether a given function f satisfies (P2) can be determined by an analytical examination of the specific function. For LP-systems, since f is linear in θ, (P1) is satisfied by changing the sign of ϕ.

2.4.2 A Condition for Parameter Convergence

Before stating the sufficient conditions for parameter convergence, the following notation is introduced. Let

$$\tilde{f}(x,y) = f\left(\phi(x), \widehat{\theta}(y)\right) - f\left(\phi(x), \theta(y)\right).$$

Let $\tilde{f}(x)$ be a short form of representing $\tilde{f}(x, x)$. Let

$$d_c(t, \theta_1) = f\left(\phi(t), \underline{\theta}\right) + \omega_s\left(\theta_1 - \underline{\theta}\right) - f\left(\phi(t), \theta_1\right),$$

where ω_s is the global slope defined in eq. (2.31).

The sufficient conditions for parameter convergence are stated in the following theorem.

Theorem 2.2 *Let ϵ, ϵ_θ and $t_0 > 0$ be given, and let f, $\phi(t)$, $t \geq t_0$ satisfy properties (P1) and (P2). For the system in Eq. (2.41), the set $B_\rho(x)$, $\rho = \sqrt{\epsilon^2 + \epsilon_\theta^2}$, is uniformly attractive if there exists a δ_0 and η, δ_0, $\eta > 0$ such that*

$$\left\| \int_{\underline{t}}^{\underline{t}+\delta_0} \tilde{f}(\tau, t)\, d\tau \right\| \geq \eta \quad \textit{and} \quad \left\| \int_{\underline{t}}^{\underline{t}+\delta_0} d_c(\tau, \theta) + \omega_s(\tau)\, \tilde{\theta}(t)\, d\tau \right\| \geq \eta \qquad (2.51)$$

Proof: The proof consists of two parts. In the first part, it will be shown that the amount of time that the systems spends in the deadzone introduced by the adaptive algorithm is finite if the parameter error is large. Having shown this, the proof then follows a the proof of a similar theorem for linear systems given in [18, 17, 19].

The first part of the proof states that in the deadzone, the system will have a large velocity in the e_c direction in the phase plane, provided that the parameter estimate is far away from the true value. This implies that the system will exit the deadzone in finite time with a large velocity. Since the system exited the deadzone with a large velocity, $\|e_\varepsilon\|$ must be large if the system is to turn around and head back towards the deadzone. The magnitude of decrease of the Lyapunov function, and thus the decrease in the magnitude of the state vector, is directly proportional to $\|e_\varepsilon\|$. Hence, the Lyapunov function will decrease by a finite amount when the system is not n the deadzone, thus establishing *uatt* of a given ball around the phase-plane origin is shown.

Let x denote the state vector, $x \triangleq [e_c, \tilde{\theta}]^T$. Without loss of generality, it is assumed that $\|x(t_0)\| > \rho$. For if $\|x(t_0)\| \leq \rho$, Theorem 2.1 implies that that $\|x(t)\| \leq \rho \;\; \forall\, t \geq t_0$, and hence Theorem 2.2 follows.

A Lyapunov function for the system in eq. (2.41) is chosen as $V = \frac{1}{2}(e_\varepsilon^2 + \tilde{\theta}^2)$. Then, $\dot{V}$ along the system trajectories is

$$\dot{V} \;=\; -k\, e_\varepsilon^2 + e_\varepsilon \left[\tilde{f} + \tilde{\theta}\,\omega - a\,sat\left(\frac{e_c}{\epsilon}\right)\right] \leq -k\, e_\varepsilon^2 \tag{2.52}$$

It will be established that over every interval $[t, t+T_0]$, $V(t)$ decreases by a nonzero value. Since $\dot{V}\,(t) = 0 \;\; \forall\, t \in \Omega_D$, the behavior of the trajectories in Ω_D (Ω_D^L and Ω_D^R) is examined. For $t \in \Omega_D$, the system dynamics are given by eq. (2.43). It is noted that for $t \in \Omega_D$ the assumption $\|x(t_0)\| > \rho$ implies that $\|\tilde{\theta}\| \geq \epsilon_\theta$, since $\|e_c(t)\| \leq \epsilon$. By the property (P1) it further implies that $\|\tilde{f}\| \geq \epsilon_F$.

First, the motion of the system on an interval $T_{D_i}^R$ is examined. Suppose that there exists an $T_{D_i}^R$ which is unbounded. By definition of $T_{D_i}^R$, this implies that ϕ on $T_{D_i}^R$ is such that convexity/concavity of f is not changed. This further implies that for all $t_1, t_2 \in T_{D_i}^R$, $\max\left(|e_c(t_1) - e_c(t_2)|\right) = \epsilon$. The system dynamics on any $T_{D_i}^R$ are given by:

$$\begin{aligned} \dot{e}_c &= \tilde{f} \\ \dot{\tilde{\theta}} &= 0. \end{aligned} \tag{2.53}$$

Since f and ϕ are assumed to be continuous, by integrating the above equation, starting with any $t_0 \in T_{D_i}^R$, it is obtained that $|e_c(t) - e_c(t_0)| > \epsilon_f\,(t - t_0)$. Hence, there exists a $t_R^* > t_0 + \epsilon/\epsilon_F$ such that $|e_c(t^*) - e_c(t_0)| > \epsilon$. Thus, the system exists $T_{D_i}^R$ after an interval of length $\Delta\, t_L = t_L^* - t_0$, and hence the assumption that $T_{D_i}^R$ is unbounded does not hold. Since $T_{D_i}^R$ was arbitrary, all $T_{D_i}^R$ are bounded intervals.

Now suppose that there exists an interval $T_{D_i}^L$ which is unbounded. Suppose, further, that, while in $T_{D_i}^L$, eq. (2.46) does not hold. Since ϕ is bounded, there exists a_{max} such that $a < a_{max}$ on $T_{D_i}^L$. Then, starting at t_0 and integrating eq. (2.42), it can be obtained that $\exists t_L^*$,

$$t_L^* > t_0 + \frac{\epsilon}{a_{max}} \log\left(1 + \frac{a_{max}}{\epsilon_F}\right)$$

such that $|e_c(t^*) - e_c(t_0)| > \epsilon$. Thus, the assumption that $T^L_{D_i}$ is unbounded is contradicted if eq. (2.46) does not hold over a finite interval $\Delta t = t^*_L - t_0$. Now, suppose that eq. (2.46) *does* hold over an interval of length of at least Δt_L. Then, property (P2) is invoked to obtain that there exists values of ϕ which will make eq. (2.46) not hold over an interval of length Δt_L. Then, once again, it is obtained that the interval $T^L_{D_i}$ is finite. Since $T^L_{D_i}$ was arbitrary, all $T^L_{D_i}$ can be made bounded by the appropriate choice of ϕ.

Therefore, if ϕ is

(i) continuous,

(ii) preserves concavity/convexity of$f$ for an interval of length at least $\Delta_C = \Delta_L + \Delta_R$, and

(iii) periodically is changed as to satisfy property (P2)

then each interval T_{D_i} is bounded. Thus, the system will spend only a finite amount of time to cross the dead-zone.

Having shown that the time intervals over which $\dot{V}$ does not decrease are bounded, it will be demonstrated that outside these time intervals $\dot{V}$ decreases by a finite amount. It will be shown that V decreases over each interval T_{P_i}. This part of the proof follows the same steps as its the linear case counterpart, which is given in [18, 17, 19].

Since all the system signals are bounded, there must exist a δ_0 such that the interval $[\underline{t}, \underline{t} + \delta_0] \in T_{P_i}$. Suppose that e_c always stays small in T_{P_i}, that is $\exists c$, $0 < c < 1$ such that $\|e_\varepsilon(t)\| < c\|x(t)\|$ $\forall t \in [\underline{t}, \underline{t} + \delta_0]$. This implies that $\|\tilde{\theta}(t)\| > (1-c)\,\|x(t)\|$ $\forall t \in [\underline{t}, \underline{t} + \delta_0]$.

From the system dynamics, it follows:

$$\begin{aligned}\|e_\varepsilon(\underline{t} + \delta_0)\| &= \|e_\varepsilon(\underline{t}) + \int_{\underline{t}}^{\underline{t}+\delta_0} -ke_\varepsilon(\tau) + \tilde{f}(\tau) + a(\tau)\,d\tau\| \\ &\geq \|\int_{\underline{t}}^{\underline{t}+\delta_0} \tilde{f}(\tau) + a(\tau)\,d\tau\| - \|e_\varepsilon(\underline{t}) + k\int_{\underline{t}}^{\underline{t}+\delta_0} e_\varepsilon(\tau)\,d\tau\|\end{aligned}$$

$$\geq \quad \| \int_{\underline{t}}^{\underline{t}+\delta_0} \tilde{f}(\tau) + a(\tau)\, d\tau \| - \|e_\varepsilon(\underline{t})\| - k \, \| \int_{\underline{t}}^{\underline{t}+\delta_0} e_\varepsilon(\tau)\, d\tau \| \quad (2.54)$$

Since the time intervals in Ω_P on which $sign(e_c)f$ changes from concave to convex, or vice versa, are separated by an interval in Ω_D, two separate cases for the values of $a(\tau)$ can be considered:

(a0) $a(\tau) = 0 \; \forall \tau \in T_{P_i}$

(a1) $a(\tau) = d_c\left(\tau, \hat{\theta}(\tau)\right) \; \forall \tau \in T_{P_i}$.

It will be assumed that τ represents any time instant on the interval $[\underline{t}, \underline{t} + \delta_0]$, and that θ represents any element on the set $\Theta = [\underline{\theta}, \overline{\theta}]$. First, consider case *(a0)*. Calculate

$$\begin{aligned} \| \int_{\underline{t}}^{\underline{t}+\delta_0} \tilde{f}(\tau)\, d\tau \| &= \| \int_{\underline{t}}^{\underline{t}+\delta_0} \tilde{f}(\tau, \underline{t})\, d\tau - \int_{\underline{t}}^{\underline{t}+\delta_0} \tilde{f}(\tau, \underline{t}) - \tilde{f}(\tau, \tau)\, d\tau \| \\ &\geq \| \int_{\underline{t}}^{\underline{t}+\delta_0} \tilde{f}(\tau, \underline{t})\, d\tau \| - \| \int_{\underline{t}}^{\underline{t}+\delta_0} \tilde{f}(\tau, \underline{t}) - \tilde{f}(\tau, \tau)\, d\tau \| \quad (2.55) \end{aligned}$$

The second integral on the right hand side of eq. (2.55) can be expressed as:

$$\begin{aligned} \| \int_{\underline{t}}^{\underline{t}+\delta_0} \tilde{f}(\tau, \underline{t}) - \tilde{f}(\tau, \tau)\, d\tau \| &\leq \int_{\underline{t}}^{\underline{t}+\delta_0} \|\tilde{f}(\tau, \underline{t}) - \tilde{f}(\tau, \tau)\|\, d\tau \\ \| \int_{\underline{t}}^{\underline{t}+\delta_0} \tilde{f}(\tau, \underline{t}) - \tilde{f}(\tau, \tau)\, d\tau \| &\leq \int_{\underline{t}}^{\underline{t}+\delta_0} \|M(\tau) \left(\tilde{\theta}(\underline{t}) - \tilde{\theta}(\tau)\right) \|\, d\tau \\ &\leq \delta_0 \, M_m \sup_\tau \|\tilde{\theta}(\underline{t}) - \tilde{\theta}(\tau)\| \\ &\leq \delta_0 M_m \int_{\underline{t}}^{\underline{t}+\delta_0} \| \dot{\tilde{\theta}}(\tau) \|\, d\tau \\ &\leq \delta_0 M_m \int_{\underline{t}}^{\underline{t}+\delta_0} \|\omega e_\varepsilon\|\, d\tau \\ &\leq \delta_0^2 M_m M_\theta \sup_\tau \|e_\varepsilon(\tau)\| \\ &\leq c\, \delta_0^2 M_m M_\theta \|x(\underline{t}\| \, . \quad (2.56) \end{aligned}$$

where $M_\theta = \max_{\tau,\theta} \nabla_\theta f(\tau)$ is the maximum gradient with respect to θ of the curve $f(\phi, \theta)$, and $M_m = \max_\tau M(\tau)$, with $M(\tau)$ such that

$$\tilde{f}(\tau, t) - \tilde{f}(\tau, \tau) = M(\tau) \left(\tilde{\theta}(t) - \tilde{\theta}(\tau)\right).$$

The quantity $M(\tau)$ is well defined, since $\tilde{f}$ is continuous, and $\tilde{\theta}(t) = \tilde{\theta}(\tau)$ implies that $\tilde{f}(\tau, t) = \tilde{f}(\tau, \tau)$. Since $\tilde{f}$ is bounded because all the system signals are bounded, M_θ and M_m exist.

On the other hand, from eq. (2.51),

$$\| \int_{\underline{t}}^{\underline{t}+\delta_0} \tilde{f}(\tau, \underline{t})\, d\tau \| \geq \eta \geq \frac{\eta}{\|x(\underline{t})\|} \|\tilde{\theta}(\underline{t})\| \tag{2.57}$$

since $\|\tilde{\theta}(\tau)\| \leq \|x(\tau)\|$.

Substituting eqs. (2.57), (2.56) into eq. (2.55), and observing that $\|\tilde{\theta}\| \leq \|x(t_0)\|$, it follows

$$\| \int_{\underline{t}}^{\underline{t}+\delta_0} \tilde{f}(\tau)\, d\tau \| \;\; \geq \;\; \frac{\eta}{\|x(\underline{t})\|} \|\tilde{\theta}(\underline{t})\| - c\, \delta_0^2 M_m M_\theta \|x(\underline{t})\| \, . \tag{2.58}$$

In case *(a1)* $\tilde{f}(\tau) + a(\tau) = d_c(\tau, \theta) + \omega_s(\tau)\tilde{\theta}(\tau)$. Thus,

$$\| \int_{\underline{t}}^{\underline{t}+\delta_0} \tilde{f}(\tau) + a(\tau)\, d\tau \| = \| \int_{\underline{t}}^{\underline{t}+\delta_0} d_c(\tau, \theta) + \omega_s(\tau)\tilde{\theta}(\underline{t}) - \omega_s(\tau) \left(\tilde{\theta}(\underline{t}) - \tilde{\theta}(\tau)\right) d\tau \|$$

By using eq. (2.51), and noting that $\omega_s = \nabla_\theta f|_\xi$, $\xi \in \Theta$, following a similar procedure to the one for case (a0) yields:

$$\| \int_{\underline{t}}^{\underline{t}+\delta_0} \tilde{f}(\tau)\, d\tau \| \geq \frac{\eta}{\|x(\underline{t})\|} \|\tilde{\theta}(\underline{t})\| - c\, \delta_0^2 M_\theta^2 \|x(\underline{t})\| \, .$$

Letting $K = \max(M_\theta, M_m)$, both cases (a0) and (a1) can be compactly expressed as

$$\| \int_{\underline{t}}^{\underline{t}+\delta_0} \tilde{f}(\tau)\, d\tau \| \geq \frac{\eta}{\|x(\underline{t})\|} \|\tilde{\theta}(\underline{t})\| - c\, \delta_0^2 K^2 \|x(\underline{t})\| \, . \tag{2.59}$$

Substituting eq. (2.59) into eq. (2.54)

$$\|e_\varepsilon(\underline{t} + \delta_0)\| \geq \frac{\eta}{\|x(\underline{t})\|} \|\tilde{\theta}(\underline{t})\| - c\, \delta_0^2 K^2 \|x(\underline{t})\| \, . - c\|x(\underline{t})\| - k\delta_0 c\|x(\underline{t})\| \tag{2.60}$$

Since $\|\tilde{\theta}(\underline{t})\| \geq \sqrt{1-c^2}\|x(\underline{t})\|$,

$$\|e_\varepsilon(\underline{t}+\delta_0)\| \geq \left\{ \frac{\eta}{\|x(\underline{t})\|}\sqrt{1-c^2} - c\left(\delta_0^2 K^2 \|x(\underline{t})\| . + 1 + k\delta_0\right)\right\} \|x(\underline{t})\| \tag{2.61}$$

Let $b_1 = \frac{\eta}{\|x(\underline{t})\|}$, and $b_2 = K^2\delta_0^2 + k\delta_0 + 1$.

If $c^2 = \frac{{b_1}^2}{{b_1}^2 + (b_2+1)^2}$ then

$$\|e_\varepsilon(\underline{t}+\delta_0)\| \;\geq\; c\|x(\underline{t})\| \;\geq\; c\|x(\underline{t}+\delta_0)\| \tag{2.62}$$

since $\|x(\underline{t})\| \geq \|x(\underline{t}+\delta_0)\|$. Hence, the assumption that e_c stays small is contradicted. Eq. (2.62) implies that there exists an interval $t \in [t_1, t_1+T]$ such that

$$\tilde{\theta}(t)^2 \leq (1-c^2)\|x(t_1)\|^2 \leq V(t_1)(1-c^2) \tag{2.63}$$

Calculating the change of V over the interval $[t_1, t_1+T]$:

$$V(t_1) - V(t_1+\delta_0) \geq k\int_{t_1}^{t_1+T} e_\varepsilon(\tau)^2\, d\tau \tag{2.64}$$

For any $t \geq t_1$ we have

$$\begin{aligned} \|e_\varepsilon(t_1)\| - \|e_\varepsilon(t)\| &\leq \|e_\varepsilon(t) - e_\varepsilon(t_1)\| && (2.65)\\ &\leq \| \int_{t_1}^{t} \| - ke_\varepsilon(\tau) + \tilde{f}(\tau) + a(\tau)\, d\tau \| && (2.66)\\ &\leq k\|x(t_1)\|(t-t_1) + \int_{t_1}^{t} \|B(\tau)\tilde{\theta}(\tau)\, d\tau\| + \int_{t_1}^{t} \|a(\tau)\, d\tau\| && (2.67) \end{aligned}$$

where, by the mean value theorem, $\tilde{f}(\tau) = B(\tau)\tilde{\theta}(\tau)$. By the definition $\|B(\tau)\| \leq M_\theta$ $\forall \tau \in [t_1, t]$. Let $A_m = \max_{\tau\in[t_1,t]} a(\tau)\frac{1}{\|x(t_1)\|}$. Thus,

$$\|e_\varepsilon(t_1)\| - \|e_\varepsilon(t)\| \leq (k + M_\theta + A_m)\|x(t_1)\|(t-t_1)$$

For some $t_2 = t_1 + c_1$, $c_1 > 0$

$$\begin{aligned} \|e_\varepsilon(t_2)\| &\geq \|e_\varepsilon(t_1)\| - c_1(k + M_\theta)\|x(t_1)\| &(2.68)\\ &\geq \sqrt{V(t_1)\,(1-c^2)} - c_1(k + M_\theta + A_m)\sqrt{V(t_1)} &(2.69)\\ &\geq \left[\sqrt{1-c^2} - c_1(k + M_\theta + A_m)\right]\sqrt{V(t_1)} &(2.70)\\ &\geq d\sqrt{V(t_1)} &(2.71) \end{aligned}$$

where $d = \sqrt{1-c^2} - c_1(k + M_\theta + A_m)$.

Hence,

$$\int_{t_1}^{t_1+T} e_\varepsilon(\tau)^2\, d\tau \geq c_1\, V(t_1)\, d^2.$$

Thus

$$V(t_1 + T) \leq (1 - c_1 d^2) V(t_1),$$

which establishes the theorem. ■

The persistent excitation conditions for the min-max algorithm in eq. (2.51) are similar the one for linearly parameterized systems. Since this discussion treats only the case of a single unknown parameter, this condition can be satisfied if $\phi(t)$ is a piecewise constant, or slowly varying signal.

2.5 Numerical examples

This section presents two numerical examples which highlight the discussed issues of parameter convergence with the min-max algorithm. Three examples are presented which emphasize the crucial role the property (P2) plays in establishing parameter convergence. The first example numerically demonstrates the behavior of the system discussed in Section 2.3. The second example shows another type of parametric nonlinearity which satisfies (P2). In the third example, the function is such that property (P2) cannot be satisfied. Thus, min-max algorithm is capable of tracking the system output, but it cannot be guaranteed apriori that the value of the unknown parameter will be estimated precisely.

2.5.1 Example A

The system in this example has the same parametric nonlinearity as the one discussed in Section (2.3). The function f used in Section 2.3 is

$$\dot{y} = -\lambda y + f(\phi, \theta) = -\lambda y + e^{-\phi\theta} \tag{2.72}$$

It is assumed that the set $\Theta = [\underline{\theta}, \overline{\theta}]$ is known, and that $\underline{\theta} \leq \theta \leq \overline{\theta}$. From eqs. (2.16)- (2.20), the min-max estimator is constructed as:

$$\begin{aligned}
\dot{\hat{y}} &= -\lambda y - k e_\varepsilon + e^{-\phi\hat{\theta}} - a\, sat\left(\frac{e_c}{\epsilon}\right) \\
\dot{\hat{\theta}} &= -e_\varepsilon \omega \\
a &= \begin{cases} 0\,, & e_c \geq 0 \\ f_{min} + \dfrac{f_{max} - f_{min}}{\overline{\theta} - \underline{\theta}}(\hat{\theta} - \underline{\theta}) - \hat{f}\,, & e_c < 0 \end{cases} \\
\omega &= \begin{cases} -\phi\, e^{-\phi\hat{\theta}}\,, & e_c \geq 0 \\ \dfrac{f_{max} - f_{min}}{\overline{\theta} - \underline{\theta}}\,, & e_c < 0 \end{cases} \\
e_\varepsilon &= e_c - \epsilon\, sat\left(\frac{e_c}{\epsilon}\right)
\end{aligned}$$

If and only if $\phi = 0$, $\tilde{f} = 0$ for all values of $\tilde{\theta}$. Hence, the property (P1) is satisfied. The function f is such that it is convex in θ for all values of ϕ. However, it can be noticed that f is monotonically decreasing when $\phi(t) > 0$, and that f is monotonically increasing when $\phi(t) < 0$. Hence, any given sign of $\tilde{f} = f(\phi, \hat{\theta}) - f(\phi, \theta)$ can be reversed by letting $\phi(t)$ sequentially take on positive and negative values. Therefore,

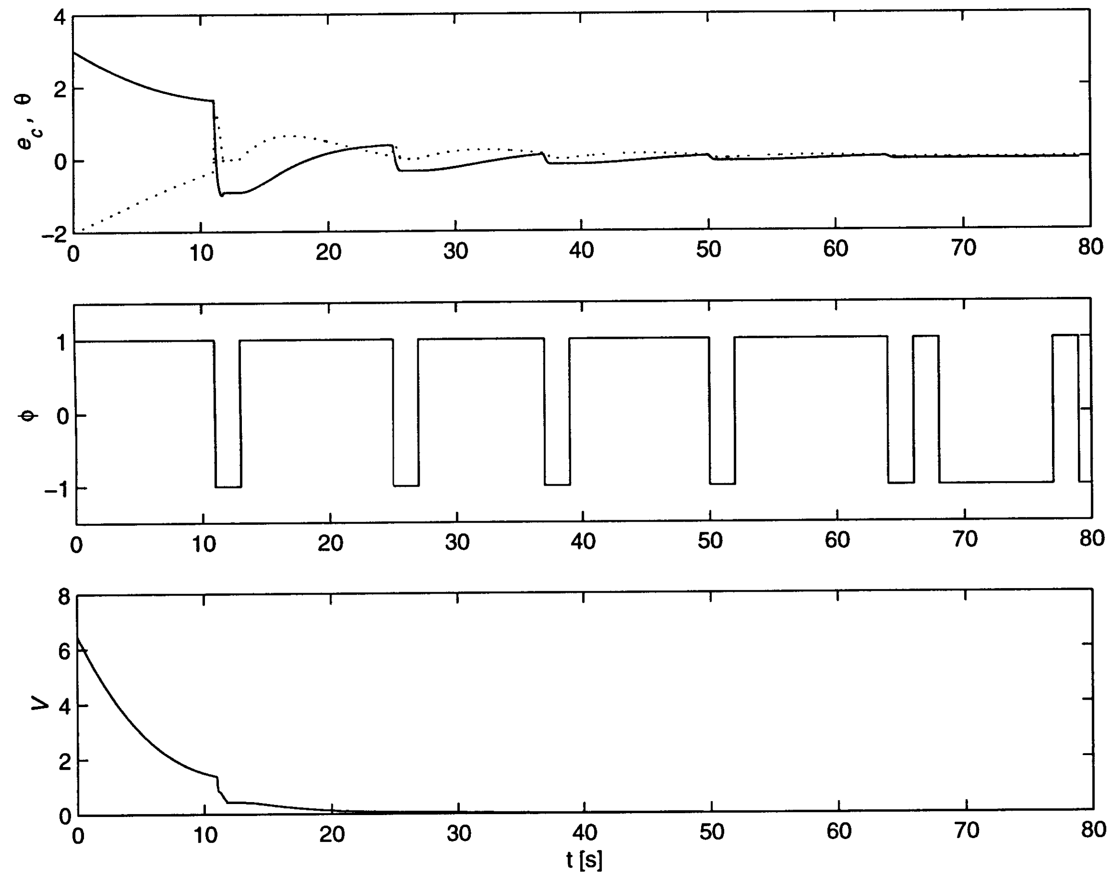

Figure 2-3: Simulation results for example A.
Top panel: Tracking and parameter error. Solid line represents the parameter error, $\tilde{\theta}$, while the tracking error, e_c, is given by the dashed line.
Middle panel: Input. The value of ϕ used in the simulation run.
Bottom panel: Lyapunov function $V = \frac{1}{2}\left(e_\varepsilon^2 + \tilde{\theta}^2\right)$.

property (P2) can be satisfied, and precise estimation of the unknown parameter is possible. For the simulation run, the following set of values were used:

$$\begin{array}{llllllll} \underline{\theta} = .5 & \widehat{\theta}(0) = 4.0 & \widehat{y}(0) = 3 & \lambda = 1 & \\ & & & & \epsilon = 0.01 \\ \overline{\theta} = 6 & \theta = 1.5 & y(0) = 5 & k = 0.1 & \end{array}$$

2.5.2 Example B

For this example, the following system is chosen:

$$\dot{y} = -\lambda y + f(\phi, \theta) = -\lambda y + (\phi - \theta)^2 \tan\left(\frac{\pi}{180}(\phi - \theta)^2\right) \tag{2.73}$$

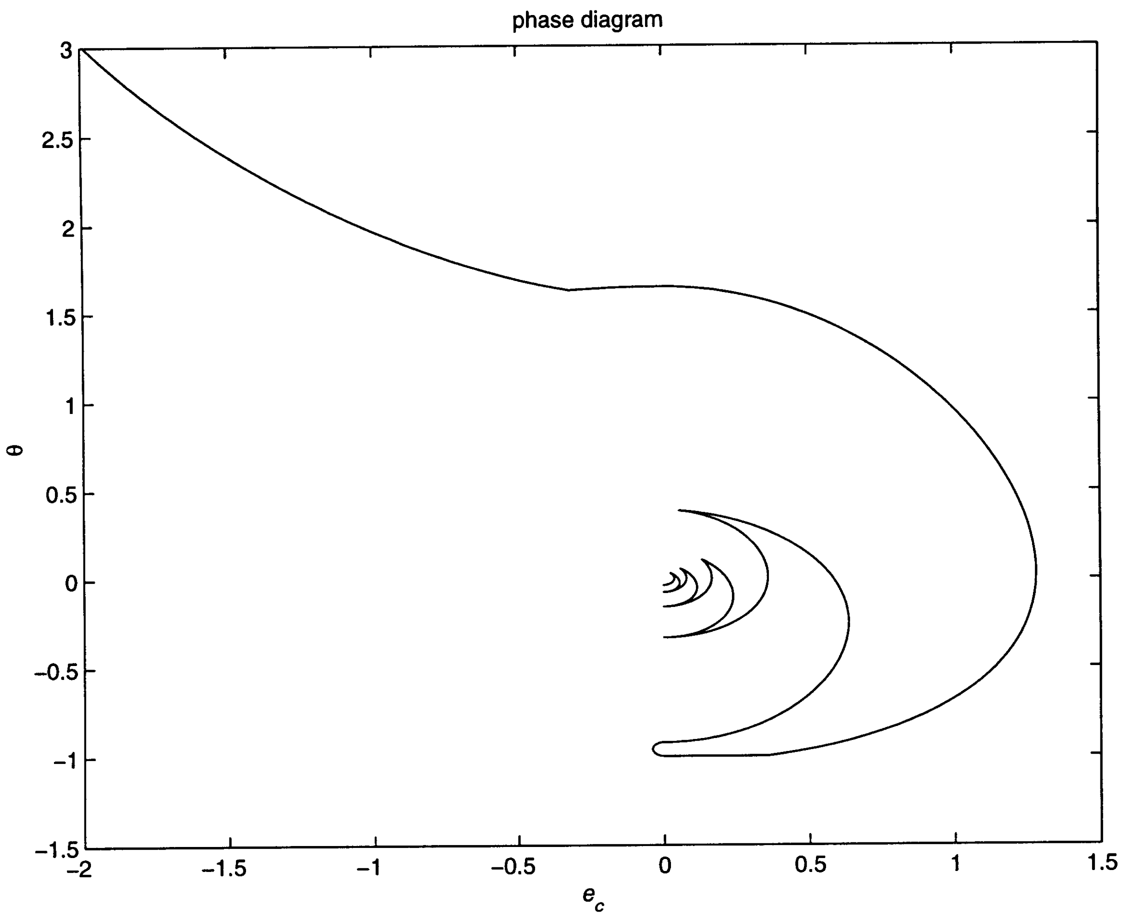

Figure 2-4: Phase plane diagram for example A. The e_c axis is horizontal, while the $\tilde{\theta}$ axis is vertical.

It can be verified that the function f is convex with respect to θ for all $\theta, \phi \in \mathbb{R}$. It is assumed that the set $\Theta = [\underline{\theta}, \overline{\theta}]$ is known and that $\underline{\theta} \leq \theta \leq \overline{\theta}$. From eqs. (2.16)- (2.20), the min-max estimator is constructed as:

$$
\begin{aligned}
\dot{\hat{y}} &= -\lambda y - k e_\varepsilon + \left(\phi - \hat{\theta}\right)^2 \tan\left(\frac{\pi}{180}\left(\phi - \hat{\theta}\right)^2\right) - a\, sat\left(\frac{e_c}{\epsilon}\right) \\
\dot{\hat{\theta}} &= -e_\varepsilon \omega \\
a &= \begin{cases} 0\,, & e_c \geq 0 \\ f_{min} + \dfrac{f_{max} - f_{min}}{\overline{\theta} - \underline{\theta}}\,(\hat{\theta} - \underline{\theta}) - \hat{f}\,, & e_c < 0 \end{cases} \\
\omega &= \begin{cases} -2\left(\phi - \hat{\theta}\right)\left(\tan\left(\frac{\pi}{180}\left(\phi - \hat{\theta}\right)^2\right) + \dfrac{\pi}{180}\dfrac{(\phi - \hat{\theta})^2}{\cos^2\left(\frac{\pi}{180}\left(\phi - \hat{\theta}\right)^2\right)}\right), & e_c \geq 0 \\ \dfrac{f_{max} - f_{min}}{\overline{\theta} - \underline{\theta}}\,, & e_c < 0 \end{cases} \\
e_\varepsilon &= e_c - \epsilon\, sat\left(\frac{e_c}{\epsilon}\right)
\end{aligned}
$$

By inspecting the function f it can be noticed that f is monotonically decreasing on the interval Θ when $\phi(t) = \overline{\theta}$, and that f is monotonically increasing on the interval Θ when $\phi(t) = \underline{\theta}$. Hence, any given sign of $\tilde{f} = f(\phi, \hat{\theta}) - f(\phi, \theta)$ can be reversed by letting $\phi(t)$ sequentially take on values of $\underline{\theta}$ and $\overline{\theta}$, respectively. Therefore, property (P2) can be satisfied, and precise estimation of the unknown parameter is possible. For the simulation run, the following set of values were used:

$$
\begin{array}{llllll}
\underline{\theta} = 1 & \hat{\theta}(0) = 9.5 & \hat{y}(0) = 3 & \lambda = 1 & & \\
& & & & & \epsilon = 0.01 \\
\overline{\theta} = 10 & \theta = 5.5 & y(0) = 5 & k = 0.1 & &
\end{array}
$$

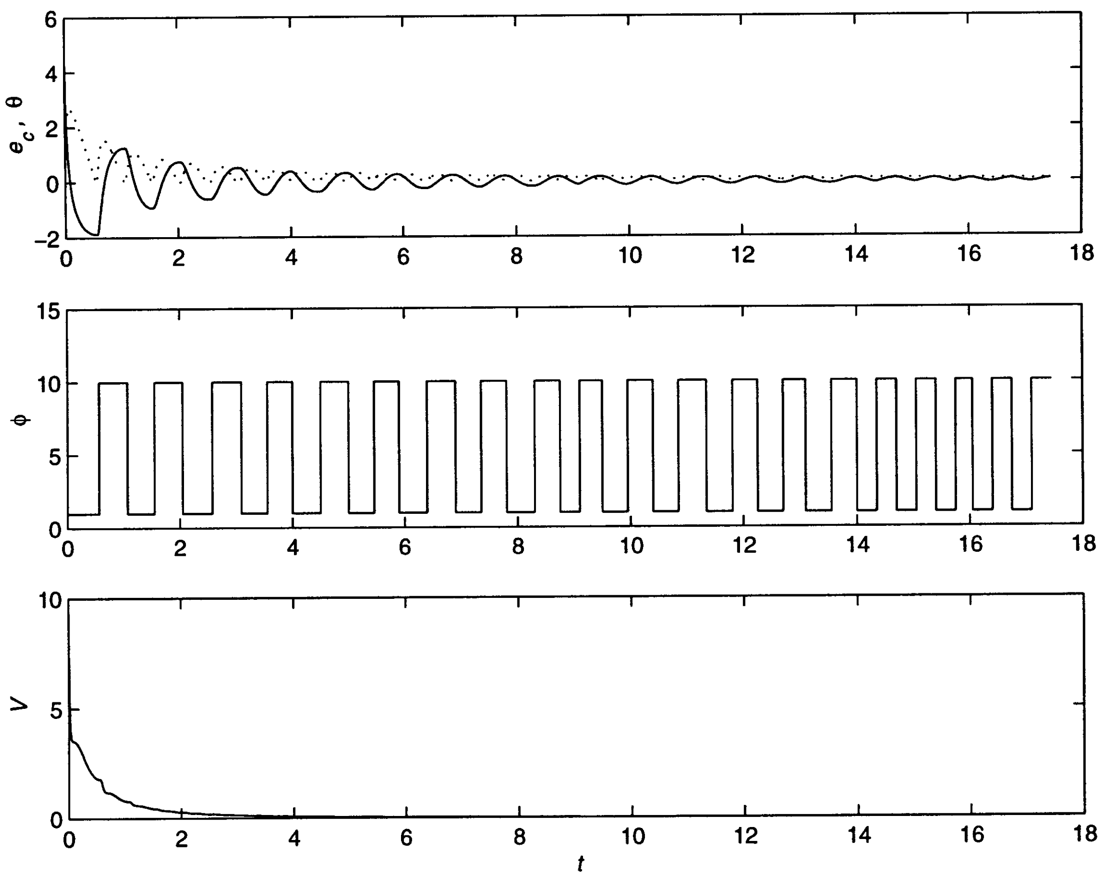

Figure 2-5: Simulation results for example B.
Top panel: Tracking and parameter error. Solid line represents the parameter error, $\tilde{\theta}$, while the tracking error, e_c, is given by the dashed line.
Middle panel: Input. The value of ϕ used in the simulation run.
Bottom panel: Lyapunov function $V = \frac{1}{2}\left(e_\varepsilon^2 + \tilde{\theta}^2\right)$.

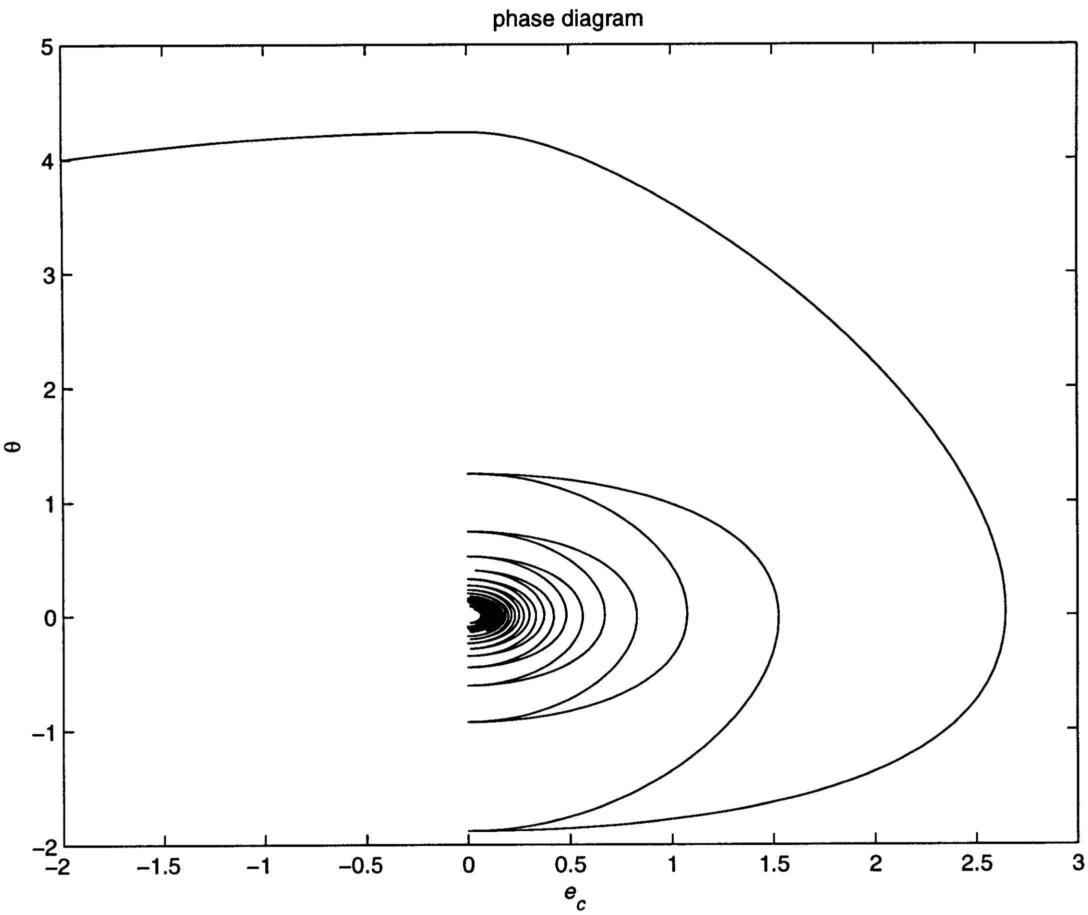

Figure 2-6: Phase plane diagram for example B. The e_c axis is horizontal, while the $\tilde{\theta}$ axis is vertical.

2.5.3 Example C

In this example, the system of interest is

$$\dot{y} = -\lambda y + f(\phi, \theta) = -\lambda y + sign(\phi)\, e^{-\phi^2 \theta} \tag{2.74}$$

The min-max estimator is constructed by using eqs. (2.16)- (2.20):

$$\begin{aligned}
\dot{\hat{y}} &= -\lambda y - k e_\varepsilon + sign(\phi)\, e^{-\phi^2 \theta} - a\, sat\left(\frac{e_c}{\epsilon}\right) \\
\dot{\hat{\theta}} &= -e_\varepsilon \omega \\
a &= \begin{cases} 0\,, & sign(\phi\, e_c) \geq 0 \\ -\, sign(e_c)\left[f_{min} + \dfrac{f_{max} - f_{min}}{\overline{\theta} - \underline{\theta}}\,(\hat{\theta} - \underline{\theta}) - \hat{f}\right], & sign(\phi\, e_c) < 0 \end{cases} \\
\omega &= \begin{cases} -\phi^2\, sign(\phi)\, e^{-\phi^2 \hat{\theta}}\,, & sign(\phi\, e_c) \geq 0 \\ \dfrac{f_{max} - f_{min}}{\overline{\theta} - \underline{\theta}}\,, & sign(\phi\, e_c) < 0 \end{cases} \\
e_\varepsilon &= e_c - \epsilon\, sat\left(\frac{e_c}{\epsilon}\right)
\end{aligned}$$

The function f is, depending on the sign of $\phi(t)$, either a convex decreasing or a concave increasing function with respect to θ. For a positive value of ϕ, f is a convex decreasing function in θ. For a negative value of ϕ, f is a concave increasing function of θ. Thus, switching ϕ between positive and negative values changes the sign of any $\tilde{f}$. However, in doing so, the convexity/concavity of f is also changed. Hence, property (P2) cannot be satisfied, and satisfactory parameter estimation is not guaranteed with this algorithm. This is illustrated in the simulation run with the

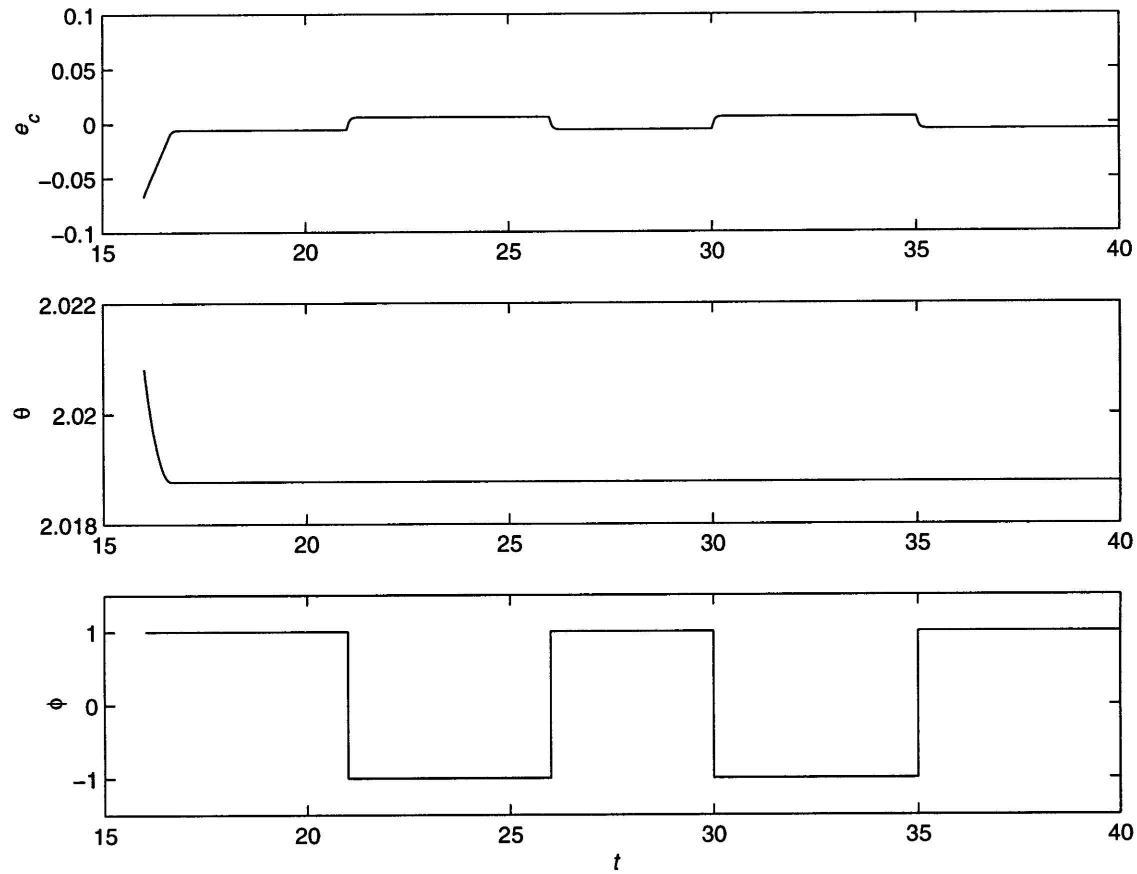

Figure 2-7: Simulation results for example C. The top panel gives the tracking error, the middle one the parameter error, and the bottom panel gives the value of ϕ used.

following set of values:

$$\begin{array}{llllllll} \underline{\theta} = 0.5 & \widehat{\theta}(0) = 5.8 & \widehat{y}(0) = 3 & \lambda = 1 & \\ & & & & \epsilon = 0.01 \\ \overline{\theta} = 6 & \theta = 2 & y(0) = 5 & k = 0.1 & \end{array}$$

Even though the parametric nonlinearity is an exponential form similar to the one in example A, the same strategy of switching the signs of ϕ does not yield satisfactory convergence of the parameter error. In fact, there exists a nonzero lower bound for the guaranteed decrease in the parameter error. The lower bound is given by the quantity $\tilde{\theta}'$ defined in Section 2.3 as $\tilde{\theta}' = \frac{1}{\omega_s}(f - f_{min}) + \underline{\theta} - \theta$. For the values used in this example, $\tilde{\theta}' = 2.7903$. Thus, if the adaptive system enters the deadzone with $\tilde{\theta}$ such that $0 \le \tilde{\theta} \le 2.7903$, it will not be able to leave the deadzone again. One possible remedy to this obstacle is to increase the value of ϕ, rendering the exponential curve

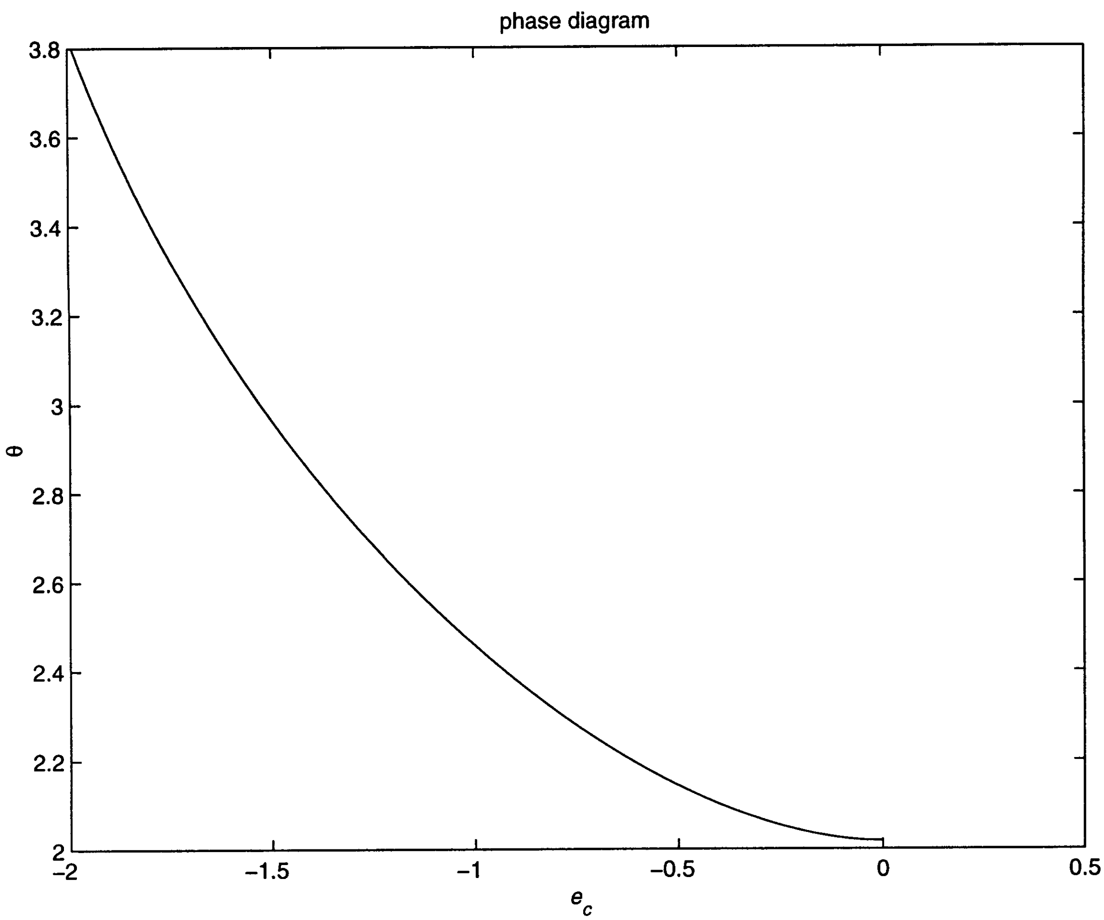

Figure 2-8: Phase plane diagram for example C. The e_c axis is horizontal, while the $\tilde{\theta}$ axis is vertical.

essentially flat. Thus, for large values of ϕ the curve becomes almost linear, implying that $\frac{1}{\omega_s}(f - f_{min}) \rightarrow \theta - \underline{\theta}$, and hence $\tilde{\theta}'$ would decrease.

Chapter 3

Monotonic Parameterization - Part I

3.1 Introduction

The previous chapter addressed the problem of parameter estimation and convergence in convex/concave functions. This chapter is concerned with the problem of parameter estimation in another specific class of nonlinearly parameterized systems. The class of systems of interest in this chapter is characterized by the presence of a nonlinearity which is monotonic in the unknown parameter.

It was seen in the last chapter that the use of the gradient parameter update law was not always sufficient to guarantee stability. Intuitively, one of the reasons why the gradient does not suffice may stem from the fact that the gradient with respect to the unknown parameter in convex/concave functions may change signs on the interval of interest. The sign of the gradient is important, because it relates the observed error in the function with the parameter error, and thereby provides a means for determining in which direction the parameter estimate should be adjusted. Monotonic functions remove the possibility of gradient changing signs, since their gradient, by definition, is always of the same sign. In fact, using this viewpoint, linearly parameterized systems can be considered as a special case within the monotonic class of functions.

The group of linear parameterized systems differ from the general monotonic class of functions because they provide an additional, albeit crucial, benefit in the task of parameter estimation. In linear parameterization, not only is the gradient of the same sign, but its value is the same everywhere. General monotonic functions considered in this chapter do not provide that luxury. As a result, it is possible for instability to occur in the case of general monotonic parameterization which cannot occur with linear parameterization. This chapter investigates under what conditions instability in monotonic parameterizations may occur.

For the sake of easier mathematical tractability, only the case of a single unknown parameter is examined in this chapter. In that case, the closed-loop adaptive system dynamics are of the second order. This allows for a direct analogy with the mass-spring-damper mechanical system. This analogy is used to gain a clearer insight into the problem of stability. It will be shown that the equivalent mechanical system has time-varying coefficients. The relationship between these coefficients, their relative rate of change and implications on stability are discussed. In particular, it will be shown that these relationships in the linear case are such that stability is guaranteed. In contrast, in the monotonic nonlinear case, it will be shown that only a part of the corresponding linear case relationships hold, thus allowing for the possibility of instability in the system. Section 3.2 starts the discussion with a proof of a simple and useful Lemma. In Section 3.3, the issue of stability of an adaptive system with a monotonic nonlinearity is treated by utilizing the mechanical system analogy and by comparing the monotonic and linear case. Finally, Section 3.4 provides a numerical example that illustrates the kind of issues addressed in the sections that precede it.

3.2 Preliminaries

Lemma 3.1 *Let $f(t)$ be an everywhere differentiable function of time such that $f(t) \geq 0$. Let t_0 be such that $f(t_0) = 0$. Then*

$$\lim_{t \to t_0} \frac{f'(t)}{f(t)} = -\infty$$

Proof: Since $\min f(t) = 0$, then $f'(t_0) = 0$. Thus, it can be suspected that the ratio $\dfrac{f'(t_0)}{f(t_0)}$ is either zero or undefined. However, it will be shown that it is not the case. The difference $f(t) - f(t_0)$ can be written as:

$$\begin{aligned} f(t) - f(t_0) &= \frac{f(t) - f(t_0)}{t - t_0} (t - t_0) \\ f(t) &= \frac{f(t) - f(t_0)}{t - t_0} (t - t_0) + f(t_0) \end{aligned}$$

Because $f(t_0) = 0$, then

$$f(t) = \frac{f(t) - f(t_0)}{t - t_0} (t - t_0)$$

Substituting this into the ratio and taking the limit it follows that:

$$\begin{aligned} \lim_{t \to t_0} \frac{f'(t)}{f(t)} &= \lim_{t \to t_0} f'(t) \frac{1}{\lim\limits_{t \to t_0} \dfrac{f(t) - f(t_0)}{t - t_0}} \frac{1}{\lim\limits_{t \to t_0} t - t_0} \\ &= f'(t_0) \frac{1}{f'(t_0)} \frac{1}{\lim\limits_{t \to t_0} t - t_0} \\ &= \lim_{\epsilon \to 0+} \frac{1}{-\epsilon} \\ &= -\infty \end{aligned}$$

■

3.3 A simple mechanical system analogy

Consider the following system

$$\ddot{x} + \left(a - \frac{\dot{k}}{2k} \right) \dot{x} + kx = 0 \tag{3.1}$$

Assume that the system satisfies the following:

(i) $0 < \alpha_{min} \leq a \leq \alpha_{max}$

(ii) $0 \leq k(t) \leq M$, $k(t)$ is differentiable

(iii) $\exists T_0$, δ_0, ϵ_0, $t_2 \in [t, t+T_0]$ such that

$$\frac{1}{T_0} \int_{t_2}^{t_2+\delta_0} k \, dt \ \geq \epsilon_0 \qquad \forall t > t_0 \tag{3.2}$$

The origin of $\mathbf{x} = [x, \dot{x}]^T$ is then uniformly asymptotically stable. Take $V(\mathbf{x}, t) = \frac{\dot{x}^2}{2k} + \frac{x^2}{2}$. Then

$$V(\mathbf{x}, t) \geq \frac{1}{2(M+1)} \left(\dot{x}^2 + x^2\right) = \frac{1}{2(M+1)} \, ||\mathbf{x}||^2$$

Thus, V is positive definite and $||\mathbf{x}||^2 \leq 2(M+1) \, V(\mathbf{x}, t)$. Taking the time derivative of V along the system trajectories it follows that

$$\begin{aligned} \dot{V}(\mathbf{x}, t) &= \frac{\dot{x} \, \ddot{x}}{k} - \frac{\dot{x}^2}{2k^2} \, \dot{k} + x \, \dot{x} \\ &= -\frac{\dot{x}^2}{k} \left(a - \frac{\dot{k}}{2k} \right) - \dot{x} \, x - \frac{\dot{x}^2}{2k^2} \, \dot{k} + x \, \dot{x} \\ &= -\frac{a}{k} \, \dot{x}^2 \leq -\frac{a}{M} \, \dot{x}^2. \end{aligned}$$

Therefore, the origin of $\mathbf{x}$ is stable. To show that the origin is asymptotically stable, an approach similar to the previous section and given by [18, 17, 19] is used. For a detailed discussion, the reader is reffered to those references. The condition in *(iii)* implies that there exist a t_3, ϵ_1, ϵ_2 such that $0 < \epsilon_2 < \epsilon_1 \leq \epsilon_0$ and

$$\frac{1}{T_0} \int_{t_3}^{t_3+\delta_0} k \, dt \ \geq \epsilon_1 \qquad \forall t > t_0 \tag{3.3}$$

with $k(t_3) > \epsilon_K > 0$. Assume that $\exists c < 1$ such that $| \, \dot{x} \, | \leq c||\mathbf{x}|| \ \forall t \in [t_0, t_0 + T_0]$.

The following is true:

$$\begin{aligned}
\int_{t_3}^{t_3+\delta_0} k(\tau)x(\tau)\, d\tau &\geq \int_{t_3}^{t_3+\delta_0} |k(\tau)x(t_3)| d\tau - \int_{t_3}^{t_3+\delta_0} |k(\tau)\ (x(t_3) - x(\tau))\,|\, d\tau \\
&\geq \epsilon_1\, T_0\, |x(t_3)| - M\delta_0^2 c \|\mathbf{x}(t_3)\| \qquad (3.4)
\end{aligned}$$

Eq. (3.4) is now employed to obtain:

$$\begin{aligned}
|\,\dot{x}\ (t_3+\delta_0)| &= \left| \dot{x}\ (t_3) + \int_{t_3}^{t_3+\delta_0} \left(a - \frac{\dot{k}}{2k}\right) \dot{x}\ (\tau)\ d\tau + \int_{t_3}^{t_3+\delta_0} -k(\tau)x(\tau)\ d\tau) \right| \\
&\geq \left| \int_{t_3}^{t_3+\delta_0} k(\tau)x(\tau)\ d\tau \right| - \left| \dot{x}\ (t_3) + \int_{t_3}^{t_3+\delta_0} \left(a - \frac{\dot{k}}{2k}\right) \dot{x}\ (\tau)\ d\tau \right| \\
&\geq \left(\epsilon_1 T_0 |x(t_3)| - M\delta_0^2 c \|\mathbf{x}(t_3)\| \right) \\
&\qquad - \left(c - \alpha_{max}\delta_0 c - c\log\left(\frac{k(t_3+\delta_0)}{k(t_3)}\right)\right) \|\mathbf{x}(t_3)\| \\
&\geq \epsilon_1 T_0 |x(t_3)| - M\delta_0^2 c \|\mathbf{x}(t_3)\| - (\alpha_{max}\delta_0 c + Kc + c)\, \|\mathbf{x}(t_3)\|
\end{aligned}$$

where $K = \log \frac{M}{\epsilon_K}$. Noting that, by the initial assumption, $|x(t_3)| \geq \sqrt{1-c^2}\ \|\mathbf{x}(t_3)\|$, and letting $p_1 = M\delta_0^2 + \alpha_{max}\delta_0 + K + 1$, it is obtained that:

$$|\dot{x}\ (t_3+\delta_0)| \geq \left(\epsilon_1 T_0 \sqrt{1-c^2}\ - p_1\, c \right) \|\mathbf{x}(t_3+\delta_0)\|$$

Thus, if

$$c^2 = \frac{\epsilon_1^2 T_0^2}{\epsilon_1^2 T_0^2 + (p_1+1)^2}$$

then,

$$|\dot{x}\ (t_3+\delta_0)| \geq c\|\mathbf{x}(t_3+\delta_0)\|$$

Hence, the initial assumption that $|\dot{x}\ (t_3+\delta_0)|$ always remains small is incorrect. Because $\dot{V}$ is negative definite, and since its magnitude depends on the magnitude of

$|\dot{x}\,(t_3+\delta_0)|$, it can be shown that

$$V(\mathbf{x}, t_3+\delta_0) \leq \left(1-\delta_0 c^2 d^2\right) V(\mathbf{x}, t_3)$$

where $d = \sqrt{\epsilon_1(1-c^2)} - 2\delta_0(K+M)(M+1)$. Thus, $V(\mathbf{x},t)$ decreases by a finite amount, implying $\lim_{t\to\infty} V(\mathbf{x},t) = 0$. Since $\|\mathbf{x}\| < 2(M+1)V$, it follows that $\lim_{t\to\infty} \|\mathbf{x}(t)\| = 0$. ■

Eq. (3.1) describes the motion of a second-order unforced oscillatory system. The damping coefficient given by $\zeta = a - \frac{\dot{k}}{2k}$ and the spring constant, k, are time-varying functions. The two quantities are not arbitrary and independent of each other. Rather, the value of the spring constant and its rate of change influence the damping coefficient through negative feedback. Thus, when the spring constant decreases, indicating a decrease in the force which tends to restore the system to the origin, the damping is increased. Increased damping means that the system will find more resistance in moving away from the origin. By applying Lemma 3.1, it follows that when $k \to 0$, the damping coefficient becomes very large, $\zeta \to \infty$. Hence, with a zero restoration force from the spring, the system will stop its motion because the resistive force will be infinitely large. On the other hand, when the restoring force is increasing due to an increase in the spring constant, the resisting force is decreased. It is interesting to compare the ratio of $\eta = \frac{\zeta}{k}$ for this case.

$$\eta = \frac{\zeta}{k} = \frac{a}{k} - \frac{\dot{k}}{2k}.$$

Then, $\eta > 1$ if

$$\dot{k} < 2k\,(a-k). \tag{3.5}$$

Since $\dot{k} > 0$ for this case, $\eta > 1$ only while $k < a$. However, for all cases when $\eta > 1$ it follows that $\zeta > k \geq 0$, implying that the system is dissipating energy due to the resistance force. For some time intervals where a very rapid exponential increase in k occurs ζ may become negative, indicating that energy is being added to the system.

However, when $\zeta < 0$, $\eta < 1$, indicating that the restoring force is dominant.

Consider now the standard form for the closed loop dynamics of a linear adaptive system with one unknown parameter:

$$\begin{aligned} \dot{x}_1(t) &= -ax_1 + u(t)\,x_2 \\ \dot{x}_2(t) &= -\Gamma\, u(t)\, x_1 \end{aligned} \tag{3.6}$$

where $\Gamma > 0$ is a constant gain. By differentiating the second equation with respect to time, it is obtained that

$$\ddot{x}_2 = -\Gamma\, \dot{u}\, x_1 - \Gamma u\, \dot{x}_1 \, .$$

Also, from the same equation $x_1 = -\dfrac{\dot{x}_2}{\Gamma u}$. Thus,

$$\begin{aligned} \ddot{x}_2 &= -\Gamma\, \dot{u} \left(-\frac{\dot{x}_2}{\Gamma u}\right) - \Gamma u \left(a \frac{\dot{x}_2}{\Gamma u} + u\, x_2\right) \\ \ddot{x}_2 &= \frac{\dot{u}}{u}\, \dot{x}_2 - a\, \dot{x}_2 - \Gamma\, u^2 x_2 \end{aligned}$$

which can be put into the form of a second-order unforced oscillatory system as:

$$\ddot{x}_2 + \left(a - \frac{\dot{u}}{u}\right) \dot{x}_2 + \Gamma\, u^2 x_2 = 0. \tag{3.7}$$

By letting $k = \Gamma\, u^2$, eq. (3.7) takes the form of eq. (3.1). It can be seen that the adaptation gain Γ has a direct influence only on the spring constant of the system. The fact that the spring constant is directly proportional to Γ explains the observed increase in oscillatory behavior with higher values of Γ.

An adaptive system with a monotonically parameterized nonlinearity is of the form:

$$\begin{aligned} \dot{x}_1(t) &= -ax_1 + \tilde{f}\,(u(t), x_2) \\ \dot{x}_2(t) &= -\Gamma\, \omega(t)\, x_1 \end{aligned} \tag{3.8}$$

where $\tilde{f}(u, x_2) = f(u, x_2 + \theta) - f(u, \theta)$, θ is a constant unknown parameter, and $\omega(t)$ is the tuning signal. By using the mean value theorem [23],

$$\tilde{f}(u, x_2) = f(u, x_2 + \theta) - f(u, \theta) = b(t)\, x_2$$

where $b(t)$ is the value of the gradient of f with respect to x_2 at some point ξ on the interval $[x_2, x_2 + \theta]$. The system of eq. (3.8) is rewritten as:

$$\begin{aligned} \dot{x}_1\,(t) &= -a x_1 + b(t)\, x_2 \\ \dot{x}_2\,(t) &= -\Gamma\, \omega(t)\, x_1 \end{aligned} \qquad (3.9)$$

Because the function is monotonic, the sign of $b(t)$ is known, and is equal to the sign of $\dfrac{\partial f}{\partial x_2}$ for any value of x_2. Since the sum $x_2 + \theta$, which represents the estimate of θ, is known, let $\omega = \left.\dfrac{\partial f}{\partial x_2}\right|_{x_2+\theta}$. Then, the sign of $\omega(t)$ is always equal to the sign of $b(t)$. By using the similar methodology as in the linear case, the adaptive system of eq. (3.8) can be reduced to the following second order system

$$\ddot{x}_2 + \left(a - \frac{\dot{\omega}}{\omega}\right)\, \dot{x}_2 + \Gamma\, \omega\, b\, x_2 = 0 \qquad (3.10)$$

The above system has the following equivalent mechanical properties: the damping coefficient $\zeta = \left(a - \dfrac{\dot{\omega}}{\omega}\right)$, and a spring constant $k = \Gamma\, \omega\, b\, > \, 0$. The time derivative of the spring constant is given by $\dot{k} = \Gamma\, \left(\dot{\omega}\ b\, +\, \omega\, \dot{b}\right)$. Therefore,

$$\frac{\dot{k}}{2k} = \frac{\dot{\omega}}{2\omega} + \frac{\dot{b}}{2b}$$

Hence, by comparing eq. (3.1) and eq. (3.8) it can be observed that for the latter system there is no full information feedback about the rate of change of the spring constant into the damping coefficient, as is the case in the former system. Thus, the stability guarantees for the system in eq. (3.1) do not directly apply to the system in eq. (3.8). Since it is assumed that u and the function f are such that both ω and b are bounded, the function $V(\mathbf{x}, t) = \dfrac{\dot{x}^2}{2k} + \dfrac{x^2}{2}$, where $\mathbf{x} = [x_1, x_2]^T$, is positive definite.

Then, the time derivative of V along system trajectories of eq. (3.8) is given by:

$$\begin{aligned}
\dot{V}(\mathbf{x},t) &= \frac{\dot{x}\,\ddot{x}}{k} - \frac{\dot{x}^2}{2k^2}\,\dot{k} + x\,\dot{x} \\
&= -\frac{\dot{x}^2}{\Gamma\,\omega\,b}\left(a - \frac{\dot{\omega}}{\omega}\right) - \dot{x}\;x - \frac{\dot{x}^2}{2\,\Gamma\,\omega^2 b^2}\left(\dot{\omega}\;b + \omega\;\dot{b}\right) + x\;\dot{x} \\
&= -\frac{\dot{x}^2}{\Gamma\,\omega\,b}\left(a - \frac{\dot{\omega}}{\omega} + \frac{\dot{b}}{b}\right)
\end{aligned}$$

Let $\gamma = \frac{\omega}{b}$. Then, $\frac{\dot{\gamma}}{\gamma} = \frac{\dot{\omega}}{\omega} - \frac{\dot{b}}{b}$. Thus,

$$\dot{V}(\mathbf{x},t) = -\frac{\dot{x}^2}{\Gamma\,\omega\,b}\left(a - \frac{\dot{\gamma}}{\gamma}\right)$$

Hence, when

$$a - \frac{\dot{\gamma}}{\gamma} < 0 \;\rightarrow\; \dot{V} > 0$$

indicating that instability can occur. In practical applications, b is unknown and the above condition cannot be a-priori checked. Because b is unknown, the damping coefficient ζ receives only partial information about the rate of change of the spring constant. In some instances, like the case when ω decreases and b increases, this partial information is sufficient to ensure that the system is stable. Likewise, there may exist other cases when the partial information is insufficient and may result in the instability of the system. One such case is demonstrated in the numerical examples section.

Thus, stability for the process of parameter identification on a monotonic nonlinearity by using the local gradient method cannot be guaranteed. It can be conjectured that the local or other purely gradient method cannot guarantee stability or convergence of the parameter estimates. To see why this is the case, consider the following energy function:

$$V(x_1, x_2) = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2$$

which for the adaptive system in eq. (3.9) yields a time derivative

$$\dot{V} = -ax_1^2 + x_1 x_2 (b - \omega)$$

Although b is unknown, the bounds on b can be obtained. Therefore, by appropriately choosing ω, the term $b - \omega$ can always be made to be of a desired sign. However, the sign of x_2 is unknown, making any choice of ω futile in guaranteeing stability for the said system. The system equations imply that x_1 is the output of a first order filter with the input of $b\,x_2$. Thus, the sign of x_1 is not equal to the sign of $b\,x_2$ at every instant of time. Because the system relies on the sign of x_1 to adjust the parameter estimate error, x_2, this implies that the system is induced to adjust the parameter estimate in the wrong direction whenever $sign(x_1) \neq sign(b\,x_2)$. For the linear case, this problem is avoided because then $b = \omega$. However, when there is noise in the system $b - \omega$ is no longer zero, and instability can occur (see [22, 19]).

3.4 Numerical example

To highlight the kind of problems that may be encountered in adaptive systems with a monotonic nonlinearity, the following numerical case is presented. The system under consideration is given by:

$$\dot{x}_1 = -ax_1 + \phi(t)\,x_2 \tag{3.11}$$

$$\dot{x}_2 = -\omega(t)\,x_1 \tag{3.12}$$

The values of ϕ and ω are chosen such that $sign(\phi\,\omega) = 1$ always. Even though the above system represents a linearly parameterized system, from the graphs it can be seen that the particular values of ϕ and ω chosen in the simulation lead to instability.

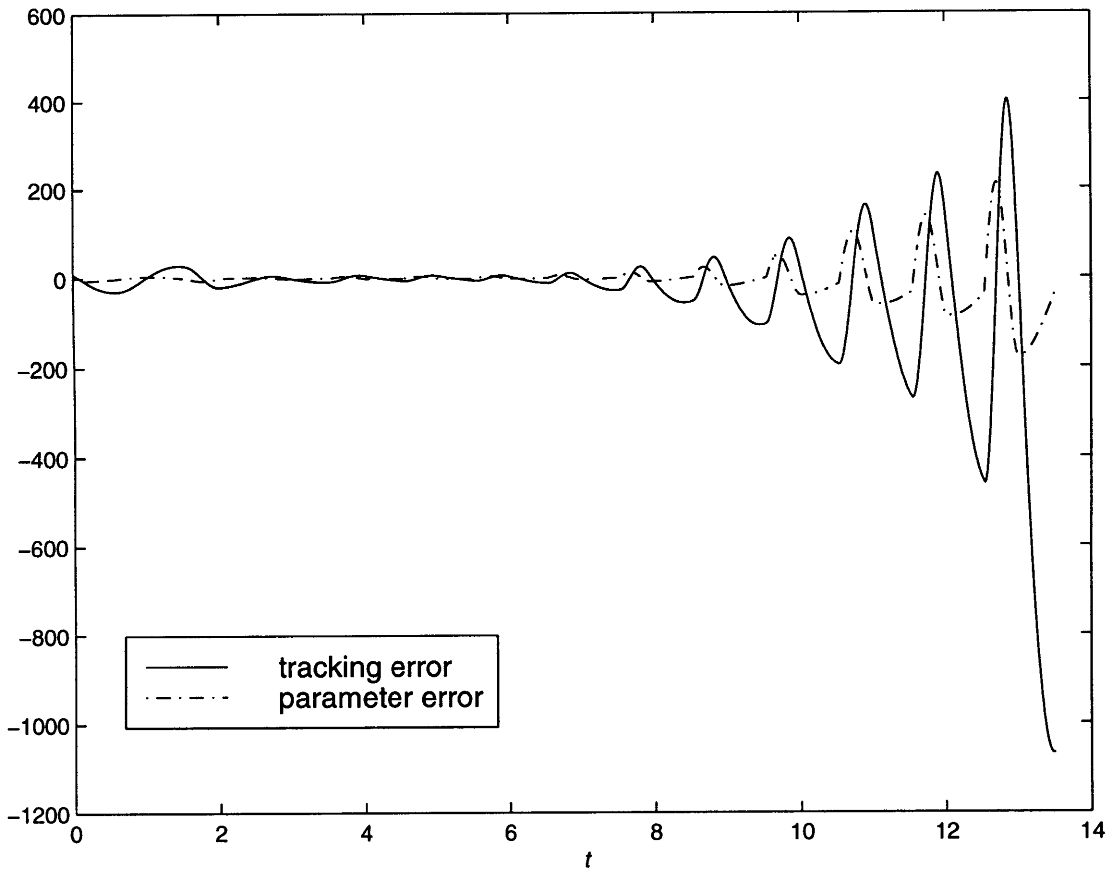

Figure 3-1: Tracking and parameter error.

The plot of the coefficient γ is given as well.

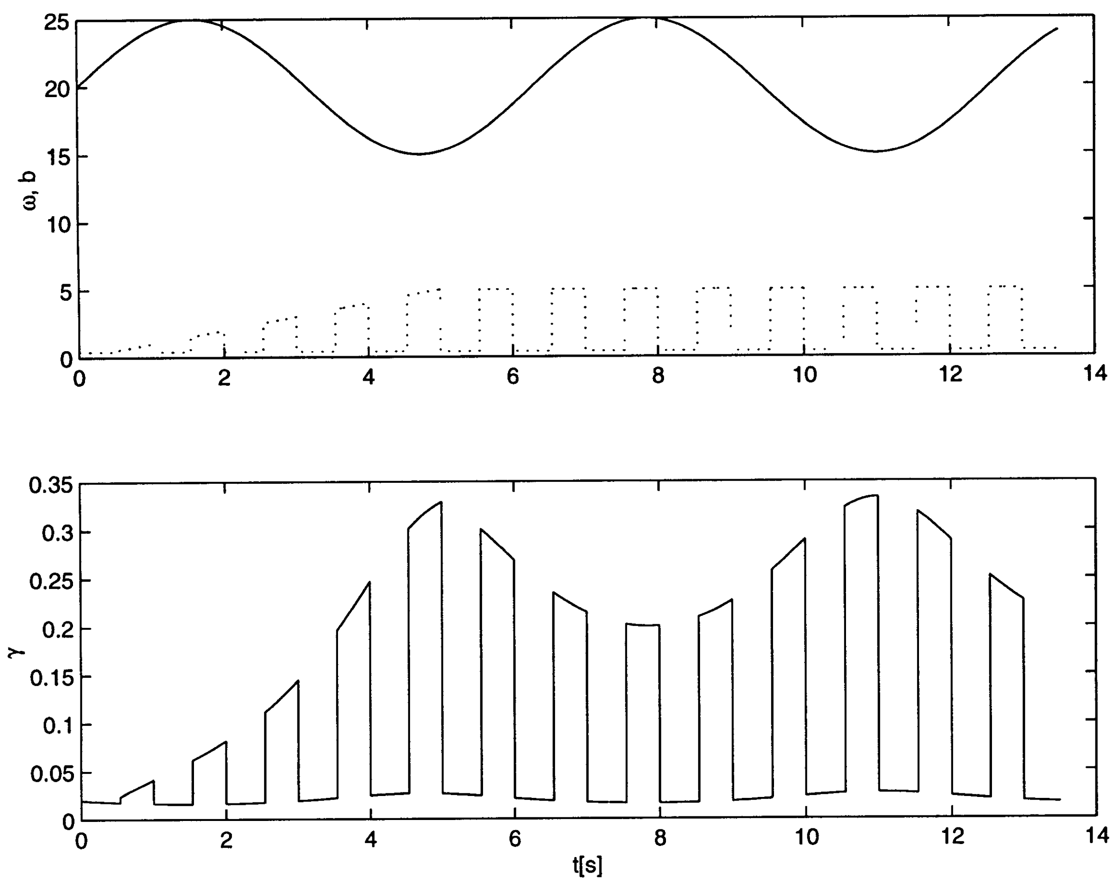

Figure 3-2: The top panel shows the values of ω(dotted line) and ϕ(solid line) used. The bottom panel depicts the value of the ratio γ.

Chapter 4

Monotonic Parameterization - Part II

4.1 Introduction

Monotonic nonlinear parameterizations are present in many systems of interest. Because linear functions are a special case of a general monotonic function, it can be expected that parameter estimation in linearly parametrized and and systems with monotonic nonlinear parametrization should have certain similarities. One important similarity is the fact that in both of these cases the correlation between the parameter error and observed error in the function value is of known sign. Knowing the sign of this correlation implies that the direction in which the parameter estimates should be adjusted is known. However in certain systems knowing this direction is not sufficient to ensure stability and convergence of parameter estimates. One such case was discussed in the previous chapter, where it was demonstrated that a gradient-based estimation technique on a monotonic nonlinearity can lead to instability when the function is present in a simple dynamic system. It was hypothesized that the principal reason the gradient-based estimation technique failed was due to the fact that direct measurements of the error in function value were not available at each instant of time. Rather, due to the dynamics and present inertia of the system, these measurements were delayed and distorted thereby allowing for the possibility of the system to become out of phase and diverge. This chapter deals with the problem of estimating the parameters in a monotonic function when there are no dynamics present between the nonlinearity and system measurements, thereby allowing the error in the function value to be directly observed at each instant of time.

This chapter limits the discussion of parameter estimation to a particular choice of the nonlinear monotonic parameterization: the sigmoidal function. The generic case of a sigmoidal function $f(.)$ is

$$f(x) \quad = \quad \frac{1-e^{-x}}{1+e^{-x}}$$

Specifically, the sigmoidal function is discussed in the context of its application in the "neural network" models. Neural network, in a fundamental form, is a a parallel system which consists of a number of processing subunits, called nodes. Each of these nodes processes the system input simultaneously, and by a superposition of the individual node outputs, the system output is formed. A node calculates its output by performing an input-output mapping through a nonlinearly parametrized function called the activation function. Based on the type of activation function used, there are many different types of neural networks. In this work, only such types of neural networks which have the sigmoidal nonlinearity as the activation function are treated.

It has been shown that neural network architectures possess powerful representation and approximation capabilities (see [11, 21]). The approximation capabilities are such that a neural network can approximate, up to a desired degree of accuracy, any sufficiently smooth function. The representation capabilities stem from the fact that a single, generic, neural network architecture can, through a process called network training, be made to approximate a desired function. The process of the neural network training consists of finding the correct values of the adjustable neural network parameters, such that the neural network can perform the desired function. Hence, the task of neural network training is a nonlinear parameter identification problem and in that context it is examined in this chapter. The classical approach, as with many similar nonlinear identification problems, is to use a gradient-based technique. Thus, the use of local gradient information for parameter updates is the basis of many of the popular neural network training methods: backpropagation [28], dynamic backpropagation [20]. For further neural network discussions, see [10]. In many various case studies, these techniques have been shown to exhibit satisfactory behavior and have enabled the neural networks to be applied to many different applications. However, because neural networks are nonlinearly parametrized, no global stability results exist. That is, in spite of the good results that have been obtained, there exist no a priori guarantees that for a particular task, the training method will converge. It is the goal of this chapter to, based on a reduced and simplified neural network model, present a methodology for establishing global convergence of gradient-based methods in neural network training.

4.2 Preliminaries

Before proceeding to the treatment of the problem, several definitions and a couple of useful Lemma are stated. Using the standard Eucledean geometry axioms of a line and length in a $\mathbb{R}^n$ space, the following concept of a vector quantity is introduced through the following definition

Definition 4.1 *A quantity is called a vector quantity if it has the following three properties:*

(i) Direction. *Direction is represented by a line along which the vector lies*

(ii) Orientation. *A line has no orientation. Along any line, there are only two possible orientations.*

(iii) Magnitude. *A scalar quantity, which graphically represents the length of the vector.*

Based on this definition, two vector quantities are said to be equal if and only if all three of their properties are identical. This will be used later in distinguishing two vectors which lie along the same direction, but have different orientations.

For the next definition, the following notation is used: S represents a bounded set, $s_1, s_2 \in S$ are any two elements of the set, $d(x, y)$ is the distance function between two points x and y, and $\rho(S) = \max\ d(s_1, s_2)$

Definition 4.2 *A bounded set S is called contiguous if for any point $s \in S$ and every ϵ, $0 < \epsilon < \rho(S)$, the set $B_\epsilon(s) = \{t \mid t \in S, d(t, s) < \epsilon\}$ is nonempty.*

Definition 4.3 *A time varying function $\alpha(t)$ is called a switching sequence between α_1 and α_2 if*

(i) $\alpha_1 \neq \alpha_2$,

(ii) for any t_a such that $\alpha(t_a) = \alpha_1$, there exists a t_b, $t_b > t_a$ such that $\alpha(t_b) = \alpha_2$ and

(iii) for any t_l such that $\alpha(t_l) = \alpha_2$, there exists a t_m, $t_m > t_l$ such that $\alpha(t_m) = \alpha_1$.

(iv) for each t_1,t_2 such that $\alpha(t_1) = \alpha_1$, $\alpha(t_2) = \alpha_2$ there exists a ρ such that for all ϵ, $0 < \epsilon \leq \rho$, the sets $B_i = \{t \mid |t - t_i| \leq \epsilon,\ \alpha(t) = \alpha_i\}$, $i = 1, 2$ are nonempty.

Definition 4.4 *A sigmoidal function defined as:*

$$f(x) = \frac{1 - e^{-x}}{1 + e^{-x}}$$

has the following properties

(i) $\lim_{x \to -\infty} f(x) = -1,\ \lim_{x \to \infty} f(x) = 0,\ f(-x) = f(x).$

(ii)

$$f'(x) = \frac{2e^{-x}}{(1 + e^{-x})^2} > 0 \qquad \forall x \in \mathbb{R}.$$

(iii)

$$f''(x) = \frac{2\left(e^{-x}\right)\left(e^{-x}-1\right)}{\left(1+e^{-x}\right)^2}; \qquad \begin{cases} f''(x) < 0, & x > 0 \\ f''(x) > 0, & x < 0 \\ f''(x) = 0, & x = 0 \end{cases}$$

Lemma 4.1 Let x, y be two positive real numbers such that $0 < y < x < 1$. Then the following inequality holds:

$$\frac{x}{y}\frac{(1+y)^2}{(1+x)^2} > 1$$

Proof: Pick a x and y which satisfy the conditions of the lemma. Suppose that for this choice of x, y the inequality does not hold, implying that:

$$x(1+y)^2 - y(1+x)^2 < 0.$$

Then:

$$\begin{aligned} 0 &> x + 2xy + xy^2 - y - 2yx - yx^2 \\ &> x - y + xy^2 - yx^2 \\ &> (x-y)(1-xy) \end{aligned}$$

Since $x > y$ and $(1 - xy) > 0$ because $x, y < 1$, the last line of the equation represents a contradiction. Due to the fact that x, y are arbitrary, Lemma 4.1 is established for any pair of x, y which satisfy the prescribed initial condition. ■

Lemma 4.2 *Let x, y, u, and w be positive real numbers and let the function $l(u,w,x,y)$ be defined as:*

$$l(u,w,x,y) = \frac{yu}{xw}\frac{(1+w)^2(1+x)^2}{(1+y)^2(1+u)^2}. \tag{4.1}$$

If the following hold,

(i) $0 < y < u < 1$

(ii) $y < x < u,\ y < w < u$

(iii) $\dfrac{yu}{xw} < 1$

then

$$l(u, w, x, y) \ < \ 1 \tag{4.2}$$

Proof: First, examine the function $\dfrac{(1+x)^2}{x}$ on the interval of interest, $x \in [0,1]$. Its first derivative is:

$$\frac{d}{dx}\left(\frac{(1+x)^2)}{x}\right) \ = \ \frac{2\,x\,(1+x) \ - \ (1+x)^2}{x^2} \ = \ \frac{x^2-1}{x^2} \qquad \forall x \in [0,1]$$

Therefore, the function $\dfrac{(1+x)^2}{x}$ is monotonically decreasing on $[0,1]$. Similarily, $\dfrac{(1+w)^2}{w}$ is also monotonically decreasing on $[0,1]$. Therefore, $l(u,w,x,y)$ is maximized when both x and w are minimized. However, because $\dfrac{xw}{yu} > 1$, it follows that $xw > yu$, implying that $x > \dfrac{yu}{w}$. So, for arbitrary y, u, the maximums of

$$k(x, w) = \frac{(1+x)^2}{x}\,\frac{(1+w)^2}{w}$$

are given by

$$\max\ k(x,w) \ = \ \max\ \frac{(1+x)^2}{x}\ \max\ \frac{(1+w)^2}{w}.$$

and they lie along the line $x = \dfrac{yu}{w}$. Along this line,

$$\begin{aligned} k_m(w) \ &= \ k(x,w)\big|_{x = \frac{yu}{w}} \\ &= \ \frac{(1+x)^2}{x}\,\frac{(1+w)^2}{w} \ = \ \frac{\left(1+\dfrac{yu}{w}\right)^2 (1+w)^2}{\dfrac{yu}{w}\,w} \ = \ \frac{\left(1+\dfrac{yu}{w}\right)^2 (1+w)^2}{yu} \end{aligned}$$

The extremums of $k_m(w)$ can be obtained by examining when its first derivative, k'_m, is equal to zero.

$$\begin{aligned} k'_m &= \left(1+\frac{yu}{w}\right)^2 2\,(1+w) \;+\; (1+w)^2\, 2\left(1+\frac{yu}{w}\right)\frac{-yu}{w^2} \\ &= 2\left(1+\frac{yu}{w}\right)(1+w)\left(1-\frac{yu}{w^2}\right) \end{aligned}$$

Thus, $k'_m(w) = 0$ when

$$1-\frac{yu}{w^2} = 0 \rightarrow w = \sqrt{yu}$$

When $w > \sqrt{yu}$, $1-\frac{yu}{w^2} > 0$, implying that $k'_m > 0$. For $w < \sqrt{yu}$, $1-\frac{yu}{w^2} < 0$, implying that $k'_m < 0$. Hence, the point $w = \sqrt{yu}$ is the minimum point of $k_m(w)$. Therefore, the function $k(x,w)$ is maximized at the ends of the allowed interval for (x,w). Because $xw > yu$, the two limit points of the interval are: $(x = y,\ w = u)$ and $(x = u,\ w = y)$. It can be easily verified that at these limit points the value of the function k is equal, $k(y,u) = k(u,y)$. Therefore,

$$\max_{x,w} k(x,w), \;<\; k(y,u) = k(u,y) = \frac{(1+u)^2)(1+y)^2}{yu}.$$

Since

$$l(u,w,x,y) \;=\; yu\,\frac{1}{(1+u)^2\,(1+y)^2}\,k(x,w)$$

it follows that for an arbitrary y, u, the maximum of $l(u,w,x,y)$ is when $k(x,w)$ is maximum. Thus, the maximum of $l(u,w,x,y)$ is

$$\begin{aligned} \max l(u,w,x,y) &= yu\,\frac{1}{(1+u)^2\,(1+y)^2}\,\max k(x,w) \\ &< yu\,\frac{1}{(1+u)^2\,(1+y)^2}\,\frac{(1+u)^2\,(1+y)^2}{yu} \\ &< 1. \end{aligned} \tag{4.3}$$

Hence, $l(u,w,x,y) < 1$ on $(x,w) \in [y,u]$. Since y, u were arbitrary, this holds for any y and u. ■

4.3 Stability in a two-node network

The system of interest is:

$$y = \sum_{i=1}^{N_n} g(u(t), \theta_i) \tag{4.4}$$

where u is a scalar function of time, θ_i, $i = 1, ..., N_n$ are scalar parameters with unknown values, and $g(u, \theta)$ is a sigmoidal function given by:

$$g(u,\theta) = \frac{1 - e^{-u\theta}}{1 + e^{-u\theta}}.$$

The general form of the system in eq. (4.4) satisfies the following assumptions:

(A1) $u > 0 \; \forall t$

(A2) $\theta_i > 0$ for all $i = 1, ..., N_n$.

(A3) The system output y is available for measurement at each instant of time.

The goal is to design an estimator which will allow the identification of the unknown parameter values in a stable manner such that the input-output behavior of the estimator is matched to that of the system for all possible choices of system input. The following estimator is proposed for such a task:

$$\widehat{y} = g(u, \widehat{\theta}_1) + g(u, \widehat{\theta}_2) \tag{4.5}$$

$$\dot{\widehat{\theta}}_i = -\widetilde{y}\, \nabla_{\widehat{\theta}_i} g(u, \widehat{\theta}_i) \qquad i = 1, ..., N_n \tag{4.6}$$

where $\widetilde{y} = \widehat{y} - y$ is the observed output error, and $\widehat{\theta}_i$ are the estimates, respectively, of θ_i, $i = 1, .., N_n$.

For the sake of mathematical tractability, only the $N_n = 2$ is considered in this chapter. In order to gain a better understanding of the behavior of the estimator in eqs. (4.5)- (4.6), the linear parametrization case when $g(u,\theta) = u\theta$ is considered. Then the estimator can be written as:

$$\begin{aligned} \tilde{y} &= u(\widehat{\theta}_1 - \theta_1 + \widehat{\theta}_2 - \theta_2) \\ \dot{\widehat{\theta}}_1 &= -\tilde{y}u \\ \dot{\widehat{\theta}}_2 &= -\tilde{y}u \end{aligned} \tag{4.7}$$

Clearly, for a given pair of parameters (θ_1, θ_2), $\tilde{y}$ is a function of $\widehat{\theta}_1, \widehat{\theta}_2$. For a particular choice of u, let $L_1 = \{(\widehat{\theta}_1^0, \widehat{\theta}_2^0) \mid \tilde{y}(\widehat{\theta}_1^0, \widehat{\theta}_2^0) = 0\}$. Thus, L_1 is the set of all parameter values for which which the estimator output is identical to that of the system. It is easy to verify that the set L_1 is

(i) independent of u

(ii) in $(\widehat{\theta}_1, \widehat{\theta}_2)$ space represented by a straight line which passes through the point (θ_1, θ_2) and has a slope of -1. That is:

$$\widehat{\theta}_2^0 = \theta_1 + \theta_2 - \widehat{\theta}_1^0.$$

Thus, for any values of $(\widehat{\theta}_1, \widehat{\theta}_2)$ which are in the set L_1, the estimator output would be indistinguishable from the system output for all u. In terms of parameter estimation, this means that there is an infinite number of points $\widehat{\theta}_1, \widehat{\theta}_2$ which produce the same input-output behavior for all possible choices of u. Therefore, in the general case, it cannot be expected of the estimator to be able to determine the specific values of θ_1, θ_2.

The convergence properties of the linear estimator are now determined by showing that the set L_1 is a globally attractive equilibrium region of the estimator for all allowed choices of u. Let π denote the region of $\widehat{\theta}_1$, $\widehat{\theta}_2$ coordinate space under consideration. That is,

$$\pi = \left\{(\widehat{\theta}_1, \widehat{\theta}_2) \mid \widehat{\theta}_1 > 0, \widehat{\theta}_2 > 0\right\}$$

Let $\mathbf{i_1}$ and $\mathbf{i_2}$ represent the unit vectors along the $\widehat{\theta}_1$ and $\widehat{\theta}_2$ axis. Then, let the change of the vector $\widehat{\theta} = [\widehat{\theta}_1, \widehat{\theta}_2]^T$ be denoted by $v = \dot{\widehat{\theta}}_1 \ \mathbf{i}_1 + \ \dot{\widehat{\theta}}_2 \ \mathbf{i}_2$. From eq. (4.7) $\dot{\widehat{\theta}}_1 = \dot{\widehat{\theta}}_2$ and it follows that the direction of v at a given point $P = (\widehat{\theta}_1, \widehat{\theta}_2)$ is independent of the choice u and lies along the line which passes through P with a slope of 1. The orientation of the vector v depends on the signs of $\dot{\widehat{\theta}}_1$ and $\dot{\widehat{\theta}}_2$. They can be examined by rewriting eq. (4.7) to obtain:

$$\dot{\widehat{\theta}}_1 = \dot{\widehat{\theta}}_2 = -u^2 h_l(\widehat{\theta}_1, \widehat{\theta}_2)$$

with $h_l(\widehat{\theta}_1, \widehat{\theta}_2) = (\widehat{\theta}_1 - \theta_1 + \widehat{\theta}_2 - \theta_2)$. Since u^2 is always positive, $sign(\dot{\widehat{\theta}}_1)$ and $sign\ (\ \dot{\widehat{\theta}}_2\ )$ depend solely on the coordinates of the adaptive system $(\widehat{\theta}_1, \widehat{\theta}_2)$ in the plane π. Therefore, the orientation of v at a given point in π is always the same as long as $u \neq 0$. To simplify further calculations, u is now chosen as a constant function $u(t) = const.$ It then follows that the time derivative of the positive quantity $J = \frac{1}{2}\widetilde{y}^2$ is:

$$\dot{J} = \widetilde{y} \left(\frac{\partial \widetilde{y}}{\partial \widehat{\theta}_1} \ \dot{\widehat{\theta}}_1 + \frac{\partial \widetilde{y}}{\partial \widehat{\theta}_2} \ \dot{\widehat{\theta}}_2 \right) = -2\widetilde{y}^2 u^2 < 0.$$

So, along the system trajectories J always decreases, implying that the vector v is oriented in a way which minimizes J. Since J has only a global minimum at $\widetilde{y} = 0$, it follows that the vector v always points towards the line L_1.

The fact that v is oriented towards L_1 implies that on any possible direction in π, there are at least two points with different orientations. Thus, L_1 divides the plane π into two distinct regions, with the distinguishing characteristic of each region being the orientation of v. Since the orientation is determined by the $sign(h_l)$, the two regions, S^+ and S^-, are defined as

$$\begin{aligned} S^+ &= \left\{ (\widehat{\theta}_1, \widehat{\theta}_2) \mid sign\left[h_l(\widehat{\theta}_1, \widehat{\theta}_2) \right] > 0 \right\} \\ S^- &= \left\{ (\widehat{\theta}_1, \widehat{\theta}_2) \mid sign\left[h_l(\widehat{\theta}_1, \widehat{\theta}_2) \right] < 0 \right\}. \end{aligned} \tag{4.8}$$

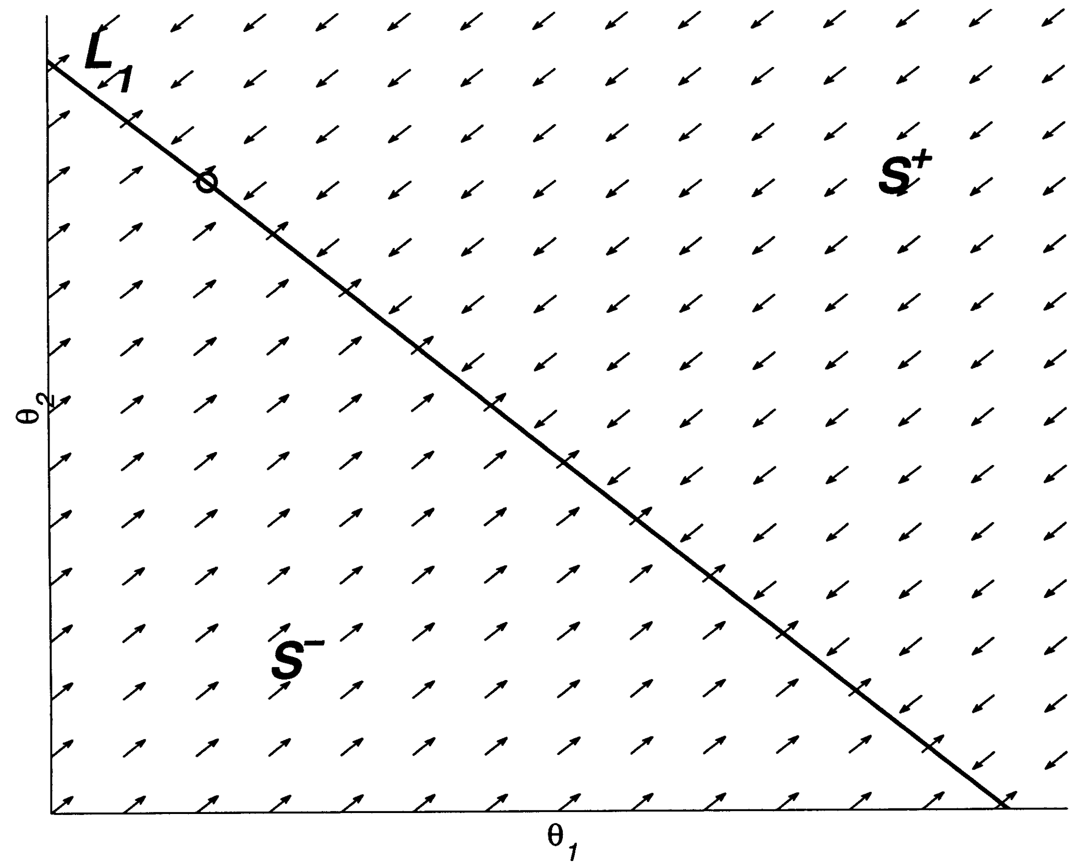

Figure 4-1: Phase plot for a linearly parameterized system. The set L_1 is given by the thick solid line, and the circle gives the location of the point (θ_1, θ_2).

Therefore,

$$\pi = S^+ \bigcup S^- \bigcup L_1$$

as is shown in Fig. (4-1).

Lemma 4.3 *For the linear version of the system in eq. (4.4) where* $g(u, \theta) = u\theta$, *which satisfies the assumptions* (A1)-(A3), *and the estimator in eq. (4.7), the set* L_1 *is a globally attractive equilibrium set.*

Proof: In proving the lemma, two elements need to be established: (a) L_1 is an invariant set and (b) outside of L_1, all system trajectories tend towards L_1. Case (a) is readily established, since on L_1, by definition, $\tilde{y} = 0$ and thus the magnitude of v is equal to zero. Hence, once in the set L_1, the system will always remain in the set L_1. Noting that $\pi - L_1 = S^+ \bigcup S^-$, it has been established that for case (b) v is always oriented towards L_1, thus implying that $\widehat{\theta}$ always tends towards L_1. Therefore, the lemma holds. ■

This concludes the analysis of the linear case. Following the same procedure, the discussion now turns to the parameter estimation problem in the system with sigmoidal parametrization in eq. (4.4). Let $A = (\theta_1, \theta_2)$ and $B = (\theta_2, \theta_1)$ be two points in π. Define the set L as:

$$L = \{A, B\}.$$

The conditions under which L is a globally attractive equilibrium for the adaptive system are stated in the following theorem.

Theorem 4.1 *For the system in eq. (4.4) which satisfies the assumptions* (A1)-(A3), *and the estimator in eqs. (4.5)- (4.6), the set L is a globally attractive equilibrium set if (a)* $\widehat{\theta}_1(t_0) \neq \widehat{\theta}_2(t_0)$ *and (b)* $u(t)$ *is a switching sequence between* u_1 *and* u_2*, where* $u_1 < u_2$.

Proof: By combining eqs. (4.4)- (4.6) the closed-loop adaptive system is obtained as:

$$\tilde{y} = g(u,\widehat{\theta}_1) + g(u,\widehat{\theta}_2) - g(u,\theta_1) + g(u,\theta_2) \qquad (4.9)$$

$$\dot{\widehat{\theta}}_i = -\tilde{y}\,\nabla_{\widehat{\theta}_i} g(u,\widehat{\theta}_i) = -\tilde{y}\,\frac{2}{u}\,\frac{e^{-u_i\widehat{\theta}_i}}{\left(1+e^{-u_i\widehat{\theta}_i}\right)^2} \qquad i = 1, ..., 2 \qquad (4.10)$$

If the condition *(a)* were not satisfied, then $\widehat{\theta}_1(t_0) = \widehat{\theta}_2(t_0)$ would imply that $\nabla_{\widehat{\theta}_1} g(u,\widehat{\theta}_1) = \nabla_{\widehat{\theta}_2} g(u,\widehat{\theta}_2)$ for all t. Let

$$F = \left\{(\widehat{\theta}_1,\widehat{\theta}_2) \mid \widehat{\theta}_1 = \widehat{\theta}_2\right\} \qquad (4.11)$$

Hence, the direction of $\widehat{\theta}$ would always be along the line defined by F. This would not hinder the possibility of convergence to L in the case that $\theta_1 = \theta_2$, but would rule out convergence for any $\theta_1 \neq \theta_2$. Since in practice the probability of $\theta_1 = \theta_2$ is zero, condition *(a)* is necessary.

For the rest of the discussion, it will be assumed that condition *(b)* is satisfied. In that case, it will be shown that convergence follows, thus making *(b)* a sufficient condition.

First, pick any $\theta_1 > 0$, $\theta_2 > 0$. Now, the sets M_1 and M_2 in π on which the estimator and system output are identical are given by:

$$M_1 = \left\{(\widehat{\theta}_1, \widehat{\theta}_2) \mid \tilde{y}(\widehat{\theta}_1, \widehat{\theta}_2, u_1) = 0\right\}$$
$$M_2 = \left\{(\widehat{\theta}_1, \widehat{\theta}_2) \mid \tilde{y}(\widehat{\theta}_1, \widehat{\theta}_2, u_2) = 0\right\}$$

The solution of the nonlinear equation

$$g(\widehat{\theta}_1, u) + g(\widehat{\theta}_2, u) = g(\theta_1, u) + g(\theta_2, u) \tag{4.12}$$

for $\widehat{\theta}_2$ in terms of $\widehat{\theta}_1$ and any value of u is obtained as:

$$\widehat{\theta}_2^0 = \frac{1}{u}\, ln \left[\frac{(1 + g(\theta_1, u) + g(\theta_2, u)) - g(\widehat{\theta}_1^0, u)}{(1 - g(\theta_1, u) - g(\theta_2, u)) + g(\widehat{\theta}_1^0, u)}\right] \tag{4.13}$$

Since $\tilde{y} = 0$ if and only if eq. (4.12) is satisfied, the sets M_1 and M_2 represent curves in the π plane which are obtained by substituting u_1 and u_2 for u, respectively, in eq. (4.13). The characteristics of the curve $\widehat{\theta}_2^0$ for a particular value of u are now analyzed by examining the first and second derivatives of the curve. To simplify the calculations, eq. (4.13) is rewritten using the following notation

$$\widehat{\theta}_2^0 = \frac{1}{u}\, ln \left(\frac{1 + a - h}{1 - a + h}\right) \tag{4.14}$$

where $a = g(\theta_1, u) + g(\theta_2, u)$, $h = g(\widehat{\theta}_1, u)$. Using these definitions, from eq. (4.12) and Definition 4.4 it follows that $0 \leq (a - h) \leq 1$. The first derivative of the curve in eq. (4.14) can be calculated by the chain rule as:

$$\frac{d\widehat{\theta}_2^0}{d\widehat{\theta}_1} = \frac{d\widehat{\theta}_2^0}{dh}\, h'\,.$$

Definition 4.4 states that $h' > 0$, and since

$$\frac{d\widehat{\theta}_2^0}{dh} = \frac{1}{u}\,\frac{1-a+h}{1+a-h}\,\frac{(1-a+h)(-1)-(1+a-h)}{(1-a+h)^2} = \frac{1}{u}\,\frac{-2}{1-(a-h)^2} < 0 \quad (4.15)$$

it follows that $\dfrac{d\widehat{\theta}_2^0}{d\widehat{\theta}_1} < 0\ \forall \widehat{\theta}_1$ and thus the curve $\widehat{\theta}_2^0(\widehat{\theta}_1)$ is monotonically decreasing. The curvature is analyzed by examining the second derivative:

$$\frac{d^2\widehat{\theta}_2^0}{d\widehat{\theta}_1^2} = \frac{d}{d\widehat{\theta}_1}\left(\frac{d\widehat{\theta}_2^0}{dh}\right) h' + \frac{d\widehat{\theta}_2^0}{dh}\, h'' = \frac{d^2\widehat{\theta}_2^0}{dh^2}\,(h')^2 + \frac{d\widehat{\theta}_2^0}{dh}\, h''$$

Definition 4.4 states that $h'' < 0$, $\dfrac{d\widehat{\theta}_2^0}{dh} < 0$ from eq. (4.15), and since

$$\frac{d^2\widehat{\theta}_2^0}{dh^2} = -2\,\frac{-1}{(1-(a-h)^2)^2}\,(-2)\,(a-h)\,(-1) = 4\,\frac{a-h}{(1-(a-h)^2)^2} > 0 \qquad \forall\, h$$

it follows that $\dfrac{d^2\widehat{\theta}_2^0}{d\widehat{\theta}_1^2} > 0\ \forall\widehat{\theta}_1$. Because the value for u in the above calculations was an arbitrary positive quantity, it follows that for any u_1, $u_2 > 0$ the corresponding curves M_1, M_2 are monotonically decreasing convex functions. This implies that they can intersect on at most two points in the π plane. It is easily obtained that

$$M_1 \bigcap M_2 = L \qquad (4.16)$$

since in the case when $\widehat{\theta} = A$ or $\widehat{\theta} = B$, eq. (4.12) is identically satisfied for all possible choices of u. Thus, L is an equilibrium set for both when $u = u_1$ and $u = u_2$. The two curves are shown in Fig. 4-2.

Having defined the characteristics of M_1 and M_2, the trajectory of the parameter estimate vector $\widehat{\theta}$ with respect to these curves is now investigated. First, let T_1 be an arbitrary time interval such that $T_1 = \{t \mid u(t) = u_1\}$. Define J_1 as $J_1 = \frac{1}{2}\,\tilde{y}^2$. Then, on the interval T_1,

$$\dot{J}_1 = \tilde{y} \sum_{i=1}^{2} \nabla_{\widehat{\theta}_i} g(u_1, \widehat{\theta}_i) \ \dot{\widehat{\theta}}_i \ = \ -\tilde{y}^2 \sum_{i=1}^{2} \left[\nabla_{\widehat{\theta}_i} g(u_1, \widehat{\theta}_i)\right]^2 \ < \ 0 \tag{4.17}$$

As in the linear case, the negative derivative of J_1 implies that at every time instant t, $t \in T_1$, the velocity vector v, $v = \dot{\widehat{\theta}}$, of the system is oriented towards the curve M_1. From eq. (4.10) it follows that:

$$sign(\dot{\widehat{\theta}}_1) = sign(\dot{\widehat{\theta}}_2) = -sign\left(\frac{\tilde{y}}{u}\right).$$

Therefore, the curve M_1 divides the plane π into two distinct regions which can be defined as:

$$\begin{aligned} S_1^+ &= \left\{(\widehat{\theta}_1, \widehat{\theta}_2) \mid sign\left(\frac{\tilde{y}(\widehat{\theta}_1, \widehat{\theta}_2, u_1)}{u_1}\right) > 0\right\} \\ S_1^- &= \left\{(\widehat{\theta}_1, \widehat{\theta}_2) \mid sign\left(\frac{\tilde{y}(\widehat{\theta}_1, \widehat{\theta}_2, u_1)}{u_1}\right) < 0\right\} \end{aligned} \tag{4.18}$$

Thus, $\pi = S_1^+ \cup S_1^- \cup M_1$. Similar analysis can be carried on for an arbitrary time interval T_2 such that $T_2 = \{t \mid u(t) = u_2\}$. On T_2 it can be obtained that J_2, $J_2 = \frac{1}{2}\tilde{y}^2$, is decreasing, implying that M_2 divides the π plane into the following two distinct regions:

$$\begin{aligned} S_2^+ &= \left\{(\widehat{\theta}_1, \widehat{\theta}_2) \mid sign\left(\frac{\tilde{y}(\widehat{\theta}_1, \widehat{\theta}_2, u_2)}{u_2}\right) > 0\right\} \\ S_2^- &= \left\{(\widehat{\theta}_1, \widehat{\theta}_2) \mid sign\left(\frac{\tilde{y}(\widehat{\theta}_1, \widehat{\theta}_2, u_2)}{u_2}\right) < 0\right\}. \end{aligned} \tag{4.19}$$

Likewise, $\pi = S_2^+ \cup S_2^- \cup M_2$.

Now define:

$$S_E = (S_1^+ \bigcap S_2^+) \bigcup (S_1^- \bigcap S_2^-) \tag{4.20}$$

$$S_D = \left\{(S_1^- \bigcap S_2^+) \bigcup (S_1^+ \bigcap S_2^-)\right\} \bigcup M_1 \bigcup M_2 \tag{4.21}$$

These sets are depicted in Fig. 4-2. Let P_E be any point in the set S_E, $P_E \in S_E$. According to eqs. (4.18)- (4.20), the orientation of v at the point P_E would be the same for all time. On the contrary, for any point P_D, $P_D \in S_D$, the orientation of v would be the opposite during time intervals T_1 and T_2.

Starting at any point P, the system would move towards M_1 during a time interval T_1, and move towards M_2 during a time interval T_2. While the system is in the set S_E, the directions during these two time intervals are equal, meaning that the system would converge to a set that contains both M_1 and M_2. That is, starting at any point in S_E the system would converge to S_D since S_D contains both M_1 and M_2. Because u is a switching sequence, the system enters S_D in a finite time.

Before examining the behavior of the system in S_D, another feature of the system is noted. Namely, the system behavior is symmetric with respect to the line F of eq. (4.11). This is because the nonlinearities in the system governing equation, eq. (4.12) appear additively. Since the operands of addition are commutable, it makes no difference if $\widehat{\theta}_1$ and $\widehat{\theta}_2$ exchange places, indicating symmetry of the problem. Therefore, without loss of generality, system motion is only examined on the set S_D^h,

$$S_D^h = \left\{(\widehat{\theta}_1, \widehat{\theta}_2) \mid (\widehat{\theta}_1, \widehat{\theta}_2) \in S_D,\, \widehat{\theta}_1 < \widehat{\theta}_2\right\} \tag{4.22}$$

Also, it is taken that the point A is in the set S_D^h, $A \in S_D^h$.

It can be shown that the set S_D, and therefore S_D^h are time invariant sets for the system dynamics. This is shown by contradiction. Suppose that S_D was not time invariant. Without loss of generality, this supposition implies that the system, starting in S_D at the beginning of a time interval T_1, was not in S_D for some time instant t_1^*. It has been established that the system during T_1 would tend towards M_1. So if the system is outside of the set S_D at time t_1^*, it implies that it must have at some time $t_{1_0} < t_1^*$, $t_{1_0} \in T_1$ crossed the curve M_1 with a non-zero velocity vector v. However, this is impossible, since by definition $v = 0$ on M_1. This contradiction establishes the claim that S_D is time invariant.

All that is now left to do in establishing the global attractiveness of L is to establish that for any point in S_D^h, the point A is attractive. A very useful insight into the behavior of the system dynamics on S_D^h can be gained by examining what happens to the value of the angle that the direction of the velocity vector makes with the $\widehat{\theta}_1$ axis, ie. the direction of $\mathbf{i_1}$, when the input value u changes from u_1 to u_2. Let $\beta(t)$ be such an angle at time t given by

$$\beta(t) \quad = \quad \arctan \frac{\dot{\widehat{\theta}}_2 \left(\widehat{\theta}_1(t), \widehat{\theta}_2(t), u(t)\right)}{\dot{\widehat{\theta}}_1 \left(\widehat{\theta}_1(t), \widehat{\theta}_2(t), u(t)\right)}. \tag{4.23}$$

For a particular point $P = (\widehat{\theta}_1, \widehat{\theta}_2)$, let

$$\begin{aligned} \beta_1 &= \arctan \frac{\dot{\widehat{\theta}}_2 \, (\widehat{\theta}_1, \widehat{\theta}_2, u_1)}{\dot{\widehat{\theta}}_1 \, (\widehat{\theta}_1, \widehat{\theta}_2, u_1)} \\ \beta_2 &= \arctan \frac{\dot{\widehat{\theta}}_2 \, (\widehat{\theta}_1, \widehat{\theta}_2, u_2)}{\dot{\widehat{\theta}}_1 \, (\widehat{\theta}_1, \widehat{\theta}_2, u_2)} \end{aligned} \tag{4.24}$$

Thus,

$$\begin{aligned} \beta_1 &= \arctan \frac{\nabla_{\widehat{\theta}_2} \, g(\widehat{\theta}_2, u_1)}{\nabla_{\widehat{\theta}_1} \, g(\widehat{\theta}_1, u_1)} = \arctan \frac{-\frac{\widetilde{y}}{u_1} \frac{2e^{-u_1\widehat{\theta}_2}}{(1+e^{-u_1\widehat{\theta}_2})^2}}{-\frac{\widetilde{y}}{u_1} \frac{2e^{-u_1\widehat{\theta}_1}}{(1+e^{-u_1\widehat{\theta}_1})^2}} \\ \beta_2 &= \arctan \frac{\nabla_{\widehat{\theta}_2} \, g(\widehat{\theta}_2, u_2)}{\nabla_{\widehat{\theta}_1} \, g(\widehat{\theta}_1, u_2)} = \arctan \frac{-\frac{\widetilde{y}}{u_2} \frac{2e^{-u_2\widehat{\theta}_2}}{(1+e^{-u_2\widehat{\theta}_2})^2}}{-\frac{\widetilde{y}}{u_2} \frac{2e^{-u_2\widehat{\theta}_1}}{(1+e^{-u_2\widehat{\theta}_1})^2}} \end{aligned} \tag{4.25}$$

$$\begin{aligned} \beta_1 &= \arctan \frac{e^{-u_1\widehat{\theta}_2}}{e^{-u_1\widehat{\theta}_1}} \frac{(1+e^{-u_1\widehat{\theta}_1})^2}{(1+e^{-u_1\widehat{\theta}_2})^2} \\ \beta_2 &= \arctan \frac{e^{-u_2\widehat{\theta}_2}}{e^{-u_2\widehat{\theta}_1}} \frac{(1+e^{-u_2\widehat{\theta}_1})^2}{(1+e^{-u_2\widehat{\theta}_2})^2} \end{aligned} \tag{4.26}$$

Letting

$$x = e^{-u_1\widehat{\theta}_1} \; y = e^{-u_1\widehat{\theta}_2} \tag{4.27}$$

it follows that $0 < y < x < 1$ on S_D^h. Then, by Lemma 4.1, it follows that $\beta_1, \beta_2 < 45°$ everywhere on S_D^h. From the aspect of establishing convergence, a more interesting quantity is the ratio of $\frac{\beta_2}{\beta_1}$ for any point on S_D^h. The relative magnitude of this ratio can be determined by examining the following ratio:

$$\gamma = \frac{e^{-u_1\widehat{\theta}_1}\, e^{-u_2\widehat{\theta}_2}}{e^{-u_1\widehat{\theta}_2}\, e^{-u_2\widehat{\theta}_1}} \frac{(1+e^{-u_2\widehat{\theta}_1})^2\,(1+e^{-u_1\widehat{\theta}_2})^2}{(1+e^{-u_1\widehat{\theta}_1})^2\,(1+e^{-u_2\widehat{\theta}_2})^2} \tag{4.28}$$

By letting

$$s = e^{-u_1\widehat{\theta}_1} \; y = e^{-u_2\widehat{\theta}_2} \; x = e^{-u_2\widehat{\theta}_1} \; w = e^{-u_1\widehat{\theta}_2} \tag{4.29}$$

eq. (4.28) takes the same form as eq. (4.1):

$$\gamma == \frac{ys}{xw} \frac{(1+w)^2\,(1+x)^2}{(1+y)^2\,(1+s)^2}. \tag{4.30}$$

The following can also be concluded about the quantities s, y, x, w of eq. (4.29) on the set S_D^h:

(i) $0 < u_1\widehat{\theta}_1 < u_2\widehat{\theta}_2 < 1$, implying that $0 < y < s < 1$.

(ii) $u_2\widehat{\theta}_2 > u_2\widehat{\theta}_1 > u_1\widehat{\theta}_1$, implying that $y < x < s$. Also, $u_2\widehat{\theta}_2 > u_1\widehat{\theta}_2 > u_1\widehat{\theta}_1$, implying that $y < w < s$.

(iii) Let r be such that:

$$e^r = \frac{e^{-u_1\widehat{\theta}_1}\, e^{-u_2\widehat{\theta}_2}}{e^{-u_1\widehat{\theta}_2}\, e^{-u_2\widehat{\theta}_1}} \tag{4.31}$$

Thus,

$$r \;=\; -u_1\widehat{\theta}_1 - u_2\widehat{\theta}_2 + u_1\widehat{\theta}_2 + u_2\widehat{\theta}_1$$

$$
\begin{aligned}
&= \widehat{\theta}_2 (u_1 - u_2) + \widehat{\theta}_1 (u_2 - u_1) \\
&= (\widehat{\theta}_2 - \widehat{\theta}_1)(u_1 - u_2) \\
&< 0.
\end{aligned}
\tag{4.32}
$$

Hence, $\dfrac{ys}{xw} < 1$.

Therefore, by applying Lemma 4.2, it follows that $\gamma < 1$, implying that $\beta_1 > \beta_2$. This relationship holds for any point in S_D^h.

It will be now demonstrated how $\beta_1 > \beta_2$ on S_D^h, together with the fact that u us a switching sequence, imply convergence. A new quantity which will facilitate the proof of convergence is now defined. Suppose that, starting at a given time instant at a certain point in S_D^h, the input value u was kept constant at either u_1 or u_2 from that time instant onwards. Then, the system would converge to a point either in M_1 or M_2. This convergence point can be calculated by integrating forward in time the system dynamics equations (eq. (4.10)) which govern the change of $\widehat{\theta}$. Thus, given a starting point denoted by $\widehat{\theta}(t)$, let

$$
\begin{aligned}
N_1(t) &= \lim_{\tau \to \infty} \widehat{\theta}\left(\tau; \widehat{\theta}(t); u_1\right) \\
N_2(t) &= \lim_{\tau \to \infty} \widehat{\theta}\left(\tau; \widehat{\theta}(t); u_2\right)
\end{aligned}
\tag{4.33}
$$

where $N_1(t) = [\widehat{\theta}_{11}(t)\ \widehat{\theta}_{21}(t)]^T$, $N_2(t) = [\widehat{\theta}_{12}(t)\ \widehat{\theta}_{22}(t)]^T$ represent the limit points of the vector $\widehat{\theta} = [\widehat{\theta}_1\ \widehat{\theta}_2]^T$, starting with the initial condition $\widehat{\theta}(t)$ and with $u = u_1$, $u = u_2$ respectively. Let each set in the sequence of sets $C_1(t)$ and $C_2(t)$ represent the trajectory along which the system would converge to each distinct $N_1(t)$ and $N_2(t)$, respectively. Thus,

$$
\begin{aligned}
C_1(t) &= \left\{P \mid P \in S_D^h,\ \lim_{\tau \to \infty} \widehat{\theta}(\tau; P; u_1) = N_1(t)\right\} \\
C_2(t) &= \left\{P \mid P \in S_D^h,\ \lim_{\tau \to \infty} \widehat{\theta}(\tau; P; u_2) = N_2(t)\right\}
\end{aligned}
\tag{4.34}
$$

Now consider the following two cases:

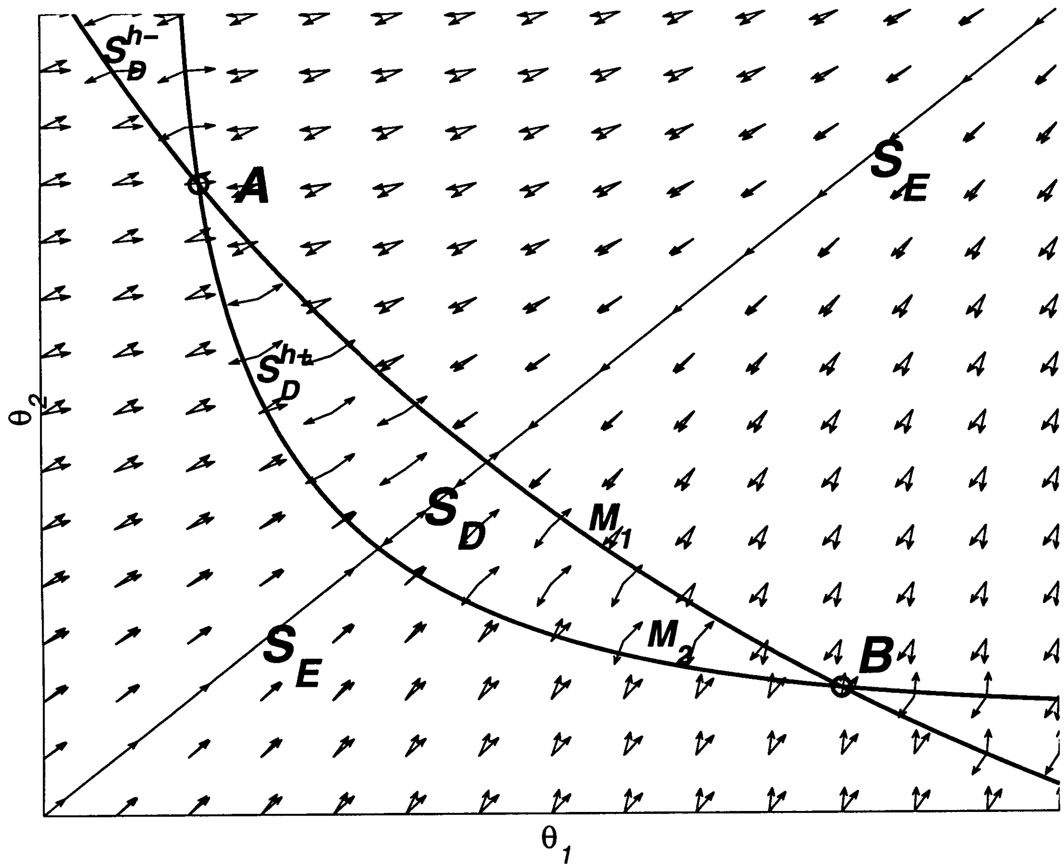

Figure 4-2: Phase plot for a two-parameter system with sigmoidal parameterization using two different values of u.

(a) The system is in the set S_D^{h-}, where S_D^{h-} is defined as:

$$S_D^{h-} = \left\{ (\widehat{\theta}_1, \widehat{\theta}_2) \mid sign\left[\dot{\widehat{\theta}}_1 \ (\widehat{\theta}_1, \widehat{\theta}_2, u_1) \right] = sign\left[\dot{\widehat{\theta}}_2 \ (\widehat{\theta}_1, \widehat{\theta}_2, u_1) \right] = -1 \right\}$$

(b) The system is in the set S_D^{h-}, where S_D^{h-} is defined as:

$$S_D^{h+} = \left\{ (\widehat{\theta}_1, \widehat{\theta}_2) \mid sign\left[\dot{\widehat{\theta}}_1 \ (\widehat{\theta}_1, \widehat{\theta}_2, u_1) \right] = sign\left[\dot{\widehat{\theta}}_2 \ (\widehat{\theta}_1, \widehat{\theta}_2, u_1) \right] = 1 \right\}$$

The following is then also true: $S_D^h = S_D^{h-} \bigcup S_D^{h+} \bigcup M_1 \bigcup M_2$ and $S_D^{h-} \bigcap S_D^{h+} = \emptyset$. Also the point A is a limit point for both sets.

In case *(a)*, it will be shown that $\widehat{\theta}_{21}(t)$ and $\widehat{\theta}_{22}(t)$ are decreasing functions of time. Suppose the system is starting its motion at time t_0 at some point $\widehat{\theta}(t_0)$ for which the condition of case *(a)* is satisfied. With the point $\widehat{\theta}(t_0)$ are associated the points $N_1(t_0) \in M_1$ and $N_2(t_0) \in M_2$, and the corresponding trajectories $C_1(t_0)$ and $C_2(t_0)$, respectively, as defined in eqs. (4.33)- (4.34). Suppose further that t_0 is contained in a contiguous interval T_2 on which $u(t) = u_2$, $t \in T_2$. Then, the system moves along the trajectory $C_2(t_0)$. Since for all t, $t \in T_2$, $sign(\dot{\widehat{\theta}}_1) = sign(\dot{\widehat{\theta}}_2) = 1$ and since $0 < \beta_2 < \beta_1 < 45°$ for any point in S_D^h, the curve representing the trajectory $C_2(t_0)$ remains always bounded from above by the curve representing trajectory $C_1(t_0)$ and from below by the line which passes through the point $\widehat{\theta}(t_0)$ and is parallel to the $\widehat{\theta}_1$ axis and its unit vector $\mathbf{i}_1$.

Because the system is moving on $C_2(t_0)$ on T_2, $\widehat{\theta}_{22}(t)$ is constant on T_2. Meanwhile, suppose that there exists a time instant t^* such that $\widehat{\theta}_{21}(t^*) > \widehat{\theta}_{21}(t_0)$. Because of the relative position of $C_1(t_0)$ and $C_2(t_0)$, this supposition implies that the curve of the trajectory $C_1(t^*)$ would have to cross the curve of the trajectory $C_1(t_0)$. Let K be the point at which they cross. The fact that they cross means that at K, the velocity vector v would have two different values for the same coordinates and the same value of u, which is impossible. Thus, the two trajectories cannot cross, and hence t^* does not exist. Therefore, on T_2, $\widehat{\theta}_{21}$ is strictly decreasing. In the case that t_0 is contained in a contiguous interval T_1 on which $u(t) = u_1$, it would follow that $\widehat{\theta}_{21}(t)$ is constant on T_1, and an argument similar to the one above can be made to show that $\widehat{\theta}_{22}(t)$ is decreasing on T_1. By Definition 4.3 of a switching sequence, each interval on which $u = u_1$ is followed by an interval on which $u = u_2$, and vice versa. Therefore, both $\widehat{\theta}_{21}$ and $\widehat{\theta}_{22}$ decrease with time.

In case *(b)*, it will be shown that $\widehat{\theta}_{21}(t)$ and $\widehat{\theta}_{22}(t)$ are increasing functions of time by a similar kind of argument as was applied in case *(a)*. The difference in the two cases is the fact that in case *(b)*, $sign\left(\dot{\widehat{\theta}}_1\ (\theta, u_1)\right) = sign\left(\dot{\widehat{\theta}}_2\ (\theta, u_1)\right) = 1$, implying that $sign\left(\dot{\widehat{\theta}}_1\ (\theta, u_2)\right) = sign\left(\dot{\widehat{\theta}}_2\ (\theta, u_2)\right) = -1$. Therefore, the curve of the trajectory $C_2(t_0)$ is bounded from above by the line parallel to the direction of $\mathbf{i}_1$, and bounded from below by the curve of the trajectory $C_1(t_0)$. Thus, if on T_1 there existed a t^* such that $\widehat{\theta}_{22}(t^*) < \widehat{\theta}_{22}(t_0)$, it would imply that the curve of the trajectory $C_2(t^*)$ would have to cross the curve of the trajectory $C_2(t_0)$. Since this cannot happen, $\widehat{\theta}_{22}$ is always increasing on T_1. Likewise, $\widehat{\theta}_{21}$ is always increasing on T_2. Since u is a switching sequence, the functions $\widehat{\theta}_{21}(t)$ and $\widehat{\theta}_{22}(t)$ are increasing functions of time.

It is interesting to note that, because the set S_D^h is invariant and bounded, $\widehat{\theta}_{11}$ and $\widehat{\theta}_{12}$ remain bounded in both cases *(a)* and *(b)*. Based on the proven behavior above, it follows that in case *(a)*, the system swings between heading towards M_1 and M_2. The swings are oriented in such a way that the system $\widehat{\theta}_2$ coordinate decreases with time. In case *(b)*, the $\widehat{\theta}_2$ coordinate is increasing with time. Since

$$\theta_2 = \min_{\widehat{\theta}_2} P(\widehat{\theta}_1, \widehat{\theta}_2) \;\; \forall P \in S_D^{h-}$$

and

$$\theta_2 = \max_{\widehat{\theta}_2} P(\widehat{\theta}_1, \widehat{\theta}_2) \;\; \forall P \in S_D^{h+}$$

it then follows that when u is a switching sequence, the system will converge to the point A starting from anywhere in the set S_D^h.

Hence, for any starting point in S_E, the system converges in finite time to S_D. Starting from any point in S_D, the system converges asymptotically to either point A or B. Thus, the theorem is established. ■

An interesting comparison can be made between the proof of Theorem 4.1 and the proof of its linear counterpart given in Lemma 4.3. The proof of stability for the linear system utilizes the standard approach currently taken in adaptive control, and control of nonlinear systems in general. The standard approach consists of taking the distance of $\widehat{\theta}$ from θ as a measure of the parameter error, and then showing that the measure is always nonincreasing. Such an approach is not beneficial for demonstrating stability in the nonlinear case. First, there is a problem in how to define the point from which the distance is measured to the current estimate $\widehat{\theta}$. From the proof of the Theorem, it follows that the point of convergence depends upon the initial estimate. If the initial estimate is such that $\widehat{\theta}_2(t_0) > \widehat{\theta}_1(t_0)$, the system converges to point A, and if $\widehat{\theta}_2(t_0) < \widehat{\theta}_1(t_0)$, the system converges to point B. If the distance for measuring the parameter error was, say, point A, it could be observed that the measure would be increasing during the course of adaptation if $\hat{\theta}_2(t_0) < \hat{\theta}_1(t_0)$. Suppose now that this problem of picking the originating point for the measure of parameter error was solved so that it was picked according to the initial condition. Then, if it was adhered to the strict notion of having the measure be nonincreasing, a second problem would appear. As can be seen from the geometry of the problem, on certain intervals of time, the measure of the parameter error can actually be increasing. Therefore, in order to prove Theorem 4.1, the strict condition that the measure is nonincreasing has to be relaxed. Rather, the asymptotic behavior is examined, implying that despite the fact that the measure may sometimes increase on bounded intervals of time, the overall behavior is such that it is decreasing. These remarks are illustrated in a numerical example in the next section.

4.4 Numerical Example

This section presents a numerical example which demonstrates the kind of system behavior discussed in the previous section. For the sake of clarity, the overall estimator dynamics are restated here.

$$\tilde{y} = g(u,\hat{\theta}_1) + g(u,\hat{\theta}_2) - g(u,\theta_1) + g(u,\theta_2) \tag{4.35}$$

$$\dot{\hat{\theta}}_i = -\tilde{y}\,\nabla_{\hat{\theta}_i} g(u,\hat{\theta}_i) = -\tilde{y}\,\frac{2}{u}\,\frac{e^{-u_i\hat{\theta}_i}}{\left(1+e^{-u_i\hat{\theta}_i}\right)^2} \qquad i = 1,...,2 \tag{4.36}$$

The example consists of simulation runs of the above adaptive system with four different initial conditions, corresponding to four different sets of estimates of θ. The following set values are used in the simulation runs.

$$\theta = [1.7813, 4.9298]^T$$

$$\hat{\theta}(t_0) = \quad 1)\begin{cases} 4 \\ 4.5 \end{cases} \quad 2)\begin{cases} 0.5 \\ 6.0 \end{cases} \quad 3)\begin{cases} 3 \\ 1 \end{cases} \quad 4)\begin{cases} 8 \\ 3 \end{cases}$$

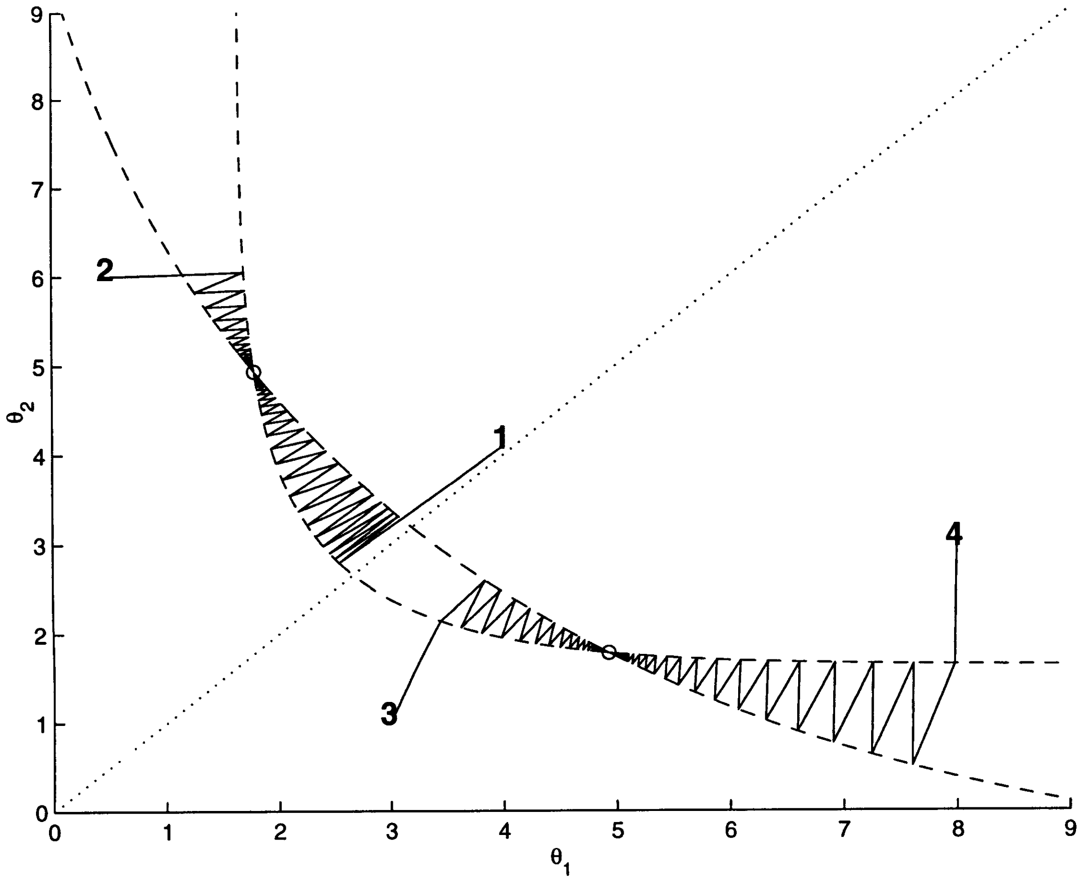

Figure 4-3: The phase plane behavior(solid line) of the adaptive system for four different initial conditions. The sets on which $\widetilde{y} = 0$ are given by the dashed line. For reference, the line K on which lie all points with equal values of the coordinates is given by the dotted line.

The input u was a switching sequence between $u_1 = 0.3$ and $u_2 = 0.8$. The simulation results are given in Figs. 4-3, 4-4. It can be seen that for all initial conditions, the system does converge to either point $A = [1.7813, 4.9298]$ or to point $B = [4.9298, 1.7813]$. Fig. 4-4 shows the behavior of the measure of the parameter error defined as $J = ||\widehat{\theta} - \theta^*||^2$, where $\theta^* = A$, or $\theta^* = B$, depending on the initial condition. It can be seen that even with such a choice of θ^*, J is, as expected, increasing on certain intervals of time. However, the asymptotic behavior is such that $J \to 0$ as $t \to \infty$.

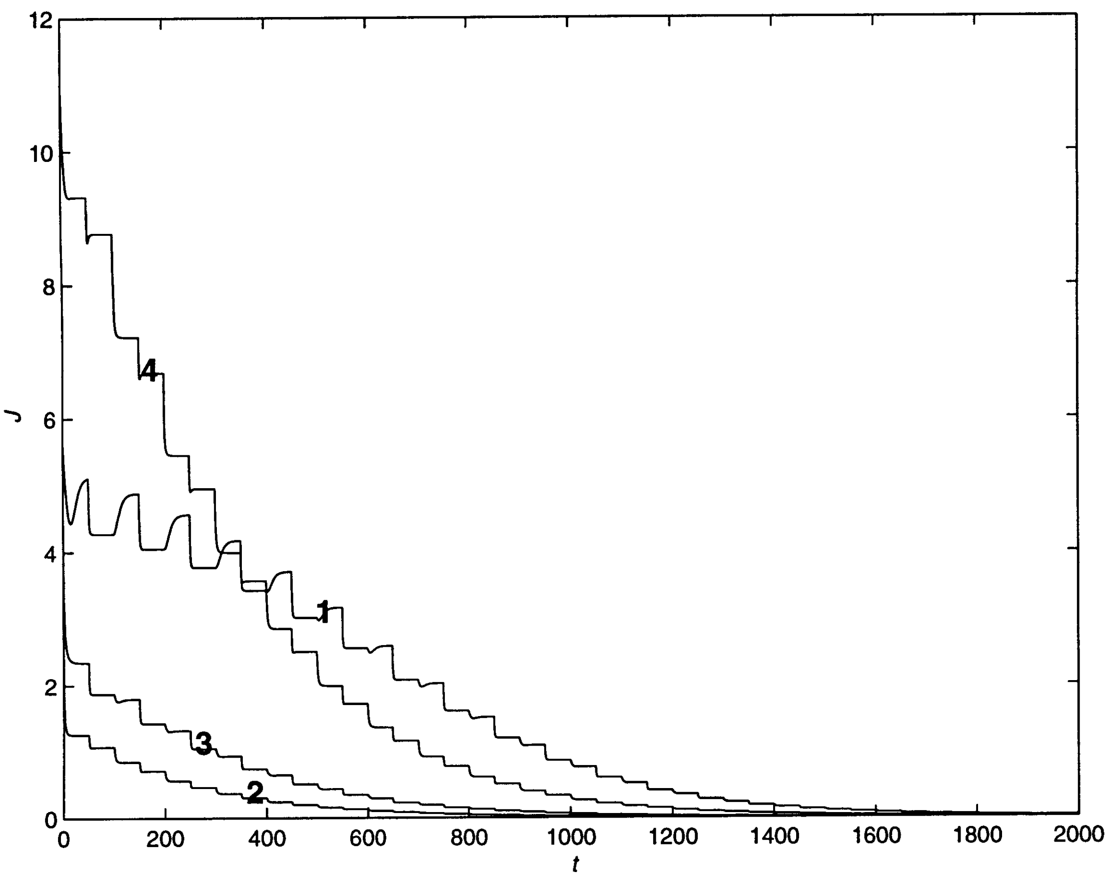

Figure 4-4: The measure of the parameter error defined as $J = \|\widehat{\theta} - \theta^*\|^2$ for four different initial conditions. θ^* is chosen according to the initial estimate $\widehat{\theta}(t_0)$.

Chapter 5

Conclusions

The problem of linear adaptive control has been extensively studied for a number of years, and powerful results can be found in the present literature. In contrast, few results exist in the field of nonlinear adaptive control, especially adaptive control of systems with nonlinear parameterization. The approaches to modify and linearly reparameterize nonlinearly parameterized systems are, for majority of the cases, successful only at achieving local results. However, these approaches have a serious drawback because they introduce a large number of virtual parameters which have no physical meaning. Thus, if the goal is to acquire information about the state of the system by learning the values of its physical parameters, linear reparameterization is inadequate. Hence, if adequate global performance is to be achieved, the problem of nonlinear parameterization must be addressed.

This thesis attempted to address the problem of parameter convergence in systems with two types of nonlinear parameterization, *(i)* convex/concave and *(ii)* monotonic parameterization. For certain convex/concave parameterizations, it was shown under what conditions and external inputs globally stable control and identification of the nonlinear parameter are achievable. The results were presented for a single parameter case. The conditions for parameter convergence apply to corresponding systems with more than one parameter, as well. However, the choice of external inputs that will satisfy these conditions for n-dimensional parameterization is still under investigation.

The second class of nonlinearly parameterized systems that is addressed in this thesis is characterized by the presence of a nonlinearity which is monotonic in the parameters. The min-max algorithm [1] that was used for the control and estimation of convexly/concavely parameterized systems differs from other methods in adaptive control in that it does not solely rely on the local gradient information for generating parameter estimates. It was demonstrated in [1] that, for nonmonotonic parameterizations, the use of local gradients can lead to instability. However, for monotonically parameterized systems, the local gradient update law seems applicable. Two types of results were derived for such systems. For the first, it was shown that for certain types of systems, gradient laws can lead to instability due to the type of the system in which the nonlinearity is present. The second result, on the other hand, showed that local gradients are sufficient for the task of estimating the parameters in monotonically parameterized neural networks during the training procedure. This was shown using a low order neural network model by introducing a methodology whereby only the asymptotic behavior of the adaptive system is examined. Currently, efforts are being taken to extend this result to the full n-dimensional case. Preliminary studies indicate that the presented methodology does allow for this extension. All the results derived are complemented with numerical simulations which graphically illustrate the issues discussed.

References

[1] A. M. Annaswamy, A.P. Loh, and F.P. Skantze. Adaptive control of continuous time systems with convex/concave parametrization. *Automatica*, 34:33–49, January 1998.

[2] A.M. Annaswamy, N. Mehta, C. Thanomsat, and A.P. Loh. Application of adaptive controllers based on nonlinear parametrization to chemical reactors and magnetic bearings". In *Proceedings of the ASME Dynamic Systems and Control Division*, Dallas, TX, November 1997.

[3] B. Armstrong-Hélouvry. *Control of Machines with Friction.* Kluwer Academic Publishers, Norwell, MA, 1991.

[4] R. Bakker and A.M. Annaswamy. Stability and robustness properties of a simple adaptive controller. *IEEE Transactions on Automatic Control*, 41:1352–1358, September 1996.

[5] Jovan D. Bosković. Stable adaptive control of a class of fed-batch fermentation processes. In *Proceedings of the Eighth Yale Workshop on Adaptive and Learning Systems*, pages 263–268, Center for Systems Science, Dept. of Electrical Engineering, Yale University, New Haven, CT 06520-1968, June 1994.

[6] Esref Eskinat, Stanley H. Johnson, and William L Luyben. Use of hammerstein models in identification of nonlinear systems. *American Institute of Chemical Engineers journal*, 37(2):255–268, February 1991.

[7] M. Fleifil, A.M. Annaswamy, and A.F. Ghoniem. A nonlinear model of unsteady

heat release in laminar premixed flames. Technical report, Adaptive Control Laboratory, MIT, Cambridge, MA, 1997.

[8] V. Fomin, A. Fradkov, and V. Yakubovich. *Adaptive control of dynamical systems.* Eds. Nauka, Moscow, 1981.

[9] G. C. Goodwin and K. S. Sin. *Adaptive Filtering, Prediction, and Control.* Prentice-Hall, 1984.

[10] John Hertz, Anders Krogh, and Richard G. Palmer. *Introduction to the Theory of Neural Computation.* Addison-Wesley Publishing Company, 1991.

[11] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2:359–366, 1989.

[12] A. Isidori. *Nonlinear Control Systems.* Springer-Verlag, New York, NY, 2nd edition, 1989.

[13] H. K. Khalil. *Nonlinear Systems.* Prentice Hall, Upper Saddle River, NJ, 1996.

[14] M. Krstić, I. Kanellakopoulos, and P. Kokotović. *Nonlinear and Adaptive Control Design.* Wiley Interscience, New York, NY, 1995.

[15] Lennart Ljung. *System Identification: Theory for the User.* Prentice-Hall, Englewood Cliffs, NJ, 1987.

[16] A. P. Loh, A.M. Annaswamy, and F.P. Skantze. Adaptation in the presence of a general nonlinear parametrization: An error model approach. *IEEE Transactions on Automatic Control*, (accepted for publication) 1998.

[17] A.P. Morgan and K.S. Narendra. "On the stability of nonautonomous differential equations $\dot{x}= [A + B(t)]x$ with skew-symmetric matrix $B(t)$". *SIAM Journal of Control and Optimization*, vol. 15:163–176, January 1977.

[18] A.P. Morgan and K.S. Narendra. "On the uniform asymptotic stability of certain linear nonautonomous differential equations". *SIAM Journal on Control and Optimization*, vol. 15:5–24, January 1977.

[19] K. S. Narendra and A. M. Annaswamy. *Stable Adaptive Systems.* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1989.

[20] Kumpati S. Narendra and Kannan Parthasarathy. Identification and control of dynamical systems using neural networks. *IEEE Transactions on Neural Networks*, 1(1):4–26, March 1990.

[21] J. Park and I. W. Sandberg. Universal approximation using radial-basis function networks. *Neural Computation*, 3:246–257, 1991.

[22] C. Rohrs, L. Valavani, M. Athans, and G. Stein. Robustness of continuous-time adaptive control algorithms in the presence of unmodeled dynamics. *IEEE Transactions on Automatic Control*, 30:881–889, 1985.

[23] Walter Rudin. *Principles of Mathematical Analysis.* McGraw-Hill, Inc., 1976.

[24] J.W. Rumsey, M. Fleifil, A.M. Annaswamy, and A.F. Ghoniem. Low-order nonlinear models of thermoacoustic instabilities and linear model-based control. *IEEE Control System Technology*, submitted 1998.

[25] F. P. Skantze. Model-based algorithms for nonlinear system identification. Master's thesis, M.I.T., Cambridge, MA., 1996.

[26] F.P. Skantze, A.P. Loh, and A.M. Annaswamy. Adaptive estimation of discrete-time systems with nonlinear parametrization. *Automatica*, (submitted) 1997.

[27] J.-J. E. Slotine and W. Li. *Applied Nonlinear Control.* Prentice Hall, Englewood Cliffs, NJ, 1991.

[28] P. J. Werbos. Generalization of backpropagation with application to recurrent gas market model. *Neural Networks*, 1:339–356, 1988.

[29] S. Yu and Anuradha M. Annaswamy. Adaptive control of nonlinear dynamic systems using θ-adaptive neural networks. *Automatica*, 42(12), 1997.