PERSPECTIVES IN ADAPTIVE SYSTEMS: FREQUENCY DOMAIN ANALYSIS AND USE OF PRIOR INFORMATION

by

Er-Wei Bai

Memorandum No. UCB/ERL M87/56

4 August 1987

COVER PAGE

PERSPECTIVES IN ADAPTIVE SYSTEMS: FREQUENCY DOMAIN ANALYSIS AND USE OF PRIOR INFORMATION

by

Er-Wei Bai

Memorandum No. UCB/ERL M87/56

4 August 1987

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

TITLE PAGE

PERSPECTIVES IN ADAPTIVE SYSTEMS: FREQUENCY DOMAIN ANALYSIS AND USE OF PRIOR INFORMATION

by

Er-Wei Bai

Memorandum No. UCB/ERL M87/56

4 August 1987

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Perspectives in Adaptive Systems: Frequency Domain Analysis and Use of Prior Information

Er-Wei Bai

Signature: ____________________
Prof. S.Shankar Sastry
Committee Chairman

Abstract

This thesis addresses frequency domain techniques and the use of prior information in the analysis of adaptive identification and control schemes.

The contribution to frequency domain analysis is twofold: First, we establish a persistency of excitation condition on the regressor vector for a reduced order identifier, i.e. it is assumed that the order of the plant is not available and only the order of the nominal model is known. The theorem states, roughly speaking, that the persistency of excitation of the regressor vector depends on the order of the nominal model, i.e. the order of the identifier and is almost independent of the existence of unmodeled dynamics. It provides a foundation for the further analysis of a reduced order identifier using averaging techniques. Then we show, under some technical conditions, that the parameter estimate will converge either to a unique tuned model or to a neighbourhood of the tuned parameter. Second, we apply frequency domain analysis techniques to the global stability proof for an indirect adaptive scheme. We present a very general indirect adaptive control scheme along with its convergence proof. We show that if the exogenous input is rich enough, then the identifier and the controller converge to their "true" values. In the thesis, only two applications have been discussed. However the scheme presented is applicable to several kinds of controller design methodologies i.e. offers a great deal of flexibility in controller design and allows for a very general richness condition on the exogenous input.

We show how prior information may be used in the analysis and application of adaptive systems by constructing a model for a wide class of partially known systems and by presenting algorithms for the adaptive identification and control of such systems. If the system is completely unknown, the methods are identical to the standard ones in the literature. However, use of the particular prior information embedded in the model results in the identification and control of a fewer number of unknown parameters and consequently better performance.

Acknowledgements

I am grateful to my advisor Professor S. Shankar Sastry for his constant encouragement, careful advice and excellent guidance during the time of this research. It was a pleasant and rewarding experience to work under his supervision.

I wish to thank Professors Charles A. Desoer and Masayoshi Tomizuka for their thoughtful comments and valuable suggestions which greatly influenced this thesis. I am also indebted to Professors Arthur R. Bergen and Shoshichi Kobayashi for serving on my qualifying examination committee.

I would like to thank my fellow graduate students who made my stay at Berkeley a pleasure. I am particularly indebted to Saman Behtash, Arlene Cole, Lichen Fu, John Hauser, Greg Heinzinger, Ping Hsu, Zexiang Li, Jeff Mason and Andy Packard.

I dedicate this thesis to my family and my wife for their love and support over the years.

This research was supported by IBM Corporation under a Faculty Development Award 83-85 and by U.S. Army Research Office under Grant #DAAG 29-85-K-0072.

Table of Contents

Chapter 1 Introduction

This thesis focuses on the two subfields, frequency domain analysis and use of prior information in adaptive systems. The philosophy of the presentation is that we only present our own research in these two fields. In order to make it accessible to a wider audience, references are given for all existing results and also a simple proof, if necessary.

Adaptive systems have been extensively studied for over a decade and numerous successful algorithms and their applications have been reported. The development of the theory has led to a much better understanding of various adaptive identification and control schemes, however, the field still lacks analytical methods -particularly in the two subfields: frequency domain analysis and use of prior information.

Frequency domain analysis is a classical method for linear time invariant systems. In this thesis, our interest is to use this machinery to analyze adaptive systems. We successfully apply this method to the study of a reduced order identifier and the global stability proof of an indirect adaptive control scheme.

On the other hand, although the importance of the use of prior information in adaptive systems has long been recognized, formal and detail study in this area is very new. Our contribution in this area is that we propose a model for a wide class of 'partially known' systems and present adaptive identification and control algorithms, including a complete convergence analysis.

We begin with a brief review of the literature.

1-1 Review of Previous Work

Frequency domain analysis techniques in parameter identification may be traced back to the work of Rabkin et al [46] (1955), Levy [37] (1959) and Kardashov [31] (1958). They used experimental frequency data to determine the coefficients of a transfer function. The work was further developed by Sanathanan et al [47] (1963), Payne [44] (1970) and Stahl [50] (1984). The idea was that the coefficients of a transfer function were obtained as a result of minimizing some given cost function. At the same time, an on-line adaptive identifier was derived and developed by Lion [38] (1967), Luder et al [39] (1973) and Kreisselmeier [34] (1977). They devised an adaptive scheme to estimate the coefficients of a transfer function. Their work, together with that of Sondhi et al [49] (1976), Anderson [1] (1977) and Morgan et al [42] (1977) showed that the schemes proposed are globally asymptotically stable and furthermore convergence rate is exponential, provided the regressor vector is persistently exciting.

Boyd and Sastry [12] (1984) went further. They used frequency analysis techniques to discuss adaptive control and then results extend easily to the case of identifier. They changed the condition of persistency of excitation on the regressor vector to a condition on the frequency content of the input, i.e. if input contains as many spectral lines as there are unknown coefficients, then the regressor vector is persistently exciting and consequently the parameter error converges to zero exponentially. Their work led to further research in this direction by Fu et al [21] (1987), Bai et al [10](1987) and Mason et al [41] (1987), in which averaging methods were applied to the analysis of identifiers.

The use of prior information in adaptive systems was suggested by Dasgupta [14] (1984) and Clary [13] (1984). They noticed that in much of adaptive literature, the standard approach was to pre-suppose a complete lack of knowledge about the unknown systems and to ignore all additional information available to the modeller. The algorithms thus estimate all the coefficients of a transfer function. However, a great deal of partial knowledge is often available. It seems intuitively plausible that the identification and control algorithms could have faster convergence rate and be more robust, if this prior information could be incorporated into the adaptive systems. It is of course clear that one could neglect the prior information embodied in the system and still be able to identify and/or control the system. However, use of the particular prior information results in the identification of fewer number of unknown parameters and consequently faster convergence rate and better transient performance.

1-2 Contribution of The Thesis

The major results of this thesis are as follows:

(1) In chapter 2-2, we establish the persistency of excitation condition on the regressor vector for a reduced order identifier. It is well known that an adaptive identifier allows asymptotic estimation of the constant parameters of a linear time invariant system, provided that the order of the system is known and the regressor vector is persistently exciting. However, when a priori knowledge about the order of the system is not available and this occurs most often in practice, how does such an identifier behave? In this thesis, we prove that the persistency of excitation condition on the regressor vector depends on the order of the nominal model and is almost independent of the existence of unmodeled part. We then show that, under some technical conditions, if input is of the form $u(t) = \sum_{i=1}^{m} \xi_i \, sinw_i t$, the identified model will either converge globally and exponentially to an unique tuned model when $m = n$ (the order of nominal model) or will converge to a neighbourhood of an unique tuned parameter when $m > n$.

(2) In chapter 2-3, we present a general stability proof for continuous time adaptive control schemes, with very general assumptions on the identifier and controller. We show that if the exogenous input signal is rich enough, then both the identifier and the controller converge to their 'true' values. To our knowledge, this is the first proof of the persistency of excitation of the regressor vector signal in the closed loop without the use of an artificial random sampling signal for continuous time case. We show persistency of excitation without preassuming the boundness of the signal. Boundness of all signals and the convergence of the compensator in turn follow from the convergence of the identifier and this is a direct consequence of persistency of excitation of the signal in the identification loop. In this thesis, only two applications of the scheme have been discussed, but the scheme presented offers a great deal of flexibility in controller design and allows for very general richness conditions

on the reference input.

(3) In chapter 3-1, we present a model for a wide class of partially known systems, then in chapters 3-3 to 3-5 we give identification and control algorithms for such systems utilizing available prior information. If the system is completely unknown, these algorithms are identical to the standard approach in the literature. However, the schemes given will prove to be particularly important when we devise algorithms for the adaptive identification and control of these 'partially known' systems, since these algorithms have better transient performance, faster convergence rate and are consequently more robust, when the system is partially known.

Chapter 2. Frequency Domain Analysis

2-1 Notation and Preliminary

This section introduces some basic definitions and results used throughout this thesis. The notation is standard, e.g. $W(t)$ denotes a function of time, $\hat{W}(s)$ is its Laplace transform. Transfer functions (matrices) of linear time-invariant systems will be denoted by upper case letters, e.g. $\hat{G}(s)$ and $\hat{G}^*(s)$ is the complex conjugate of the transpose of $\hat{G}(s)$.

Definition 2-1.1 (Persistency of Excitation (PE))

A signal $W(t)\colon R_+ \to R^n$ is said to be *persistently exciting* if and only if there exist some $\alpha, \delta > 0$ such that

$$\int_{t_0}^{t_0+\delta} W(t)W^T(t)\,dt \geq \alpha I \qquad \text{for } all\ t_0 \geq 0$$

Roughly speaking, the intuition of the PE condition is that $W(t)$ spans the whole space R^n uniformly over time interval δ. There is an interesting frequency domain interpretation for PE condition. First, let us recall the definition of a spectral line.

Definition 2-1.2 (Spectral Line)

A signal $W(t)\colon R_+ \to R^n$ is said to have a *spectral line* at frequency w of amplitude $\overline{W}(w) \in C^n$ if and only if

$$\frac{1}{T}\int_{t_0}^{t_0+T} W(t)\, e^{-jwt}\, dt$$

converges to $\overline{W}(w)$ as $T \to \infty$ uniformly in t_0. When $\overline{W}(w) \neq 0$, we say $W(t)$ has a spectral line at w.

Then it follows,

Lemma 2-1.3 (PE and Spectral Lines)

Let $W(t) \in R^n$ have spectral lines at frequencies $w_1, \ldots, w_n$. Further assume that { $\overline{W}(w_1),...,\overline{W}(w_n)$ } are linearly independent in C^n. Then $W(t)$ is PE.

Proof: See [11].

Lemma 2-1.4 (Filter Lemma of Spectral Lines)

Let $u(t): R_+ \rightarrow R$, $y(t): R_+ \rightarrow R^n$ be the input and output, respectively, of a stable linear time-invariant system with transfer function (matrix) $\hat{G}(s)$. If u has a spectral line at frequency w, then so does y with amplitude

$$\bar{y}(w) = \hat{G}(jw)\,\bar{u}(w)$$

Proof: See [11].

Definition 2-1.5 (Sufficient Richness (SR))

A scalar signal $r(t): R_+ \rightarrow R$ is said to be *sufficiently rich* (of order n) if and only if it has at least n spectral lines.

It is shown [12], in a stochastic context that the PE condition is directly related to the positivity of the covariance of the signal. For the deterministic case, they are also very closely related.

Definition 2-1.6 (Autocovariance)

A signal $W(t): R_+ \rightarrow R^n$ is said to have an *autocovariance* $R_W(\tau) \in R^{n \times n}$ if and only if

$$\lim_{T \to \infty} \frac{1}{T} \int_{t_0}^{t_0+T} W(t) W^T(t+\tau)\, dt = R_W(\tau)$$

with the limit uniformly in t_0.

The concept is reminiscent of the theory of wide sense stationary stochastic processes. But we emphasize that an autocovariance is a completely deterministic notation. Its relation to the PE is simple.

Lemma 2-1.7 (PE Lemma)

Suppose $W(t)$ has an autocovariance $R_W(\tau)$. Then $W(t)$ is PE if and only if $R_W(0) > 0$.

Proof: See [12].

We will present a few more definitions and lemmas which will be used in the main body of the thesis. The proofs can be found in [12].

Definition 2-1.8 (Stationary)

A signal $W(t): R_+ \rightarrow R^n$ is said to be *stationary* if and only if it has an autocovariance $R_W(\tau)$.

Lemma 2-1.9 (Bochner Representation)

Suppose $R_W(\tau)$ is a positive semidefinite function. Then $R_W(\tau)$ has a Bochner representation

$$R_W(\tau) = \int e^{jw\tau} S_W(dw)$$

where S_W is a positive semidefinite matrix of bounded measure, called the spectral measure of $W(t)$.

Lemma 2-1.10 (Filter Lemma of Spectral Measure)

Suppose $u(t): R_+ \rightarrow R^n$ has an autocovariance $R_u(\tau)$ with its spectral measure S_u, and $\hat{G}(s)$ is a stable transfer matrix. Then $\hat{y}(s) = \hat{G}(s)\,\hat{u}(s)$ has an autocovariance $R_y(\tau)$. Its spectral measure is given by

$$S_y(dw) = \hat{G}(jw) S_u(dw)\, \hat{G}^*(jw)$$

The following lemma is concerned with the measure of a proper set, which is a very useful concept in section 2.2.

Lemma 2-1.11

Let $f: R^n \rightarrow R^k$ be real analytic and not identically zero. Then the set $V = f^{-1}(0)$ has measure zero in R^n.

Proof: See [27].

2-2 Convergence Analysis of a Reduced Order Identifier

2-2-1 Problem Statement

An adaptive identifier (e.g Kreisselmeier [34] and Lunders & Narendra [39]) allows asymptotic estimation of the constant parameters of a linear time invariant plant, provided that the order of the plant is known and the regressor satisfies some persistency of excitation conditions. However, for the case where a priori knowledge about the order of the plant is not avaiable, which occurs for most practical situations, how do such identifiers behave? We will study the problem in this section, i.e. we consider the problem of identifying a reduced order single-input single-output transfer function. The transfer function to be identified is of the form

$$\hat{G}(s) = \hat{G}_0(s)\frac{n_u(s)}{d_u(s)}$$

with $\hat{G}_0(s)$ nominal model which is to be identified and $\frac{n_u(s)}{d_u(s)}$ unmodeled part. System identifications with no unmodeled dynamics, i.e. $\frac{n_u(s)}{d_u(s)} \equiv 1$, have been investigated by many authors (see e.g. Kreisselmeier [34] and Luders & Narendra [39]). We consider the case where $\frac{n_u(s)}{d_u(s)} \neq 1$. The main technique we used here is the frequency domain analysis, which we think is intuitively insightful and technically rigorous. First we prove that the persistently exciting condition of the regressor vector depends on the order of the nominal part $\hat{G}_0(s)$ (i.e. the order of identifier) and is almost regardless of the existence of unmodeled part. This result is very useful from the engineering point of view, since the plant to be identified is always lower order model and usually only little knowledge is available about the unmodeled dynamics. We then show that, under some technical conditions, if input is of the form $u(t) = \sum_{i=1}^{m} \xi_i \sin w_i t$, the identified model will either converge globally and exponentially to an unique tuned model $\hat{G}_T(s)$, (which depends on the choice of the input frequencies w_i), when $m=n$ (where n is the order of identifier), such that

$$\hat{G}(jw_i) = \hat{G}_T(jw_i) \qquad i=1,\ldots,n$$

or converge to a neibourhood of an unique tuned parameter when $m > n$.

2-2-2 The Identifier Structure and The PE Condition on The Regressor

The plant under consideration is of the form

$$\hat{G}(s)=\hat{G}_0(s)\frac{n_u(s)}{d_u(s)}=\frac{n_0(s)}{d_0(s)}\cdot\frac{n_u(s)}{d_u(s)}=\frac{n(s)}{d(s)} \tag{2-2-2.1}$$

with the following assumptions:

(A1) $\hat{G}(s)$ is a unknown strictly proper, stable, finite order and coprime transfer function. The order N is unknown and could be very large.

(A2) The nominal part $\hat{G}_0(s)=\dfrac{n_0(s)}{d_0(s)}$ is a unknown *nth* order, strictly proper transfer function.

The identification problem is to identify nominal part $\hat{G}_0(s)$ (i.e. the coefficients of the numerator and denominator of $\hat{G}_0(s)$, or Certainty Equivalence) from the input-output measurements of the plant.

Remarks:

(1) The unmodeled part may represent some high frequency dynamics, nearly pole-zero cancellations, and other kinds of unmodeled dynamics.

(2) Let $\Delta\hat{G}_1(s)=\dfrac{n_u(s)}{d_u(s)}-1$ and $\Delta\hat{G}_2(s)=\hat{G}_0(s)\,\Delta\hat{G}_1(s)$, then the plant $\hat{G}(s)$ may be written as

$$\hat{G}(s)=\hat{G}_0(s)(1+\Delta\hat{G}_1(s))$$

or

$$\hat{G}(s)=\hat{G}_0(s)+\Delta\hat{G}_2(s)$$

which are standard representations of unmodeled dynamics (multiplicative and additive). Notice that no assumption has been made about the properness of unmodeled part $\dfrac{n_u(s)}{d_u(s)}$, hence it is possible that

$$|\Delta\hat{G}_1(jw)|\rightarrow\infty \quad as\ |w|\rightarrow\infty$$

The identifier considered here is an adaptive observer/identifier [34,39] as in figure 2-2-2.1.

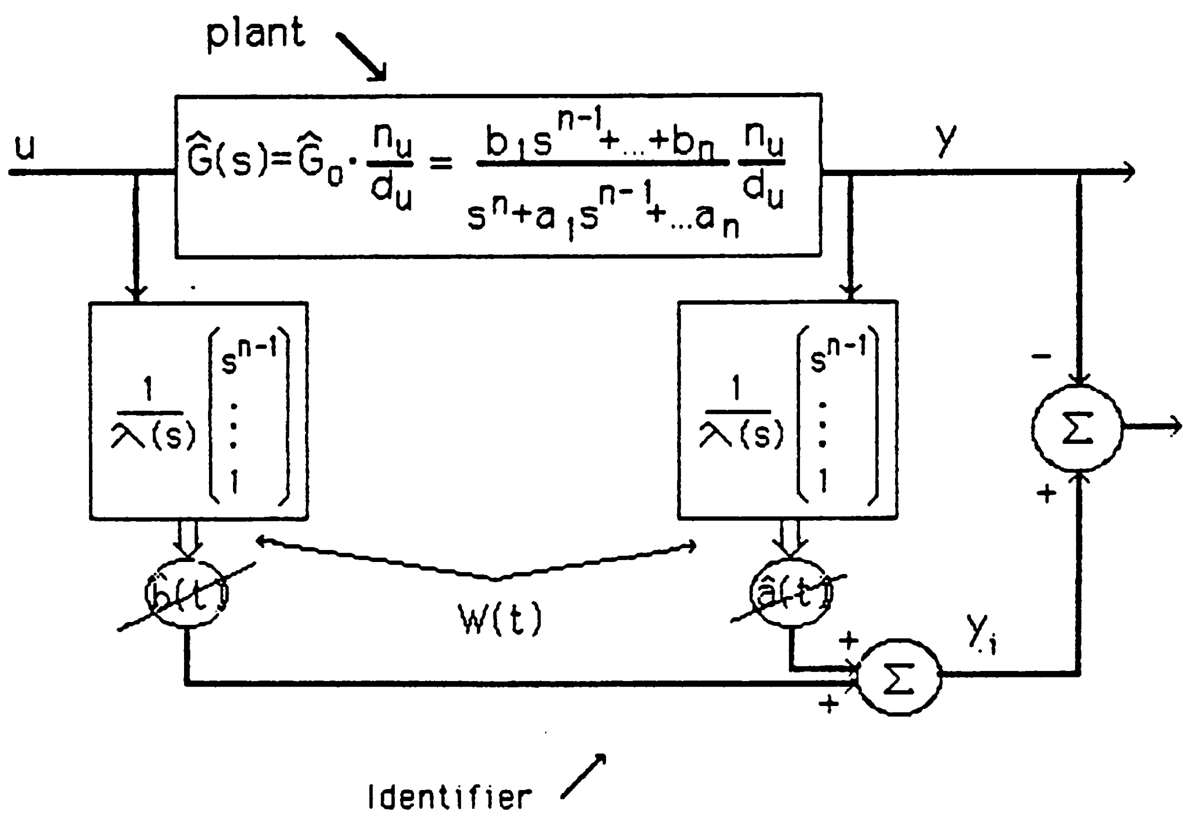

Fig. 2-2-2.1

Define

$$\hat{W}^T(s)=(\frac{s^{n-1}}{\lambda(s)}\hat{y}(s),...,\frac{1}{\lambda(s)}\hat{y}(s),\frac{s^{n-1}}{\lambda(s)}\hat{u}(s),...,\frac{1}{\lambda(s)}\hat{u}(s))$$

where $\lambda(s)=s^n+\lambda_1 s^{n-1}+\lambda_2 s^{n-2}+...+\lambda_n$ is an arbitrary *nth* order stable monic polynomial and let

$$\hat{\theta}^T(t)=(\hat{a}^T(t),\hat{b}^T(t))=(\hat{a}_1(t),...,\hat{a}_n(t),\hat{b}_1(t),...,\hat{b}_n(t))$$

denote the parameter estimate of the nominal model $\hat{G}_0(s)$ and $W(t)$ denote the regressor vector (Laplace inverse of $\hat{W}(s)$), we see that the output of the identifier can be written as

$$y_i(t)=W^T(t)\,\hat{\theta}(t)$$

The parameter update law is defined by

$$\dot{\hat{\theta}}(t)=-W(t)(y_i(t)-y(t))$$

$$=-W(t)W^T(t)\hat{\theta}(t)+W(t)y(t) \qquad (2\text{-}2\text{-}2.2)$$

It is well known (see e.g. Boyd & Sastry [11]) that in the case of no unmodeled dynamics, i.e. $\frac{n_u(s)}{d_u(s)}\equiv 1$, the necessary and sufficient condition for parameter estimate to converge to the true value exponentially is the persistently exciting condition on the regressor $W(t)$, i.e.

$$\int_{t_0}^{t_0+\delta} WW^T dt \ge \alpha I$$

for some $\alpha, \delta > 0$ and all $t_0 > 0$ (see definition 2-1.1) or input $u(t)$ is sufficiently rich (of order $2n$), if $u(t)$ is stationary (see section 2-1)

In the presence of unmodeled part, the situation is more complicated since $W(t)$ involves unmodeled part $\frac{n_u(s)}{d_u(s)}$. It is easy to see that necessary condition for $W(t)$ to be persistently exciting is that input u(t) is sufficiently rich (of order $2n$) since $W(t)$ is $2n$ dimensional. However, it is not clear for the sufficient condition. To this end, we have the following lemma and theorem.

Lemma 2-2-2.1

Consider the plant $\hat{G}(s)$ in (2-2-2.1) and the identifier (2-2-2.2). Then for each l, $1 \le l \le n$, there is a nowhere dense, measure zero subset V_{2l} in R^{2l} such that if the input u(t) has $2l$ spectral lines at w_i, $i=1,...2l$ and $(w_1,...,w_{2l}) \in R^{2l} - V_{2l}$, the resultant regressor $W(t)$ has $2l$ linearly independent spectral lines at w_i, $i=1,...,2l$.

Proof: We prove this lemma in 2 steps. First we show that there exists some point $(w_1,...,w_{2l}) \in R^{2l}$ such that if the input has $2l$ spectral lines at these frequencies, then so does $W(t)$ and $\{ \overline{W}(w_1),...,\overline{W}(w_{2l})\}$ are linearly independent.

Proceed by similar arguments as in Boyd & Sastry [11], we have that if the input u(t) has $2N$ (N is the order of plant $\hat{G}(s)$) spectral lines at $(w_1,...,w_{2N})$, then the matrix defined by

$$M_{2N}(w_1,\ldots,w_{2N}) = \begin{bmatrix} (jw_1)^{N-1}n(jw_1) & \cdots & (jw_{2N})^{N-1}n(jw_{2N}) \\ \vdots & & \vdots \\ n(jw_1) & \vdots & n(jw_{2N}) \\ (jw_1)^{N-1}d(jw_1) & \cdots & (jw_{2N})^{N-1}d(jw_{2N}) \\ \vdots & & \vdots \\ d(jw_1) & \cdots & d(jw_{2N}) \end{bmatrix}$$

is nonsingular. This implies that there is a nonsingular submatrix

$$M_{2l}(w_{i_1},\ldots,w_{i_{2l}}) = \begin{bmatrix} (jw_{i_1})^{l-1}n(jw_{i_1}) & \cdots & (jw_{i_{2l}})^{l-1}n(jw_{i_{2l}}) \\ \vdots & & \vdots \\ n(jw_{i_1}) & \vdots & n(jw_{i_{2l}}) \\ (jw_{i_1})^{l-1}d(jw_{i_1}) & \cdots & (jw_{i_{2l}})^{l-1}d(jw_{i_{2l}}) \\ \vdots & & \vdots \\ d(jw_{i_1}) & \cdots & d(jw_{i_{2l}}) \end{bmatrix}$$

where $(w_{i_1},\ldots,w_{i_{2l}}) \in (w_1,\ldots,w_{2N})$. Using this fact and lemma 2-1.4, it follows that if input $u(t)$ has $2l$ spectral lines at $(w_{i_1},\ldots,w_{i_{2l}})$, then $W(t)$ has $2l$ linearly independent spectral lines at these frequencies, i.e.

$$(\overline{W}(w_{i_1}),\ldots,\overline{W}(w_{i_{2l}})) = \begin{bmatrix} (jw_{i_1})^{n-1}n(jw_{i_1}) & \cdots & (jw_{i_{2l}})^{n-1}n(jw_{i_{2l}}) \\ \vdots & & \vdots \\ n(jw_{i_1}) & \vdots & n(jw_{i_{2l}}) \\ (jw_{i_1})^{n-1}d(jw_{i_1}) & \cdots & (jw_{i_{2l}})^{n-1}d(jw_{i_{2l}}) \\ \vdots & & \vdots \\ d(jw_{i_1}) & \cdots & d(jw_{i_{2l}}) \end{bmatrix}$$

$$\cdot\, diag\left(\frac{\overline{u}(w_{i_1})}{\lambda(jw_{i_1})\,d(jw_{i_1})},\ldots,\frac{\overline{u}(w_{i_{2l}})}{\lambda(jw_{i_{2l}})\,d(jw_{i_{2l}})}\right)$$

This completes the first step. Now define a map $f_{2l} : R^{2l} \to C$ $(or\ R^2)$ by

$$f_{2l}(w_1, \ldots, w_{2l}) = \det \begin{bmatrix} (jw_1)^{l-1} n(jw_1) & \cdots & (jw_{2l})^{l-1} n(jw_{2l}) \\ \vdots & \vdots & \vdots \\ n(jw_1) & \vdots & n(jw_{2l}) \\ (jw_1)^{l-1} d(jw_1) & \cdots & (jw_{2l})^{l-1} d(jw_{2l}) \\ \vdots & \vdots & \vdots \\ d(jw_1) & \cdots & d(jw_{2l}) \end{bmatrix}$$

$$= \det M_{2l}(w_1, \ldots, w_{2l}) \qquad (2\text{-}2\text{-}2.3)$$

f_{2l} is analytic and not identically zero by the first step of proof. This implies that the set

$$V_{2l} = f_{2l}^{-1}(0)$$

is measure zero from lemma 2-1.11 and then the conclusion follows.

Theorem 2-2-2.2 (Persistency of Excitation of W(t))

Consider the plant $\hat{G}(s)$ in (2-2-2.1) and the identifier (2-2-2.2). Then for almost any $2n$ spectral lines which input u(t) has, the resultant regressor $W(t)$ is persistently exciting. More precisely, there exist a nowhere dense, measure zero subset V in R^{2n}, such that if input u(t) has $2n$ spectral lines at w_i, i=1,...,$2n$ and $(w_1, \ldots, w_{2n}) \in R^{2n} - V$, the resultant regressor $W(t)$ is persistently exciting.

Proof: Follows from lemmas 2-1.3 and 2-2-2.1.

Remarks:

(1) The theorem 2-2-2.2 is very useful since the persistency of excitation condition of $W(t)$ is almost satisfied if input u(t) has $2n$ spectral lines which does not depend on the order and the "bigness" of the unmodeled part.

(2) We conjecture that the set V is a union of finite points and finite smooth manifolds $\{(w_1, \ldots, w_{2n}) \in R^{2n} \mid w_i = w_j \text{ for } some\ i \neq j\}$

We now in a position to analyze the convergence of the parameter estimate. Three cases will be discussed seperately where the input has exactly 2n, less than 2n and more than 2n spectral lines.

2-2-3 Tuned Model (Input u(t) Has Exactly 2n Spectral Lines)

In this section, we discuss the case where input has exactly $2n$ spectral lines. First we give the following lemma.

Lemma 2-2-3.1

Suppose input u(t) has $2n$ spectral lines at $(w_1,...w_{2n}) \in R^{2n} - V$ (this is almost true since V is measure zero). Then there exist two unique polynomials

$$\gamma(s) = \gamma_1 s^{n-1} + ... + \gamma_n$$

and

$$\beta(s) = s^n + \beta_1 s^{n-1} + ... + \beta_n$$

such that

$$n(jw_i)\,\beta(jw_i) = d(jw_i)\,\gamma(jw_i) \quad i=1,...,2n$$

If

$$\beta(jw_i) \neq 0, \qquad \text{for } i=1,...,2n \tag{2-2-3.1}$$

then

$$\hat{G}(jw_i) = \hat{G}_T(jw_i) \quad i=1,...,2n$$

where

$$\hat{G}_T(s) = \frac{\gamma(s)}{\beta(s)} = \frac{\gamma_1 s^{n-1} + ... + \gamma_n}{s^n + \beta_1 s^{n-1} + ... + \beta_n}$$

Moreover if input is of the form $u(t) = \sum_{i=1}^{n} \xi_i \sin w_i t$, all coefficients γ_i and β_i are real.

Proof: The hypothesis implies that the matrix M_{2n} defined in equation (2-2-2.3) is nonsingular at $(w_1, \ldots, w_{2n})$. Consequently, there is a unique solution $(\beta_1, \ldots, \beta_n, -\gamma_1, \ldots, -\gamma_n)$ for the following equation

$$(\beta_1, \ldots, \beta_n, -\gamma_1, \ldots, -\gamma_n)\, M_{2n}(w_1, \ldots, w_{2n}) = (-(jw_1)^n n(jw_1), ..., -(jw_{2n})^n n(jw_{2n}))$$

i.e.

$$n(s)\,\beta(s) = d(s)\,\gamma(s)\ ,\quad s = jw_1, \ldots, jw_{2n}$$

where

$$\gamma(s) = \gamma_1 s^{n-1} + \ldots + \gamma_n$$

$$\beta(s) = s^n + \beta_1 s^{n-1} + \ldots + \beta_n$$

If $u(t) = \sum_{i=1}^{n} \xi_i sinw_i t$, i.e. $u(t)$ has spectral lines at $\pm w_i$ $(i=1,...,n)$, then

$$\overline{(\beta_1, \ldots, \beta_n, -\gamma_1, \ldots, -\gamma_n)} = (\beta_1, \ldots, \beta_n, -\gamma_1, \ldots, -\gamma_n)$$

Here -- denotes the complex conjugate. This completes the proof.

Remarks:

(1) Nothing has been said about the stability of $\hat{G}_T(s)$. In fact, $\hat{G}_T(s)$ could be unstable, depending on the frequencies w_i chosen.

(2) If we assume the stability of $\hat{G}_T(s)$, then it is clear under the input $u(t)=\sum_{i=1}^{n} \xi_i \sin w_i t$ the output $y_T(t)$ of $\hat{G}_T(s)$ equals to the plant output $\underset{\sim}{y}(t)$ up to an exponentially decaying term due to initial conditions. The following lemma tells us that under some technical conditions this fact is also true even though $\hat{G}_T(s)$ is unstable.

Lemma 2-2-3.2

Consider a strictly proper, *nth* order transfer function $\hat{G}_T(s)$. Suppose $\hat{G}_T(s)$ has no pole on the jw axis. Then for input of the form $u(t)=sinwt$, there exists some initial condition $I_T(0)$ for $\hat{G}_T(s)$, such that the output $y_T(t)$ of $\hat{G}_T(s)$ is pure sinusoid with the same frequency w as the input. More precisely,

$$y_T(t) = \mathrm{Im}(\hat{G}_T(jw)\, e^{jwt}) \quad \text{for } all\ t \qquad (2\text{-}2\text{-}3.2)$$

where Im denotes the imaginary part of a complex number.

Before giving the proof, we need the following fact

Fact 2-2-3.3

If a matrix A has no eigenvalues on the jw axis, then $(A^2+w^2I)^{-1}$ exists. Moreover

$$(sI-A)^{-1}\frac{w}{s^2+w^2}=(sI-A)^{-1}w(w^2I+A^2)^{-1}-A(w^2I+A^2)^{-1}\frac{w}{s^2+w^2}-(w^2I+A^2)^{-1}\frac{ws}{s^2+w^2}$$

Proof: See [55].

Proof of lemma 2-2-3.2: Let (A,b,c) be a minimal realization of $\hat{G}_T(s)$, the Laplace transform of output $y_T(t)$ under input $u(t)=sinwt$ is

$$\hat{y}_T(s)=c(sI-A)^{-1}l_T(0)+c(sI-A)^{-1}b\frac{w}{s^2+w^2}$$

where $l_T(0)$ is the initial condition. From the fact 2-2-3.3, it follows that

$$\hat{y}_T(s)=c(sI-A)^{-1}l_T(0)+c(sI-A)^{-1}w(w^2I+A^2)^{-1}b$$

$$-cA(w^2I+A^2)^{-1}b\frac{w}{s^2+w^2}-c(w^2I+A^2)^{-1}b\frac{ws}{s^2+w^2}$$

Choose $l_T(0)=-w(w^2I+A^2)^{-1}b$, we obtain

$$y_T(t)=-cA(w^2+A^2)^{-1}b\ sinwt-c(w^2I+A^2)^{-1}b\ coswt$$

Thus the output is pure sinusoid, the equation (2-2-3.2) can be found in Desoer and Kuh [16,pg280].

Definition: 2-2-3.4

$(\hat{G}_T(s),l_T(0))$ is called a *tuned model* of $\hat{G}(s)$ at frequencies $(\pm w_1,\ldots,\pm w_n)$, if under the input $u(t)=\sum_{i=1}^{n}\xi_i \sin w_i t$, the output $y_T(t)$ of $\hat{G}_T(s)$ with initial condition $l_T(0)$ equals to the output $y(t)$ of $\hat{G}(s)$ up to an exponentially decaying term.

Similar to the definition of the regressor $\hat{W}(s)$ for the identifier in Fig. 2-2-1.1, we define

$$\hat{W}_T(s)=\left(\frac{s^{n-1}}{\lambda(s)}\hat{y}_T(s),\ldots,\frac{1}{\lambda(s)}\hat{y}_T(s),\frac{s^{n-1}}{\lambda(s)}\hat{u}(s),\ldots,\frac{1}{\lambda(s)}\hat{u}(s)\right)^T$$

It is easy to verify that under input $u(t)=\sum_{i=1}^{n}\xi_i \sin w_i t$, the output of the tuned model may be written as

$$y_T(t)=\theta_T{}^T W_T(t)+\Delta y_T(t)$$

for $\theta_T=(\lambda_1-\beta_1,\ldots,\lambda_n-\beta_n,\gamma_1,\ldots,\gamma_n)^T \in R^{2n}$ and some exponentially decaying term $\Delta y_T(t)$ due to the initial condition of the filter $\frac{1}{\lambda(s)}$. Thus we can relate $W(t)$ to $W_T(t)$ by

$$W(t)=W_T(t)+\Delta W(t)$$

where $\Delta W(t)$ goes to zero exponentially because of $y(t)=y_T(t)$ up to an exponentially decaying term.

Now under the assumption of existence of a tuned model, let us reconsider the parameter update law (2-2-2.2)

$$\begin{aligned}\dot{\hat\theta}&=-W(y_i-y)\\ &=-W(y_i-y_T+y_T-y)\\ &=-W(W^T\hat\theta-W_T{}^T\theta_T)-W(y_T-y)\\ &=-WW^T(\hat\theta-\theta_T)-W(\theta_T{}^T\Delta W+y_T-y)\\ &=-WW^T(\hat\theta-\theta_T)+\Delta e\end{aligned}$$

where Δe indicates an exponentially decaying term. By defining

$$\phi_T=\hat\theta-\theta_T \tag{2-2-3.3}$$

We have

$$\dot\phi_T=-WW^T\phi_T+\Delta e \tag{2-2-3.4}$$

This procedure can be seen clearly as in figure 2-2-3.1.

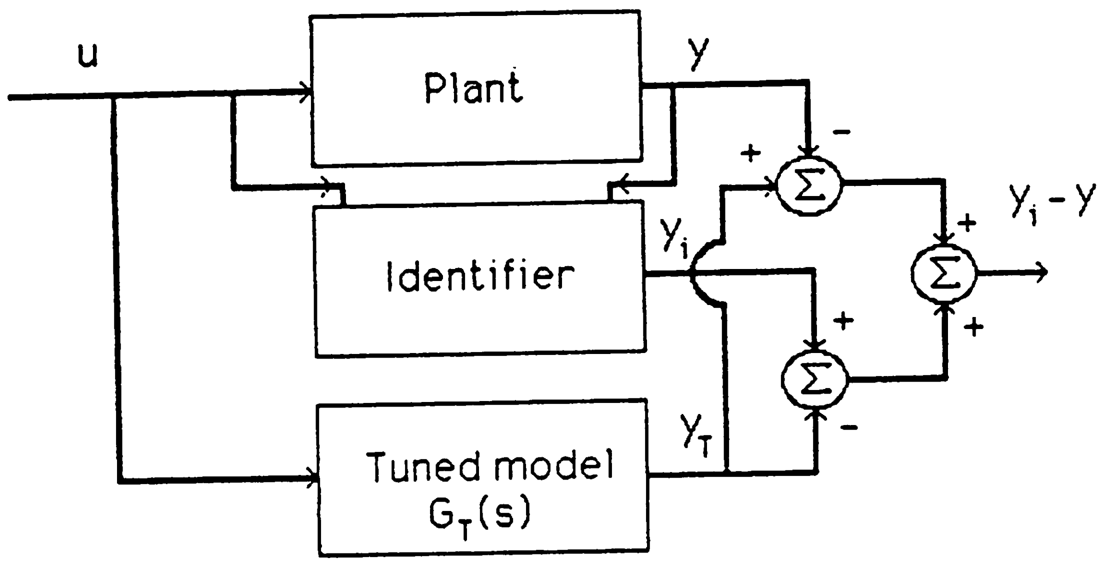

Fig.2-2-3.1

Noticed that in the Fig.2-2-3.1, the tuned model is pure fictitious. Consequently, no robustness and sensitivity problems exist even though the tuned model is unstable. But by doing so, we relate the parameter estimate to the tuned model. From equation (2-2-3.4), we see that the identified model (parameter estimate) converges globally and exponentially to the tuned model. In summary, we have

Theorem 2-2-3.5 (Convergence of The Identifier)

Consider the plant $\hat{G}(s)$ in (2-2-2.1) and the identifier (2-2-2.2) with input $u(t) = \sum_{i=1}^{n} \xi_i \sin w_i t$. Suppose $W(t)$ is persistently exciting and the tuned model exists. Then the identified nominal model (estimation of the nominal model $\hat{G}_0(s)$) converges globally and exponentially to an unique tuned model $\hat{G}_T(s)$, which depends on the input frequencies, such that

$$\hat{G}(s) = \hat{G}_T(s) \qquad s = \pm jw_1, \ldots, \pm jw_n$$

Proof: Follows from the equation (2-2-3.4).

Remark:

The theorem 2-2-3.5 guarantees that if input is of the form $u(t) = \sum_{i=1}^{n} \xi_i \sin w_i t$ satisfying some technical conditions, the identified nominal model will converge to the tuned model. However nothing has been said about how close between the nominal model $\hat{G}_0(s)$ which is to be identified and the tuned model $\hat{G}_T(s)$ which we actually get. In fact, they may be quite different. This can be seen from the following example.

The plant we consider is a first order model with high frequency unmodeled part as follows

$$\hat{G}(s) = \frac{k}{s+p} \cdot \frac{\delta p}{s+\delta p} = \hat{G}_0(s) \frac{\delta p}{s+\delta P} \qquad \delta, p > 0$$

where δ represents the ratio of modeled pole and unmodeled pole. When input $u(t) = \xi \sin wt$, for any stable filter $\frac{1}{\lambda(s)} = \frac{1}{s+\alpha}$, the regressor $W(t)$ is always persistently exciting for all $w \neq 0$. (Note that the theorem 2-2-2.2 only guarantees that for almost any frequency w, The regressor $W(t)$ is PE. However, in this example, any frequency $w \neq 0$ will produce a persistently exciting regressor $W(t)$). Since the matrix

$$\left[\frac{k\delta p}{(jw+p)(jw+\delta p)}, \ \frac{k\delta p}{(-jw+p)(-jw+\delta p)} \right]$$

is nonsingular for all $w \neq 0$ and this implies $W(t)$ always has two linearly independent spectral lines for all $w \neq 0$. From the theorem 2-2-3.5, the identified nominal model will converge to the tuned model $\frac{k_T}{s+p_T}$ such that

$$\frac{k}{jw+p} \cdot \frac{\delta p}{jw+\delta p} = \frac{k_T}{jw+p_T}$$

By solving this equation, we have

$$k_T(w) = k \cdot \frac{\delta}{\delta+1}$$

$$p_T(w) = \frac{\delta p^2 - w^2}{p(\delta+1)}$$

We see that if the input frequency $w > \delta^{1/2}p$, the tuned model is unstable (even $\hat{G}(s)$ is stable). In fact, identified pole p_T could be any number in the interval $(-\infty, p\frac{\delta}{\delta+1})$ depending on the input frequency w chosen. Fig. 2-2-3.2 and Fig. 2-2-3.3 show the simulation results where $k=p=1$, $\delta=10$ with input $u(t)=10sint$ and $u(t)=10sin5t$ respectively. In order to get a good estimation of $\hat{G}_0(s)$, the input frequency w must be in the low frequency range which will be discussed in later sections.

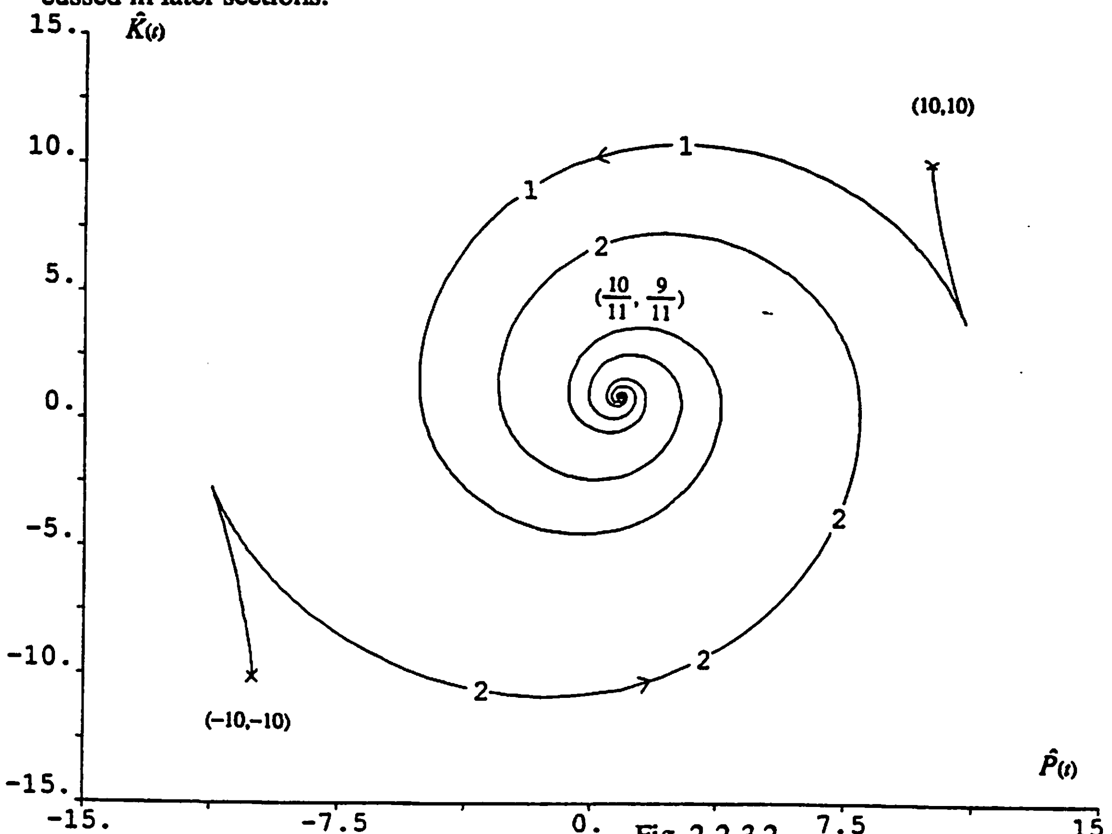

Fig. 2-2-3.2

$u(t) = 10sint$, $(\hat{k}(t), \hat{p}(t)) \rightarrow (k_T(1), p_T(1)) = (\frac{10}{11}, \frac{9}{11})$

-1- $(\hat{k}(0), \hat{p}(0)) = (10, 10)$, -2- $(\hat{k}(0), \hat{p}(0)) = (-10, -10)$

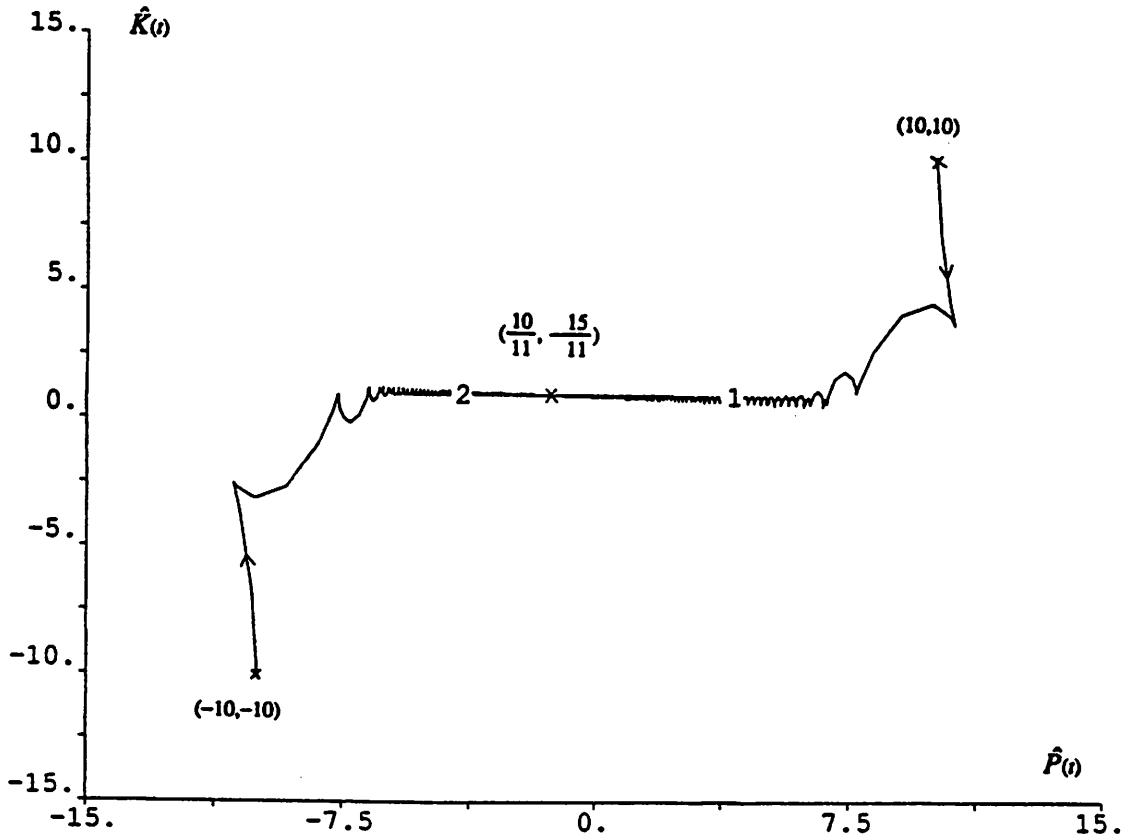

Fig.2-2-3.3

$u(t) = 10 sin\, 5t$, $(\hat{k}(t), \hat{p}(t)) \rightarrow (k_T(5), p_T(5)) = (\frac{10}{11}, \frac{-15}{11})$

-1- $(\hat{k}(0), \hat{p}(0)) = (10, 10)$, -2- $(\hat{k}(0), \hat{p}(0)) = (-10, -10)$

2-2-4 Lack of Persistency of Excitation (Input u(t) Has Less Than 2n Spectral Lines)

In this case, the regressor $W(t)$ is not persistently exciting. However, from lemma 2-2-2.1, we see that if input u(t) has $2l$ spectral lines at

$$(w_1, \ldots, w_{2l}) \in R^{2l} - V_{2l} \tag{2-2-4.1}$$

then $W(t)$ has $2l$ linearly independent spectral lines at these frequencies. By similar arguments as in lemma 2-2-3.1, there exist two polynomials (not unique)

$$\gamma(s) = \gamma_1 s^{n-1} + \ldots + \gamma_n$$

$$\beta(s) = s^n + \beta_1 s^{n-1} + \ldots + \beta_n$$

such that

$$n(jw_i)\,\beta(jw_i) = d(jw_i)\,\gamma(jw_i) \qquad i=1,\ldots,2l$$

If

$$\beta(jw_i) \neq 0 \quad i=1,\ldots,2l \tag{2-2-4.2}$$

then

$$\hat{G}(jw_i) = \frac{\gamma(jw_i)}{\beta(jw_i)} = \hat{G}_T(jw_i) \quad i=1,\ldots,2l$$

further if $w_{2i-1} = -w_{2i}$, $i=1,\ldots,l$, then all γ_i, β_i's are real numbers i.e. if input $u(t) = \sum_{i=1}^{l} \xi_i \, sinw_i t$ $(l<n)$ satisfying (2-2-4.1) and (2-2-4.2), there is a tuned model (not unique) $(\hat{G}_T(s), I_T(0))$ of $\hat{G}(s)$, such that the output $y_T(t)$ of $\hat{G}_T(s)$ with initial condition $I_T(0)$ equals to the output $y(t)$ of plant $\hat{G}(s)$, up to an exponentially decaying term. Based on the above discussion, we have

Theorem 2-2-4.1

Consider the plant (2-2-2.1) and the identifier (2-2-2.2)

$$\dot{\theta} = -WW^T\theta + Wy = -W(y_i - y)$$

with input of the form $u(t) = \sum_{i=1}^{l} \xi_i sinw_i t$, $(l<n)$. Suppose the conditions (2-2-4.1) and (2-2-4.2)

are satisfied, then

(1) $||\hat{\theta}(t)|| \leq M$ for *all* t and *some* $M < \infty$

(2) $\dot{\hat{\theta}}(t) \to 0, \quad as\ t \to \infty$

Proof: By using same notations as in the previous section, for some tuned model $(\hat{G}_T(s), I_T(0))$, we may rewrite the parameter update law as

$$\dot{\hat{\theta}} = -W(y_i - y_T + y_T - y)$$

$$= -W(W^T\hat{\theta} - W_T{}^T\theta_T) - W(y_T - y)$$

Since $y_T \to y$ exponentially, it follows that

$$\dot{\hat{\theta}} = -WW^T(\hat{\theta} - \theta_T) + W(\theta_T{}^T\, \Delta W + y_T - y)$$

$$= -WW^T(\hat{\theta} - \theta_T) + \Delta e$$

where Δe and $\Delta W = W(t) - W_T(t)$ are exponentially decaying terms. As is standard in the literature, we will drop such terms in our analysis, since the presence of these terms does not change any of the conclusions that follow. Let $\phi = \hat{\theta} - \theta_T$, we have

$$\dot{\phi} = -WW^T\phi \qquad (2\text{-}2\text{-}4.3)$$

Define a Lyapunov function

$$v = \frac{\phi^T\phi}{2}$$

the derivative of v along the solution of equation (2-2-4.3) becomes

$$\dot{v} = -(W^T\phi)^2$$

This implies that $||\phi(t)||$ is positive, nonincreasing function, so that $||\hat{\theta}(t)|| = ||\phi(t) + \theta_T|| \leq M$ for some $0 < M < \infty$. Also note that $(W^T\phi) \in L_2$ and $d/dt\ (W^T\phi)$ is bounded, hence $W^T\phi \to 0$ as $t \to \infty$. Then it follows that

$$\dot{\hat{\theta}}(t) = \dot{\phi}(t) = -W(t)W^T(t)\phi(t) \to 0, \qquad as\ t \to \infty$$

Remarks:

(1) The above theorem guarantees that if sinusoid input satisfies the technical conditions (2-2-4.1) and (2-2-4.2), the parameter estimate is bounded and asymptotically slows down, though the homogeneous part of parameter update equation (2-2-2.2) is not exponentially stable. Results are similar to that of no unmodeled dynamics case.

(2) It has been shown in the theorem that $\hat{\theta}$ is bounded and $\dot{\hat{\theta}} \to 0$ as $t \to \infty$, but $\hat{\theta}$ may not converge at all. This fact has been proven even in the case of no unmodeled dynamics [11,27]. The exception is DC input.

Proposition 2-2-4.2

Consider the plant $\hat{G}(s)$ in (2-2-2.1) and the identifier (2-2-2.2). Suppose the input $u(t) \equiv c$ for some constant c. Then parameter estimate $\hat{\theta}(t) \to \theta_0$ as $t \to \infty$ for some $\theta_0 \in R^{2n}$.

Proof: The hypothesis implies that

$$W(t) \to W_0 \quad \textit{(a constant vector) exponentially}$$

$$y(t) \to y_0 \quad \textit{(a constant) exponentially}$$

Consequently, the update law (2-2-2.2) becomes asymptotically

$$\dot{\hat{\theta}} = -W_0 W_0^T \hat{\theta} + W_0 y_0$$

from linear algebra, there is an orthonormal matrix J, such that

$$J^{-1} W_0 = \begin{bmatrix} \lambda^{1/2} \\ 0 \\ \cdot \\ \dot{0} \end{bmatrix}$$

where $\lambda \geq 0$ is the maximum eigenvalue of $W_0 W_0^T$. Define

$$\hat{\theta}(t) = J\, x(t)$$

we have

$$\dot{x}(t) = -\begin{bmatrix} \lambda & 0 & & 0 \\ 0 & 0 & & \cdot \\ \cdot & \cdot & & \cdot \\ \dot{0} & \dot{0} & & \dot{0} \end{bmatrix} x(t) + \begin{bmatrix} 1 \\ 0 \\ \cdot \\ \dot{0} \end{bmatrix} \lambda^{1/2} y_0$$

From this equation, we may conclude that $\hat{\theta} = J\, x$ converges to some point. This completes the proof.

In general, input u(t) having less than 2n spectral lines is not recommend for the identification problem, since the homogeneous part of the parameter update law is not exponentially stable. Hence parameter estimate might go to unbounded for even small enough disturbance.

2-2-5 Tuned Parameter (Input u(t) Has More Than 2n Spectral Lines)

In this case, if $W(t)$ is persistently exciting (this is almost true by the theorem 2-2-2.2), the homogeneous part of parameter update law is exponentially stable. We can conclude that the parameter estimate $\hat{\theta}(t)$ is bounded. Further more, we have

Proposition 2-2-5.1

Consider the parameter update law (2-2-2.2)

$$\dot{\hat{\theta}}(t)=-W(t)W^T(t)\hat{\theta}(t)+W(t)y(t)$$

Suppose input $u(t)$ is periodic and the regressor $W(t)$ is persistently exciting. Then $\hat{\theta}(t)$ is asymptotically periodic with same period as input $u(t)$.

Proof: See appendix 2-2-7.

Though the above discussion guarantees that the parameter estimate is bounded and is periodic if input $u(t)$ is periodic, it is not much useful and does not provided much insight. We will use concept of tuned parameter to facilitate our analysis. First we consider an easy case where input is sinusoid and periodic, i.e. we assume that

(A3) Input $u(t)=\sum_{i=1}^{m}\xi_i sinw_i t$, $m>n$ and is periodic with period T.

(A4) There exists some $w_{i_j}, j=1,\ldots,n$ such that $(w_{i_1},-w_{i_1},\ldots,w_{i_n},-w_{i_n})\in R^{2n}-V$, where $(w_{i_1},-w_{i_1},\ldots,w_{i_n},-w_{i_n})\in(w_1,-w_1,\ldots,w_m,-w_m)$.

Consider a nth order strictly proper transfer function

$$\hat{G}_T(s)=\frac{\gamma(s)}{\beta(s)}=\frac{\gamma_1 s^{n-1}+\ldots+\gamma_n}{s^n+\beta_1 s^{n-1}+\ldots+\beta_n}$$

If we are only interested in the steady-state, the output $y_T(t)$ of $\hat{G}_T(s)$ and the output $y(t)$ of plant $\hat{G}(s)$, under input $u(t)=\sum_{i=1}^{m}\xi_i sinw_i t$, become

$$y_T(t)=\sum_{i=1}^{m}\mathrm{Im}\frac{\gamma(jw_i)}{\beta(jw_i)}\xi_i\, e^{jw_i t}$$

and

$$y(t)=\sum_{i=1}^{m} \mathrm{Im}\frac{n(jw_i)}{d(jw_i)}\xi_i e^{jw_it}$$

Note that even though $\hat{G}_T(s)$ is unstable, the expression for $y_T(t)$ is still true provided that it has no poles on the jw axis (see lemma 2-2-3.2). Define an output error term by

$$e(t)=y_T(t)-y(t)$$

By the nature of the gradient type parameter update law (2-2-2.2), we see that the aim of the parameter update is to minimize $e^2(t)$, i.e. to minimize

$$\left(\sum_{i=1}^{m}\left(\mathrm{Im}\frac{\gamma(jw_i)}{\beta(jw_i)}\xi_i e^{jw_it}-\mathrm{Im}\frac{n(jw_i)}{d(jw_i)}\xi_i e^{jw_it}\right)\right)^2$$

$$=\left(\sum_{i=1}^{m}\mathrm{Im}\left(\frac{\gamma(jw_i)}{\beta(jw_i)}-\frac{n(jw_i)}{d(jw_i)}\right)\xi_i e^{jw_it}\right)^2$$

Obviously, when

$$\frac{\gamma(jw_i)}{\beta(jw_i)}=\frac{n(jw_i)}{d(jw_i)} \qquad i=1,\ldots,m \tag{2-2-5.1}$$

the minimum is achieved. In the case where input $u(t)$ has no more than $2n$ spectral lines, i.e. $m \le n$, from the results of last two sections , solutions of equation (2-2-5.1) always exist and are what we call tuned models. However if $m > n$, solution of (2-2-5.1) generally does not exist since a nth order transfer function may not match $2m > 2n$ points in the complex plane. This implies that in the case $m > n$, the tuned model defined in section 2-2-3 generally does not exist. In the following, we define a tuned parameter instead, to facilitate our analysis.

Consider the identifier in Fig. 2-2-2.1. Suppose $\hat{\theta}(t)^T=(\hat{a}(t)^T,\hat{b}(t)^T)$ is fixed, say $\hat{\theta}(t)\equiv\theta$, then

$$y_i(\theta,t)=W^T(t)\theta$$

Define output error $e(\theta,t)$ by

$$e(\theta,t)=y_i(\theta,t)-y(t)$$

As discussed above, there may not exist such a θ in the case of $m > n$ that

$$e^2(\theta, t) \equiv 0 \quad \text{for } all\ t \tag{2-2-5.2}$$

However, if there were a θ_T such that equation (2-2-5.2) would be true, then it is obvious that

$$\theta_T = \underset{\theta}{argmin} \frac{1}{T} \int_{nT}^{(n+1)T} e^2(\theta, t)\, dt \qquad \text{for } any\ n \tag{2-2-5.3}$$

Notice that since we are only interested in the tail properties, i.e. we neglect all exponentially decaying terms. Consequently $W(t)$ and $y(t)$ are periodic with period T and so does $e^2(\theta, t)$. This implies that the equation (2-2-5.3) is equivalent to

$$\theta_T = \underset{\theta}{argmin} \frac{1}{T} \int_0^T e^2(\theta, t)\, dt \tag{2-2-5.4}$$

Such θ_T is called a tuned parameter. The formal definition, which valids for general cases, is as follows:

Definition 2-2-5.2

Consider the plant $\hat{G}(s)$ (2-2-2.1) and the identifier in Fig.2-2-2.1. Then a constant vector $\theta_T \in R^{2n}$ is called a *tuned parameter* of $\hat{G}(s)$ (with input u(t)) if and only if θ_T satisfies

$$\theta_T = \underset{\theta}{argmin} \lim_{s \to \infty} \frac{1}{S} \int_0^s e^2(\theta, t)\, dt \tag{2-2-5.5}$$

Theorem 2-2-5.3

Consider the plant $\hat{G}(s)$ and the identifier in Fig.2-2-2.1. Suppose input u(t) satisfies the conditions (A3) and (A4). Then there is an unique tuned parameter θ_T for the plant $\hat{G}(s)$, and

$$\theta_T = R^{-1}{}_w\, R_{wy} \tag{2-2-5.6}$$

where

$$R_w = \lim_{s \to \infty} \frac{1}{S} \int_0^s W(t) W^T(t)\, dt \tag{2-2-5.7a}$$

$$R_{wy} = \lim_{s \to \infty} \frac{1}{S} \int_0^s W(t) y(t)\, dt \tag{2-2-5.7b}$$

Note these limits exist since $u(t)$ is periodic.

Proof: Since input is sinusoid and periodic, θ_T is the tuned parameter if and only if it satisfies equation (2-2-5.4). Notice that

$$\hat{W}^T(s) = \left(\frac{s^{n-1}}{\lambda(s)}\hat{y}(s),\ldots,\frac{1}{\lambda(s)}\hat{y}(s),\ \frac{s^{n-1}}{\lambda(s)}\hat{u}(s),\ldots,\frac{1}{\lambda(s)}\hat{u}(s)\right)$$

$$= \left(\frac{s^{n-1}}{\lambda(s)}\hat{G}(s),\ldots,\frac{1}{\lambda(s)}\hat{G}(s),\ \frac{s^{n-1}}{\lambda(s)},\ldots,\frac{1}{\lambda(s)}\right)\hat{u}(s)$$

$$= h(s)\,\hat{u}(s)$$

This implies that

$$e^2(\theta,t) = (y_i(\theta,t) - y(t))^2$$

$$= \left(\sum_{i=1}^{m} \mathrm{Im}\,(\theta^T h(jw_i) - \frac{n(jw_i)}{d(jw_i)})\,\xi_i\, e^{jw_i t}\right)^2$$

Let

$$\theta^T h(jw_i) - \frac{n(jw_i)}{d(jw_i)} = z(\theta, jw_i) = \rho(\theta, w_i)\, e^{j\alpha(\theta, w_i)}$$

where

$$\rho(\theta, w_i) = |\, z(\theta, jw_i)\,|$$

and

$$\alpha(\theta, w_i) = ang\ \ z(\theta, jw_i)$$

By the orthogonality of sinusoid functions, we have

$$\int_0^T e^2(\theta,t)\,dt = \int_0^T \left(\sum_{i=1}^{m} \rho(\theta,w_i)\,\xi_i \sin(w_i t + \alpha(\theta,w_i))\right)^2 dt$$

$$= \int_0^T \sum_{i=1}^{m} (\rho(\theta,w_i)\,\xi_i\, sin\,(w_i t + \alpha(\theta,w_i)))^2\,dt$$

$$= \frac{T}{2}\sum_{i=1}^{m} \rho^2(\theta,w_i)\,\xi_i^{\,2}$$

$$= T\ ((z(\theta, jw_1), \ldots, z(\theta, jw_m)) \begin{bmatrix} \frac{\xi_1^2}{2} & \cdot & 0 \\ 0 & \cdot & \vdots \\ \vdots & \cdot & 0 \\ 0 & \cdot & \frac{\xi_m^2}{2} \end{bmatrix} \begin{bmatrix} z^*(\theta, jw_1) \\ \vdots \\ z^*(\theta, jw_m) \end{bmatrix})$$

From the definition of $z(\theta, jw_i)$, it follows that

$$(z(\theta, jw_1), \ldots, z(\theta, jw_m)) = (\theta^T (h(jw_1), \ldots, h(jw_m)) - (\frac{n(jw_1)}{d(jw_1)}, \ldots, \frac{n(jw_m)}{d(jw_m)}))$$

Denote the above equation, for simplicity, as

$$Z = \theta^T H - Y$$

and let $Q = diag\ (\frac{\xi_1^2}{2}, \ldots, \frac{\xi_m^2}{2})$, we obtain

$$\frac{1}{T} \int_0^T e^2(\theta, t)\, dt = Z\ Q\ Z^*$$

$$= (\theta^T H - Y)\, Q\ (\theta^T H - Y)^*$$

The tuned parameter θ_T is defined to be a value such that the above equation achieves minimum. This is a typical least-square problem, it has an unique solution

$$\theta_T = (Re\ H\ Q\ H^*)^{-1} (Re\ H\ Q\ Y^*)$$

Where Re denotes the real parts. Note that the assupmtion (A4) and the theorem 2-2-2.2 guarantee that $(Re\ (H\ Q\ H^*))^{-1}$ exist. This completes the existence and uniqueness. Now by calculation, we have

$$\mathrm{Re}\, H\ Q\ H^* = \sum_{i=1}^{m} h(jw_i)\, h^*(jw_i) \frac{\xi_i^2}{2}$$

On the other hand, recall the definition of autocovariance R_w (2-2-5.6), we have

$$\mathrm{Re}\, H\ Q\ H^* = R_w$$

Similarly

$$\mathrm{Re}\, H\ Q\ Y^* = R_{wy}$$

i.e.

$$\theta_T = (Re\ H\ Q\ H^*)^{-1}(Re\ H\ Q\ Y^*) = R_w^{-1} R_{wy}$$

This completes the proof.

The theorem above holds for the sinusoid inputs only. However it is not difficult to show that the conclusion holds for general cases as stated in the following theorem.

Theorem 2-2-5.4

Consider the plant $\hat{G}(s)$ and the identifier in Fig.2-2-2.1. Suppose input u(t) is stationary and has more than 2n spectral lines such that W(t) is persistently exciting. Then there is an unique tuned parameter θ_T and

$$\theta_T = R_W^{-1} R_{Wy}$$

with R_W and R_{Wy} defined in (2-2-5.7).

Remarks:

(1) The tuned parameter is the value which minimize the mean squared power of the output error.

(2) The tuned parameter may also be interpreted as follows: Consider the parameter update law (2-2-2.2) with input $u(t)$ satisfying assumption (A4)

$$\dot{\hat{\theta}} = -WW^T\,\hat{\theta} + Wy$$

$$= -WW^T\,(\hat{\theta} - \theta_T) + (-WW^T\,\theta_T + Wy) \qquad (2\text{-}2\text{-}5.8)$$

Defining $\phi = \hat{\theta} - \theta_T$, it follows

$$\dot{\phi} = -WW^T\,\phi + (-WW^T\,\theta_T + Wy) \qquad (2\text{-}2\text{-}5.9)$$

Suppose θ_T be the tuned parameter, then from the definition (2-2-5.2) of θ_T,

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T (-WW^T\,\theta_T + Wy)\,dt = 0 \qquad (2\text{-}2\text{-}5.10)$$

We now see that the tuned parameter is a value such that if we consider the error equation (2-2-5.9), the time average of the driving term is zero. In this sense, we may say that error $\phi(t)$ is around the origin or that the parameter estimate $\hat{\theta}(t)$ is in the neighborhood of the tuned parameter θ_T. This fact may be seen clearer when the averaging technique is applied.

Theorem 2-2-5.5

Consider the plant $\hat{G}(s)$ (2-2-2.1) and the identifier with small gain $\varepsilon > 0$

$$\dot{\hat{\theta}}(t) = -\varepsilon W(t)(W^T(t)\hat{\theta}(t) - y(t))$$

Suppose input $u(t)$ is stationary and has $m \geq 2n$ spectral lines so that $W(t)$ is PE. Then the parameter estimate $\hat{\theta}(t)$ satisfies

$$\lim_{t \to \infty} \| \hat{\theta}(t) - \theta_T \| \leq \eta(\varepsilon)$$

for some class K function $\eta(\varepsilon)$ (see [52]).

Proof: See [41].

Remark:

The theorem guarantees that if the update gain is small then the parameter estimate converges to a ball centered at the tuned parameter θ_T with radius $\eta(\varepsilon)$.

2-2-6 Concluding Remarks

We have investigate the convergence of a reduced order identifier. The results are very general since there is almost no assumptions on the unmodeled part. If the unmodeled dynamics is small enough in the sense

$$|\hat{G}_0(jw) - \hat{G}(jw)| \le \delta \quad \text{for } \textit{any } w \in (-\infty, \infty)$$

Then from the results of chapter 3-4, we will see that the parameter estimate converges to a ball centered at true value (not tuned value, true value means true coefficients of the nominal model $\hat{G}_0(s)$, or Certainty Equivalence) with radius $\eta(\delta)$ which is linearly in δ.

So far, we only analyze the projection type update law, but the techniques could be easily extend to the Least-squares type update law.

One related issue to identification is that what is the purpose for identifying the nominal model $\hat{G}_0(s)$. From the results which we have presented, it can be seen that the estimated model might not be close to the nominal model $\hat{G}_0(s)$ even input contains resonable low frequency contents and the regressor is persistently exciting. However the estimated model could be a better approximation of the plant $\hat{G}(s)$ at these frequencies than the nominal model $\hat{G}_0(s)$ does. Let us see the following example:

Plant: $\hat{G}(s) = \dfrac{1}{s+1} \cdot \dfrac{s+11}{s+10}$

Nominal model: $\hat{G}_0(s) = \dfrac{1}{s+1}$

Estimated model (input frequency w*=2)*: $\dfrac{1.087}{s+0.957}$

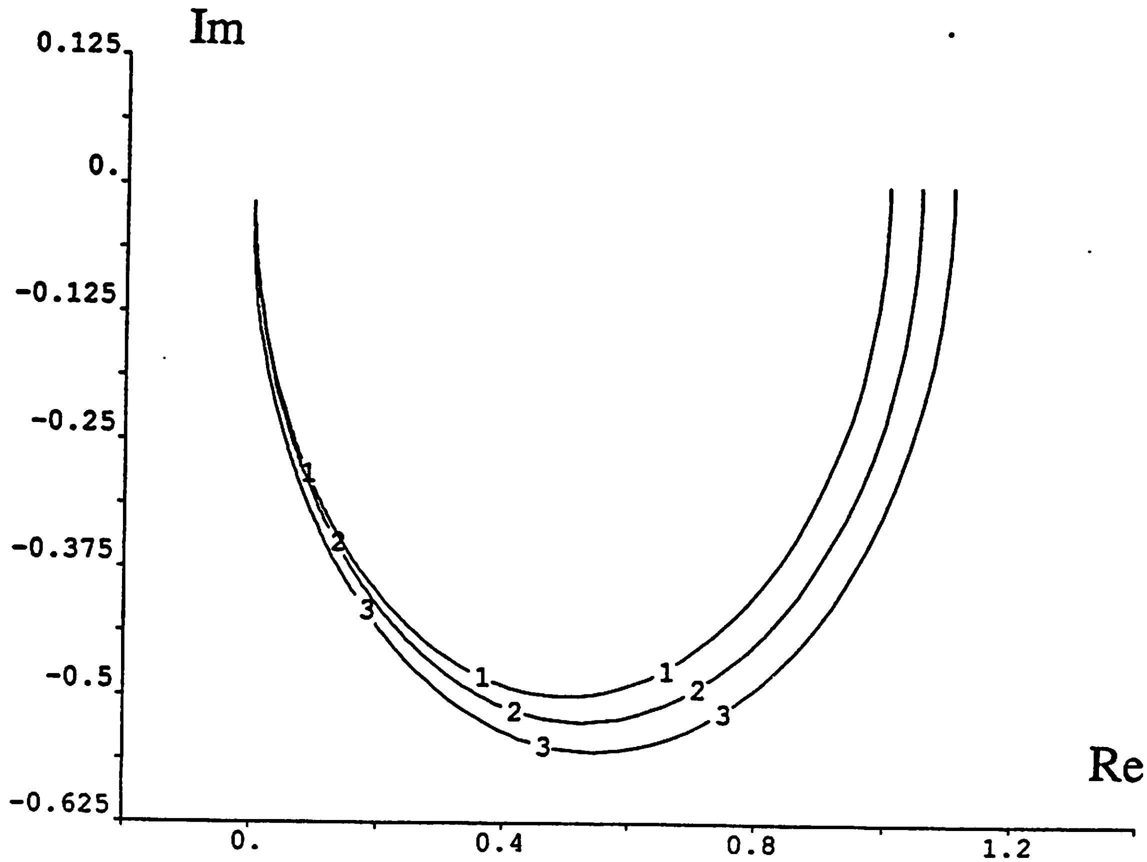

Fig. 2-2-6.1

-1-, nominal model

-2-, estimated model

-3-, plant

Fig.2-2-6.1 shows the Nyquist plots of these three transfer functions. We see that estimated model is closer to the plant than the nominal model does. Thus if the control is concerned, the estimated model could be a better model than the nominal model $\hat{G}_0(s)$ does for this purpose. We leave it as a further research problem.

2-2-7 Appendix to Chapter 2.2

Proof of The Proposition 2-2-5.1

We prove the general case. Consider a linear periodic system

$$\dot{x} = A(t)\,x + f(t)$$

where $x \in R^n$, $A(t+T) = A(t)$, $f(t+T) = f(t)$. Suppose $A(t)$ is asymptotically stable, then $x(t)$ is asymptotically periodic with period T for any bounded initial conditions.

Proof: By 2 steps.

Step 1. If the initial condition of above equation is

$$x(0) = (I - \Phi(T,0))^{-1} \int_0^T \Phi(T,\tau)\, f(\tau)\, d\tau$$

Then $x(t)$ is periodic with period T. Where $\Phi(t,\tau)$ is the state thansition matrix of $A(t)$. Note that $A(t)$ is asymptotically stable, so that $(I - \Phi(T,0))^{-1}$ exists. Now it suffices to show that

$$x(t+T) = x(t) \qquad \text{for } all\ t$$

We first show that $x(T) = x(0)$, this is because

$$x(T) = \Phi(T,0)\, x(0) + \int_0^T \Phi(T,\tau)\, f(\tau)\, d\tau$$

$$= (\Phi(T,0)\,(I - \Phi(T,0))^{-1} + I\,) \int_0^T \Phi(T,\tau)\, f(\tau)\, d\tau$$

$$= (I - \Phi(T,0))^{-1} \int_0^T \Phi(T,\tau)\, f(\tau)\, d\tau$$

$$= x(0)$$

It follows that

$$x(t+T) = \Phi(t+T,T)\, x(T) + \int_T^{t+T} \Phi(t+T,\tau)\, f(\tau)\, d\tau$$

Let $s = \tau - T$, we have

$$x(t+T) = \Phi(t,0)\,x(0) + \int_0^t \Phi(t+T,s+T)\,f(s+T)\,ds$$

$$= \Phi(t,0)\,x(0) + \int_0^t \Phi(t,s)\,f(s)\,ds$$

$$= x(t)$$

Step 2. For any bounded initial condition $x(0)$, we have

$$x(t) = \Phi(t,0)\,(x(0) - (I - \Phi(T,0))^{-1} \int_0^T \Phi(T,\tau)\,f(\tau)\,d\tau)$$

$$+ \Phi(t,0)\,(I - \Phi(T,0))^{-1} \int_0^T \Phi(T,\tau)\,f(\tau)\,d\tau + \int_0^t \Phi(t,\tau)\,f(\tau)\,d\tau$$

Note that the first term on the right hand side goes to zero asymptotically and the rest terms on the right hand side give a periodic solution. This completes the proof.

2-3 Global Stability Proofs for Indirect Adaptive Control Schemes

2-3-1 Problem Statement

A popular technique of adaptive control is the so-called indirect technique: a non-adaptive controller is designed parametrically i.e. the controller parameters are written as a function of plant parameters. This scheme is made adaptive by replacing the plant parameters in the design calculation by their estimates at time t, obtained from an on-line identifier. Reasons for the popularity of indirect adaptive controllers stem from the considerable flexibility in choice of both the controller and identifier. Global stability of indirect schemes have been shown in the discrete time case (Goodwin&Sin [25] , Anderson&Johnstone [3] Polak,Salcudean&Mayne [45]) but less so in the continuous time context. A recent paper of Elliot et al [19] uses random sampling to establish convergence results in the continuous time case. Other papers have assumed that the plant parameters lie in a convex set in which no unstable pole-zero cancellations occur (Kreisselmeier [35,36]).

In this section, we discuss a general, indirect adaptive control scheme for SISO continuous time systems using frequency analysis techniques. We show that when the reference input to the closed loop system is sufficiently rich, then the regressor vector of the identifier is persistently exciting so as to cause parameter convergence. In turn the controller is updated only when adequate information has been obtained for a ' meaningful' update. Thus, roughly speaking, the adaptive system consists of a fast parameter identification loop and a slow controller update loop. A sufficient richness condition on the exogenous reference input is used to give an insightful global stability proof with no restrictions on the parameter estimate lying in a convex set or lack of unstable pole-zero cancellations in the identifier.

In section 2-3-4, we show the specialization of our general scheme to a pole placement type adaptive controller.

The second contribution of this section is the application of our techniques to the adaptive stabilization of a SISO system using the factorization approach (factorization over the ring of stable, proper rational functions) that has proven to be a useful and elegant tool (see [17], [20]) for the study of robust multivariable design. Since it is known [51] that when the stable coprime factorization approach is used, a plant with unstable unmodeled dynamics is really no different from a plant with stable unmodeled dynamics as far as the effect of the unmodeled dynamics on the robustness of the system is concerned. We feel that our techniques lay the groundwork for obtaining an adaptive version of H^∞ optimal controller design by the factorization approach. In this context our work has contact with a recent paper of Ma & Vidyasagar [40]. In this section, we only discuss the SISO continuous time case, the extension to the discrete time case is trivial. We feel that our results could be extended to MIMO cases as well, if a good MIMO identifier structure is obtained.

Our major concern in this section is the proof of stability of the scheme with no assumptions of unmodelled dynamics, output disturbances in the plant. However, we note that the kind of stability we prove is exponential with its attendant margins of tolerance to both unmodelled dynamics and output disturbances as has been well documented in the literatures. Other techniques such as the use of a deadzone in the adaptation law may also be used as has been suggested in the context of discrete time adaptive control [26].

2-3-2 General Structure of The Indirect Adaptive Controller

The basic structure of the adaptive controller is as shown in Fig.2-3-2.1

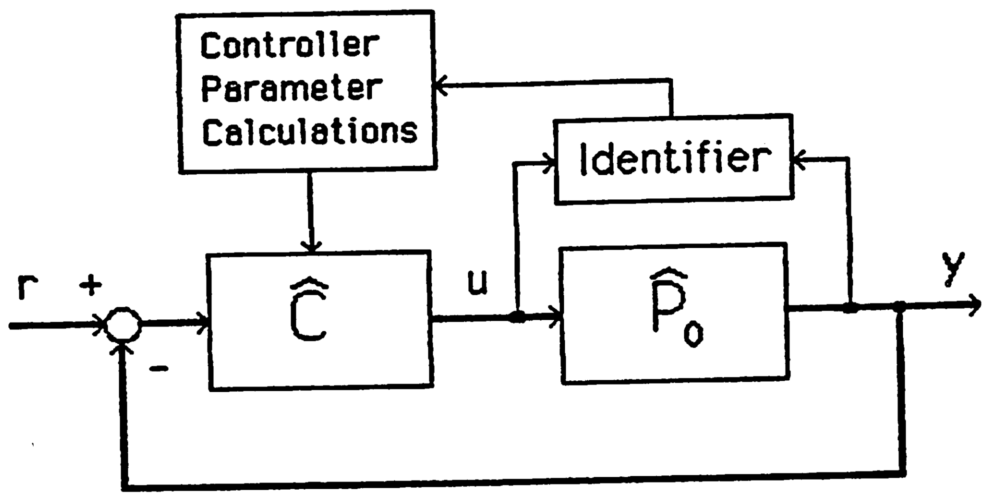

Fig. 2-3-2.1

The unknown plant is assumed to be of the form

$$\hat{P}_0(s)=n_p(s)/d_p(s)=\frac{\alpha_1 s^{n-1}+...+\alpha_n}{s^n+\beta_1 s^{n-1}+...+\beta_n} \qquad (2\text{-}3\text{-}2.1)$$

where $\hat{P}_0$ is a strictly proper transfer function with $n_p(s)$ and $d_p(s)$ coprime. We will assume that the order n of the plant is known and that the α_i and β_j are unknown. Note that some of the α_i's may be zero so that (2-3-2.1) can denote a plant of relative degree not necessarily 1. The proper *mth* order compensator is defined by

$$\hat{C}(s)=n_c(s)/d_c(s)=\frac{a_0 s^m+...+a_m}{b_0 s^m+...+b_m} \qquad (2\text{-}3\text{-}2.2)$$

The adaptive scheme proceeds as follows: the identifier gets an estimate of the plant parameters. The compensator design (pole placement, model reference,...) is performed assuming that the plant parameter estimate corresponds to the true parameter value (Certainty Equivalence). We will assume that there exists a unique choice of compensator $\hat{C}(s)$ of the form (2-3-2.2) for the plant estimate $\hat{P}_0$. The hope is that as $t \to \infty$ the identifier identifies the plant correctly and that the compensator converges asymptotically to the desired one. In this section, we discuss indirect adaptive control abstractly without restricting attention to any specific control scheme--pole placement, model reference, etc. In later sections, we specialize to a pole-placement type controller and a controller derived using the factorization approach.

Basically the most important element of the adaptive loop is the convergence of the identifier. We design an identifier which uses the input and output of the not necessarily stable plant as follows: the equation (2-3-2.1) relating the transform of the input and output of the plant can be written (with initial condition terms unspecified) as

$$s^n y(t) = \theta^{*T} v(t) \tag{2-3-2.3}$$

where s denotes the differentiator and

$$\theta^{*T} = (-\beta_1, \ldots -\beta_n, \alpha_1, \ldots, \alpha_n)$$

$$v^T(t) = (s^{n-1} y(t), \ldots, y(t), s^{n-1} u(t), \ldots, u(t))$$

Since the signal $v(t)$ involves differentiation of the input and output of the plant, we filter both side of (2-3-2.3) by the transfer function $1/(s+\alpha)^n$, $\alpha > 0$, to get

$$\frac{s^n}{(s+\alpha)^n} y(t) = \theta^{*T} W(t) \tag{2-3-2.4}$$

where

$$W^T(t) = \left(\frac{s^{n-1}}{(s+\alpha)^n} y(t), \ldots, \frac{1}{(s+\alpha)^n} y(t), \frac{s^{n-1}}{(s+\alpha)^n} u(t), \ldots, \frac{1}{(s+\alpha)^n} u(t)\right)$$

Note that the regressor $W(t)$ may be obtained by proper, stable filtering of the input and output of the plant. The equation error for identification of θ^* is developed as follows: let $\hat{\theta}(t)$ be the estimate of the parameter θ^* at time t. Then, define the equation error to be

$$e(t) = \hat{\theta}(t)^T W(t) - \frac{s^n}{(s+\alpha)^n} y(t) \tag{2-3-2.5}$$

If $\phi(t)$ denotes the parameter error $(\hat{\theta}(t)-\theta^*)$, then it follows that, up to exponentially decaying terms, we have

$$e(t)=\phi^T(t)W(t) \tag{2-3-2.6}$$

As is standard in the literature, we will in future drop the exponentially decaying terms. The interested reader may wish to confirm that the presence of such terms does not change any of the proofs (or conclusions) that follow.

The identification technique used is of the least squares type with resetting, given by

$$\dot{\hat{\theta}}(t)=\dot{\phi}(t)=-P(t)W(t)e(t) \tag{2-3-2.7a}$$

$$\dot{P}(t)=-P(t)W(t)W^T(t)P(t) \quad P(t_i)=\beta I>0 \tag{2-3-2.7b}$$

where $\{t_i\}=\{0,t_1,t_2\cdots\}$ will be specified shortly. It is easy to verify (using the Lyapunov function $\phi^T P^{-1}\phi$) that the parameter error ϕ is bounded even though y(t) may not be and further $\phi(t)\to 0$ asymptotically, if W(t) is persistently exciting, i.e. there exist $\alpha,\delta>0$ such that

$$\int_t^{t+\delta} W(\tau)W^T(\tau)d\tau \geq \alpha I \quad \text{for } all\ t$$

It has been shown in [43] that W is persistently exciting if u is rich enough i.e. u has no less than $2n$ spectral lines (assuming that u(t) is stationary).

The design of the compensator is based on the plant parameter estimate namely $\hat{\theta}(t)$. It would appear to be intuitive that if as $t\to\infty$, $\hat{\theta}(t)\to\theta^*$ that the time varying compensator will converge to the true compensator and that the closed loop system will be asymptotically stable. In this section, we do not deal with a specific compensator design; however the system of Fig.2-3-2-1 can be understood to be a time varying linear system which is asymptotically time invariant and stable. Such systems are themselves stable; more precisely, using standard Lyapunov function arguments, we have

Lemma 2-3-2.1

Consider a time varying system

$$\dot{x}=(A+\Delta A(t))x \tag{2-3-2.8}$$

where A is a constant matrix and $\Delta A(t)$ is time varying. Assume that $\|\Delta A(\cdot)\|_\infty$ is bounded and converges to zero as $t \to \infty$. Suppose that all eigenvalues of A lie in the open left half plane, then (2-3-2.8) is asymptotically stable. Furthermore , there exist $T, M, \lambda > 0$ such that the state transition matrix $\Phi(t,\tau)$ of the equation (2-3-2.8) satisfies

$$\|\Phi(t,\tau)\| \leq M \exp(-\lambda(t-\tau)) \quad \text{for } all \;\; t > \tau > T$$

Proof: See appendix 2-3-7.

2-3-3 Update Law and The Stability Proof

Though the update law (2-3-2.7a) and (2-3-2.7b) for the identifier is easily shown to be asymptotically convergent when W is persistently exciting, it is of practical importance to limit the update of the controller to instants when sufficiently new information has been obtained. The amount of information is measured through the 'information matrix'

$$\int_{t}^{t+\delta} W(\tau)W^T(\tau)d\tau$$

Thus given $\gamma > 0$, we choose a sequence of update times $\{t_i\}$, by $t_0=0$ and $t_{i+1}=t_i+\delta_i$, where δ_i satisfies

$$\delta_i := \underset{\Delta}{argmin} \int_{t_i}^{t_i+\Delta} WW^T d\tau \geq \gamma I \qquad (2\text{-}3\text{-}3.1)$$

The compensator $\hat{C}$ is held constant between t_i and t_{i+1}. Further, we assume that the compensator parameters are continuous functions of θ^*.

Remark:

(1) The idea of updating the controller only when new data becomes available was first proposed by [45] for the discrete time case. A similar idea was proposed by Elliot, et al [19], but they use a sequence of independent random variables to generate the update sequence.

(2) The update times are based on a monitoring of the excitation contained in the regressor W.

We may state the following lemma relating the richness of the reference signal r(t) in the scheme of Fig.2-3-2-1 to the convergence of the identifier.

Lemma 2-3-3.1 (Convergence of The Identifier)

Consider the system of Fig. 2-3-2.1 with identifier described in equation (2-3-2.7) and resetting times $\{t_i\}$ given by (2-3-3.1). Further assume that there is a unique choice of controller for each estimate of the plant and that the controller is updated only at $\{t_i\}$. If the input r(t) is stationary and has no less than $3n+m$ spectral lines, then the identifier parameter error converges to zero exponentially as $t \to \infty$. More precisely, there exists $0 < \rho < 1$ such that

$$||\phi(t_i)|| \le \rho^i ||\phi(0)|| \tag{2-3-3.2}$$

and $\{\delta_i = t_{i+1} - t_i\}$ is a bounded sequence.

Proof: By lemma A3 (in the Appendix 2-3-7), it is enough to show that $\{\delta_i\}$ is a bounded sequence. Suppose, for the sake of contradiction that $\{\delta_i\}$ is an unbounded sequence, then one of the two following scenarios occurs;

(i) There exist $i < \infty$ such that $\delta_i = \infty$, or

(ii) $\{\delta_i\} \to \infty$ *as* $i \to \infty$.

Consider the scenario (i) first. If (i) happens, then the system becomes time invariant after time t_i, since the controller is not updated. Consequently one obtains the transfer function (not necessarily stable) from r to u to be

$$\hat{H}_{ur} = \frac{\hat{n}_c(t_i) d_p}{n_p \hat{n}_c(t_i) + d_p \hat{d}_c(t_i)} = n/d \tag{2-3-3.3}$$

where $\hat{d}_c(t_i)$ and $\hat{n}_c(t_i)$ are the denominator and numerator of controller at time t_i respectively. Using (2-3-3.3), we may write the transfer function from r to $\underset{\sim}{W}$ to be

$$\hat{H}_{wr}(s) = \frac{n}{(s+\alpha)^n d_p d}(s^{n-1} n_p, \ldots, n_p, s^{n-1} d_p, \ldots, d_p)^T$$

Since the degree of n is (n+m), no more than (n+m) of the spectral lines of the input can correspond to zeros of the numerator polynomial. Even assuming that (n+m) of the spectral lines do, in fact, coincide with the zeros of n, we can see that under the assumption of n_p, d_p being coprime, W is persistently exciting. The proof of this for the stable case was given by Boyd and Sastry [11,12]. For the unstable case, the idea is as follows; we have a minimal state space realization of $\hat{H}_{wr}(s)$ as

$$\dot{x} = Ax + br$$

$$W = cx$$

where $A \in R^{k \times k}$ $(k \le 3n+m)$. Then, the persistency of excitation of x(t) follows from the hypothesis on the input r(t) and the fact that (A,b) is controllable (see, Nordstrom and Sastry [43]). Further, notice that, from the coprimeness of d_p and n_p, the rows of $\hat{H}_{wr}(s)$ are linearly independent. Now, since

$$\hat{H}_{wr}(s) = c(sI - A)^{-1}b$$

we see that c has full row rank i.e.

$$cc^T \geq \alpha I \quad \text{for } some \ \alpha > 0$$

Thus,

$$\int WW^T dt = c\int xx^T c^T dt \geq \gamma cc^T \geq \gamma\alpha I$$

where

$$\int xx^T dt \geq \gamma I$$

This implies that W(t) is persistently exciting. This fact however contradicts the assumption that $\delta_i = \infty$.

Now consider scenario (ii). First notice that when the plant parameters are known, then the closed loop system is time invariant and stable, so that we may write the following equation relating input r(t) to signal $W_0(t)$ ($W_0(t)$ means W(t) in the case when $\phi(t) \equiv 0$).

$$\dot{z}_0 = Az_0 + br$$

$$W_0 = cz_0$$

where A is a constant stable matrix. For the adaptive control situation, the plant parameters are unknown, i.e. parameter error $\phi(t) \neq 0$. However, we may write the following equation relating r(t) to W(t)

$$\dot{z}(t) = (A + \Delta A(t))z(t) + (b + \Delta b(t))r(t)$$

$$W = (c + \Delta c(t))z(t)$$

where $\Delta A(t)$, $\Delta c(t)$ and $\Delta b(t)$ are continuous functions of $\phi(t)$ and $\Delta A(t)$, $\Delta b(t)$ and $\Delta c(t) \rightarrow 0$ as $\phi(t) \rightarrow 0$. Now if scenario (ii) happens, we still have that $\phi(t) \rightarrow 0$ as $i \rightarrow \infty$ from lemma A3 (in the Appendix 2-3-7). It follows from lemma A1 (in the Appendix 2-3-7) that $W_0(t)$ and $W(t)$ are arbitrarily close when t is large enough. Then the persistency of excitation of W(t) follows as a consequence of the result of lemma A2 (in the Appendix 2-3-7) and the fact that $W_0(t)$ is persistently exciting. This, however, contradicts the assumption that $\delta_i \rightarrow \infty$ as $i \rightarrow \infty$. This completes the proof.

We are now in a position to prove the main theorem in this section.

Theorem 2-3-3.2 (Stability of The Closed Loop System)

Consider the system of Fig.2-3-2.1. Assume that the plant and compensator are described as in lemma 2-3-3.1. Suppose that input r(t) is stationary and has no less than $3n+m$ spectral lines, then the overall system is asymptotically time invariant and stable.

Proof: Follows from lemma 2-3-2.1 and lemma 2-3-3.1.

2-3-4 Application to Pole Placement Type Adaptive Control

In this section, we consider an indirect, adaptive pole placement scheme. Pole placement is easily described in the context of the Fig.2-3-2.1. Given a plant $\hat{P}_0$ of the form n_p/d_p as in (2-3-2.1), find a (n-1)th order compensator $\hat{C}$ so that the closed loop poles lie at the zeros of a given characteristic polynomial $d^*(s)$ of order (2n-1), i.e. find n_c, d_c to satisfy

$$n_c n_p + d_c d_p = d^* \tag{2-3-4.1}$$

When the plant $\hat{P}_0$ is unknown, the 'adaptive' pole placement scheme is mechanised by using the estimates $\hat{n}_p(t_i)$ and $\hat{d}_p(t_i)$ of the numerator and denominator polynomials respectively. It is easy to verify (see lemma A4, in the Appendix 2-3-7) that if $\hat{n}_p(t_i)$ and $\hat{d}_p(t_i)$ are coprime then there exist $\hat{n}_c(t_i)$ and $\hat{d}_c(t_i)$ of the order n-1 such that

$$\hat{n}_c(t_i)\hat{n}_p(t_i) + \hat{d}_c(t_i)\hat{d}_p(t_i) = d^* \tag{2-3-4.2}$$

The estimates for $\hat{n}_p(t_i)$ and $\hat{d}_p(t_i)$ follow from the plant parameter estimates $\hat{\theta}(t)$ of section 2-3-2 (the estimates of the coefficients of the denominator followed by those of the numerator). In analogy to the plant parameter vector θ^*, we have the parameter vector of the compensator

$$\theta_c = (b_0, \ldots, b_{n-1}, a_0, \ldots, a_{n-1}) \tag{2-3-4.3}$$

Recall from equation (2-3-2.2), with m=n-1, that the compensator is given by

$$\hat{C} = \frac{a_0 s^{n-1} + \ldots + a_{n-1}}{b_0 s^{n-1} + \ldots + b_{n-1}}$$

Further, to guarantee that $\hat{n}_p(t_i)$ and $\hat{d}_p(t_i)$ are coprime at t_i, we need to modify the definition of the update times as follows:

$$t_{i+1} = t_i + \delta_i \tag{2-3-4.4}$$

where δ_i is the smallest real number satisfying

$$(i) \quad \int_{t_i}^{t_i+\delta_i} w(t) w^T(t)\, dt \geq \gamma I \tag{2-3-4.5}$$

and

$$(ii)\ \hat{n}_p(t_i+\delta_i) \text{ and } \hat{d}_p(t_i+\delta_i) \ \textit{are coprime.} \tag{2-3-4.6}$$

More precisely (2-3-4.6) is verified by guaranteeing that the smallest singular value of the Sylvester matrix (A.11) (see appendix 2-3-7) measuring the extent of coprimeness exceeds a number $\sigma > 0$.

Then, we have

Theorem 2-3-4.1 (Convergence of The Pole Placement Scheme)

Consider the adaptive pole placement law (2-3-4.2) applied to the system of (2-3-2.1), along with the least squares identifier of (2-3-2.7) and the update sequence t_i defined by (2-3-4.4 & 2-3-4.6). Now, if the input r(t) is stationary and has no less than 4n-1 spectral lines. Then all signals in the loop are bounded and the characteristic polynomial of the closed loop system tends to $d^*(s)$. Moreover

$$||\hat{\theta}_c(t_i)-\theta_c|| \rightarrow 0 \quad \textit{exponentially}$$

Proof: The first half of the theorem is a direct consequence of lemmas 2-3-2.1 and 2-3-3.2. For the second half, note from (A.11) (in the Appendix 2-3-7) that

$$A(\hat{\theta}(t_i))\hat{\theta}_c(t_i)=d_* \tag{2-3-4.7}$$

with d_* the vector of coefficients of d^*.

It is easy to see from (A.11) (in the Appendix 2-3-7) that there is an $M_1>0$ such that

$$||A(\hat{\theta}(t_i))-A(\theta^*)|| \leq M_1||\hat{\theta}(t_i)-\theta^*|| \tag{2-3-4.8}$$

Now,

$$A(\theta^*)\theta_c=d_* \tag{2-3-4.9}$$

Subtracting (2-3-4.9) from (2-3-4.7) we get

$$-(A(\hat{\theta}(t_i))-A(\theta^*))\hat{\theta}_c(t_i)=A(\theta^*)(\hat{\theta}_c(t_i)-\theta_c)$$

Using the estimate

$$||\hat{\theta}_c(t_i)-\theta_c|| \leq ||A^{-1}(\theta^*)||\ ||A(\hat{\theta}(t_i))-A(\theta^*)||\ ||\hat{\theta}_c(t_i)||$$

Noting that $\hat{\theta}_c(t_i)$ is bounded (see equation (2-3-4.6) and the remark following it), we get

$$\| \hat{\theta}_c(t_i) - \theta_c \| \leq M_2 \| \hat{\theta}(t_i) - \theta^* \| \quad \text{for } \mathit{some}\ M_2 > 0 \qquad \text{(2-3-4.10)}$$

Since $\hat{\theta}(t_i)$ converges to θ^* exponentially, it follows that $\hat{\theta}_c(t_i) \rightarrow \theta_c$ exponentially.

2-3-5 Adaptive Stabilization Using The Factorization Approach to Controller Design

First we briefly review the (non-adaptive version of) factorization approach to controller design. Consider the linear time-invariant system shown in Fig. 2-3-5.1

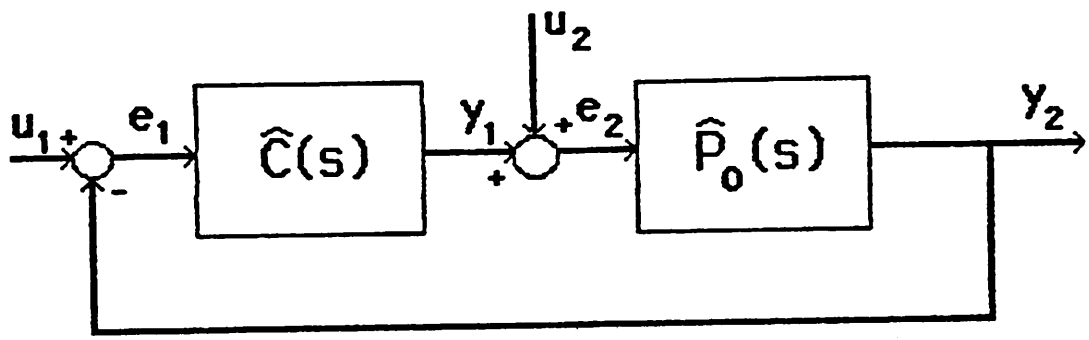

Fig. 2-3-5.1

The plant $\hat{P}_0(s)$ is defined as in equation (2-3-2.1) and the compensator $\hat{C}(s)$ as in (2-3-2.2). The equations relating e_1, e_2 to u_1, u_2 are

$$\begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = \frac{1}{1+\hat{P}_0\hat{C}} \begin{bmatrix} 1 & -\hat{P}_0 \\ \hat{C} & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \tag{2-3-5.1}$$

The system (2-3-5.1) is BIBO stable if and only if each of the four elements in (2-3-5.1) is stable, i.e. belongs to **R** the ring of proper, stable rational functions. The ring **R** is a more convenient ring than the ring of polynomials for the study of robust control systems, since a plant with unstable unmodeled dynamics is really no different from a plant with stable unmodeled dynamics.. Thus, we assume that $\hat{P}_0$ and $\hat{C}$ are factored coprimely in **R** (not uniquely!) as

$$\hat{P}_0(s) = d_p^{-1}(s)\, n_p(s)$$

$$\hat{C}_0(s)=d_c^{-1}(s)n_c(s) \qquad (2\text{-}3\text{-}5.2)$$

From (2-3-5.1) it follows that (for details see [51]) the system of Figure 2-3-5.1 is BIBO stable if and only if $(n_p n_c+d_p d_c)^{-1}\in$ R, or equivalently $n_p n_c+d_p d_c$ is a unimodular element of the ring R. Without loss of generality, then, we can state that a compensator stabilizes the system of Figure 2-3-5.1 if and only if

$$n_p n_c+d_p d_c=1 \qquad (2\text{-}3\text{-}5.3)$$

Let $(A\,,b\,,c\,)$ be a controllable canonical realization of $\hat{P}_0$ i.e

$$\hat{P}_0(s)=c\,(sI-A\,)^{-1}b \qquad (2\text{-}3\text{-}5.4)$$

with

$$A=\begin{bmatrix} 0 & 1 & & 0 \\ 0 & 0 & & . \\ . & . & & . \\ . & . & & 1 \\ -\beta_1 & -\beta_2 & & -\beta_n \end{bmatrix} \quad b=\begin{bmatrix} 0 \\ . \\ . \\ 0 \\ 1 \end{bmatrix} \qquad (2\text{-}3\text{-}5.5)$$

$$c=(\alpha_1,\ldots,\alpha_n)$$

If $f^T\in R^n$ and $l\in R^n$ are chosen so that $A_f=A-bf$ and $A-lc=A_l$ are stable(such a choice is possible by the minimality of the realization of (2-3-5.4) and (2-3-5.5)), then it is shown [, pg.83] that all the solutions of (2-3-5.3) can be written in the form

$$n_p=c\,(sI-A_l)^{-1}b \qquad (2\text{-}3\text{-}5.6)$$

$$d_p=1-c\,(sI-A_l)^{-1}l \qquad (2\text{-}3\text{-}5.7)$$

$$d_c=1+c\,(sI-A_f)^{-1}l-q\,(s)c\,(sI-A_f)^{-1}b \qquad (2\text{-}3\text{-}5.8)$$

$$n_c=f\,(sI-A_f)^{-1}l+q\,(s)(1-f\,(sI-A_f)^{-1}b\,) \qquad (2\text{-}3\text{-}5.9)$$

with $q(s)$ an arbitrary element of R which is chosen to meet other performance criteria (for instance, minimization of the disturbance to output map , obtaining the desired closed loop transfer function, optimal desensitization to unmodeled dynamics, etc...).

The optimal choice of $q(s)$ depends on the plant parameters. However, such a choice of q(s) may not be unique or depend continuously on plant parameters. This may give rise to difficulties in applying the method discussed in section 2-3-2, since the design of the compensator may not be unique as required by the assumptions of the scheme. We defer this to further investigation. However, if our only concern is the problem of adaptive stabilization of the unknown plant, then any fixed $q(s) \in$ **R** will do. For simplicity, we fix $q(s)=0$ in what follows. Note this implies that the compensator $\hat{C}(s)$ is of order n (see equations (2-3-5.8) & (2-3-5.9)).

We now discuss the adaptive version of factorization approach to controller design. The objective is to design a compensator $\hat{C}$ adaptively, i.e. based on the estimate $\hat{\theta}$ of plant parameters, using the factorization approach, so that the closed loop system is asymptotically time invariant and stable. In what follows, we assume that $u_2(t) \equiv 0$.

The identifier and compensator update times $\{t_i\}$ are defined as in (2-3-3.7) and (2-3-4.4 to 2-3-4.6) respectively. The first difficulty in choosing the compensator is the choice of $l(t_i)$ and $f(t_i)$ at time t_i (see equations (2-3-5.8) and (2-3-5.9)). From linear system theory, we know that a controllable canonical realization of plant $\hat{P}_0(s)$.

$$\dot{x}=Ax+bu$$

$$y=cx$$

can be transformed through a linear change of coordinatea $\bar{x}=Mx$ to get the observable canonical form of $\hat{P}_0(s)$.

$$\dot{\bar{x}}=MAM^{-1}\bar{x}+Mbu=\bar{A}\bar{x}+\bar{b}u$$

$$y=cM^{-1}\bar{x}=\overline{cx} \qquad (2\text{-}3\text{-}5.10)$$

with

$$\bar{A}=\begin{bmatrix} 0 & & 0 & -\beta_1 \\ 1 & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & 0 & \cdot \\ 0 & & 1 & -\beta_n \end{bmatrix} \quad \bar{b}=\begin{bmatrix} \alpha_1 \\ \cdot \\ \cdot \\ \cdot \\ \alpha_n \end{bmatrix}$$

$$\bar{c}=(0,\ldots,0,1)$$

Then for (any) given Hurwitz polynomial

$$p(s)=s^n+p_1s^{n-1}+\cdots+p_n \tag{2-3-5.11}$$

there exists a vector

$$\bar{l}^T=(\bar{l}_1,\ldots,\bar{l}_n)=(p_1,\ldots,p_n)-(\beta_1,\ldots,\beta_n)$$

such that the matrix

$$\bar{A}-\bar{l}\bar{c}=\begin{bmatrix} 0 & 0 & -p_1 \\ 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \dot{0} & \dot{1} & -p_n \end{bmatrix}$$

is stable and has a characteristic polynomial p(s). Define

$$l=M^{-1}\bar{l}$$

With this definition, it is easy to see that (A-lc) is stable and has characteristic polynomial p(s).

Now the controller design procedure can be stated as follow:

(Step1)

At time t_i, the parameter estimate $\hat{\theta}(t_i)$ generated by identifier is used to obtain the estimates $A(t_i)$, $b(t_i)$, and $c(t_i)$.

(Step2)

By calculation, we obtain $M^{-1}(t_i)$ as described in (2-3-5.10). Define

$$f(t_i)=\bar{l}^T(t_i)=(p_1,\ldots,p_n)-(\beta_1(t_i),\ldots,\beta_n(t_i)) \tag{2-3-5.12}$$

with $(p_1,\ldots,p_n)$ as defined in (2-3-5.11) and

$$l(t_i)=M^{-1}(t_i)\bar{l}(t_i) \tag{2-3-5.13}$$

We now see that the matrices

$$A_l(t_i)=A(t_i)-l(t_i)c(t_i)$$

and

$$A_f(t_i)=A(t_i)-b(t_i)f(t_i)$$

are stable with characteristic polynomial p(s). Furthermore, $f(t_i)$ and $l(t_i)$ converge to some constant vectors as $i \to \infty$.

(Step3)

Choose a compensator $\hat{C}(t_i)=n_c(t_i)d_c^{-1}(t_i)$ as follows

$$n_c(t_i)=f(t_i)(sI-A_f(t_i))^{-1}l(t_i) \tag{2-3-5.14}$$

$$d_c(t_i)=1+c(t_i)(sI-A_f(t_i))^{-1}l(t_i) \tag{2-3-5.15}$$

This compensator can be easily implemented. Then, as expected, we have

Theorem 2-3-5.1 (Convergence of The Overall System)

Assume that the identifier and controller update described above are applied to the plant $\hat{P}_0(s)$ (2-3-2.1). Suppose that the input r(t) is stationary and has no less than $4n$ spectral lines. Then, the closed loop system is asymptotically stable and all signals are bounded.

2-3-6 Concluding remarks

This section has presented a proof of global stability for indirect adaptive control. In the section, only two applications (pole placement and factorization approach) have been discussed, however the results are applicable to several kinds of controller design methodologies. The key assumption is a richness condition on the reference input. To our knowledge, this is the first verification of the persistency of excitation of the regressor signal in the closed loop (which is time varying) without the use of an artificial random sampling signal (see [19]) for the continuous time case. We show persistency of excitation without preassuming the boundedness of the signals. Boundedness of all signals and the convergence of the compensator in turn follow from the convergence of the identifier, which is a direct consequence of the persistency of excitation of the signal in the identification loop.

The scheme presented here offers a great deal of flexibility in the controller design and allows for very general richness conditions on the reference input. The results of this section are easily extendable to the discrete time case.

2-3-7 Appendix to Chapter 2.3

In this appendix, we prove some lemmas of use in the main body of the section.

Lemma A1

Consider the following linear systems

$$\dot{z}_0 = Az_0 + br \tag{A1}$$

$$\dot{z} = (A + \Delta A(t))z + (b + \Delta b(t))r \tag{A2}$$

with A stable and ΔA , Δb both bounded and converging to zero as $t \to \infty$. Assume that the input r(t) is bounded. Then given $\varepsilon > 0$, there exists $k > 0$ (k is independent of the choice ε) and a $T(\varepsilon)$ such that

$$||z(t) - z_0(t)|| \le \varepsilon k \quad \text{for all } t \ge T \tag{A3}$$

Proof: From lemma 2.1, it follows that (A2) is asymptotically stable and that there exists T_1 such that the state transition matrix of (A2) satisfies

$$||\Phi(t,\tau)|| \le M \exp -\lambda(t-\tau)$$

for some $M, \lambda > 0$ and for all $t > \tau > T_1$. Using this estimate it is easy to show that z(t) is bounded. Defining the error $e(t) := z(t) - z_0(t)$ we see that

$$\dot{e} = Ae + \Delta Az + \Delta br$$

Using the facts that ΔA , $\Delta b \to 0$ as $t \to \infty$; that z,r are bounded and A is stable, it is easy to establish (A3).

Lemma A2

Suppose that $W_0(t) \in R^n$ is persistently exciting, i.e there exist $\delta, \alpha > 0$ such that

$$\int_s^{s+\delta} W_0 W_0^T \, dt \ge \alpha I \qquad \text{for all } s$$

Then any signal $W \in R^n$ satisfying

$$||W(t)-W_0(t)||<\alpha/(\delta)^{1/2}$$

is also persistently exciting.

Proof: Can be found in [8].

Lemma A3

Consider the least squares identification algorithm described by (2-3-2.7) with resetting sequence $\{0,t_1,t_2,...\}$, that is

$$\dot{\phi}=-PWW^T\phi \tag{A4}$$

and

$$\dot{P}^{-1}=WW^T \quad t\neq 0,t_1,t_2,... \tag{A5}$$

$$P^{-1}(t_i^+)=\alpha I \quad t=0,t_1,t_2,... \tag{A6}$$

If W is persistently exciting, that is

$$\int_{t_i}^{t_{i+1}} WW^T dt \geq \gamma I \quad \text{for } all\ t_i \tag{A7}$$

Then, there exists $1>\rho>0$ such that

$$||\phi(t_i)||\leq\rho^i||\phi(0)|| \tag{A8}$$

Proof: Note that for $t\neq\{0,t_1,...\}$,

$$\frac{d}{dt}P^{-1}\phi=0$$

Thus

$$P^{-1}(t_i^-)\phi(t_i)=P^{-1}(t_{i-1}^+)\phi(t_{i-1})$$

so that

$$\phi(t_i)=\alpha P(t_i^-)\phi(t_{i-1})$$

and we get

$$||\phi(t_i)|| \le (\frac{\alpha}{\alpha+\gamma})||\phi(t_{i-1})|| \tag{A9}$$

In the last step we use equation (A7). Recursion on (A9) yields the conclusion (A8).

Lemma A4

Consider two coprime polynomials d_p monic of order n and n_p of order n-1. Then given an arbitrary polynomial d^* of order 2n-1, there exist unique polynomials n_c and d_c of order n-1 so that

$$n_c n_p + d_c d_d = d^* \tag{A10}$$

Proof: Is a standard result from algebra (see [25]). It is useful for the proof of theorem 2-3-4.1 to note that if

$$d_p = s^n + \beta_1 s^{n-1} + ... + \beta_n$$

$$n_p = \alpha_1 s^{n-1} + ... + \alpha_n$$

$$n_c = a_1 s^{n-1} + ... + a_n$$

$$d_c = b_1 s^{n-1} + ... + b_n$$

Then, the linear equation relating the coefficients of n_c , d_c to those of d^* is

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 0 & & 0 \\ \beta_1 & 1 & 0 & \cdots & & \alpha_1 & 0 & & \cdot \\ \beta_2 & \beta_1 & 1 & \cdots & & \alpha_2 & \alpha_1 & & \cdot \\ \cdot & \cdot & \cdot & \cdots & & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & \cdots & & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & \cdots & 1 & \cdot & \cdot & & \cdot \\ \beta_n & \beta_{n-1} & \cdot & \cdots & \beta_1 & \alpha_n & \alpha_{n-1} & & \alpha_1 \\ 0 & \beta_n & \cdot & \cdots & & 0 & \alpha_n & & \cdot \\ 0 & 0 & \beta_n & \cdots & & 0 & 0 & & \cdot \\ \cdot & \cdot & 0 & \cdots & & \cdot & \cdot & & \cdot \\ 0 & 0 & 0 & \cdots & \beta_n & 0 & 0 & & \alpha_n \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_n \\ a_1 \\ a_2 \\ \cdot \\ \cdot \\ a_n \end{bmatrix} = \begin{bmatrix} d_1^* \\ d_2^* \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ d_{2n}^* \end{bmatrix} \tag{A11}$$

i.e.

$$A(\theta^*)\theta_c = d_*$$

Proof of the Lemma 2-3-2.1

Since A is stable, there is a positive definite matrix P such that

$$A^T P + PA = -I$$

and by the hypothesis $\Delta A(t) \to 0$ as $t \to \infty$, so that there is a $T > 0$ such that

$$I - \Delta A^T(t)P - P\,\Delta A(t) \geq \frac{1}{2} I \qquad \text{for } all\ t \geq T$$

Now define a Lyapunov function by

$$v(x) = x^T P\, x$$

then the derivative of $v(x)$ along the solution of equation (2-3-2.8) becomes

$$\dot{v} = -x^T (I - \Delta A^T P - P \Delta A^T)\, x$$

$$\leq \frac{-1}{2} x^T x \qquad \text{for } all\ t \geq T$$

The conclusion follows easily from the Lyapunov theorem [52].

Chapter 3 Use of Prior Information

3-1 Introduction

A great deal of effort has been devoted to establishing conditions for robust stability of adaptive control algorithms. There are two sets of approaches to this issue: In the first approach, an internal signal in the adaptive loop is made persistently exciting to guarantee exponential stability of the scheme. Robustness of the scheme follows as a consequence of the robustness of exponential stability. In the second approach, the adaptive algorithm is modified, using for instance, a deadzone, forgetting factor or σ and e modification in the adaptation law to prevent the algorithm from responding to spurious signals such as those arising from noise and unmodeled dynamics. Both approaches model the plant to be controlled as being completely unknown. In this section, we discuss the identification and control of systems which are partially known (in a sense that is made explicit shortly). It seems intuitively plusible that the identification and control algorithms could be robust if this prior information could be incorporated into the adaptive controller. It is of course clear that one could neglect the prior information embodied in the system and still be able to identify and/or control the system. However, usage of the particular prior information results in the identification of a fewer number of unknown parameters and consequently faster convergence rate and better transient performance as we will see in this section. With this as motivation, we discuss the problem of adaptive identification and control of 'partially known ' systems. Our work presented here was especially motivated by the recent Ph.D thesis of Dasgupta [14].

3-2 Model for a Class of Systems with Prior Information

The system to be identified and/or controlled is a single input-single output linear time-invariant system of the form

$$\frac{y}{u} = \frac{n_0 + \sum_{i=1}^{l} \alpha_i n_i}{d_0 - \sum_{j=1}^{k} \beta_j d_j} \tag{3-2.1}$$

where α_i , β_j's are unknown constants, n_i , d_j 's are known polynomials in z^{-1} (discrete time) or s (continuous time). The model (3-2.1) is general enough for several kinds of 'partially known' systems. We give the following examples:

(1) Network functions of RLC circuits with some elements unknown. Consider the circuit of Fig.3-2.1 with the resistor R unknown (drawn as a two port to exhibit the unknown resistance).

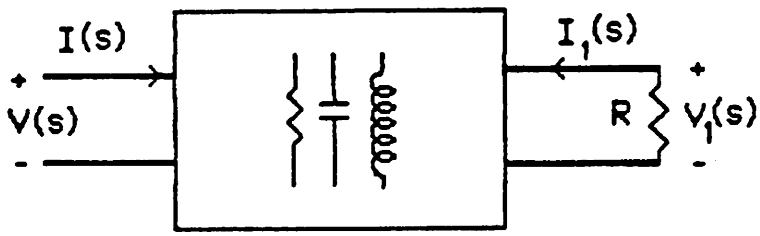

Fig. 3-2.1

If the short circuit admittance matrix of the two port in Fig.3-2.1 is

$$\begin{bmatrix} I \\ I_1 \end{bmatrix} = \begin{bmatrix} y_{11}(s) & y_{12}(s) \\ y_{21}(s) & y_{22}(s) \end{bmatrix} \begin{bmatrix} V \\ V_1 \end{bmatrix} \tag{3-2.2}$$

that a simple calculation yields the admittance function to be

$$\frac{I(s)}{V(s)} = \frac{y_{11}+R\,(y_{11}y_{22}-y_{12}y_{21})}{1+Ry_{22}}$$

which is of the form (3-2.1). Circuits with more than one unknown element can be drawn as multiports to show that the admittance function is of the form (3-2.1)

(2) Interconnection of several systems with unknown interconnection gains. Consider the simple discrete time configuration of Fig.3-2.2 with the polynomials $n(z)$ and $d(z)$ known.

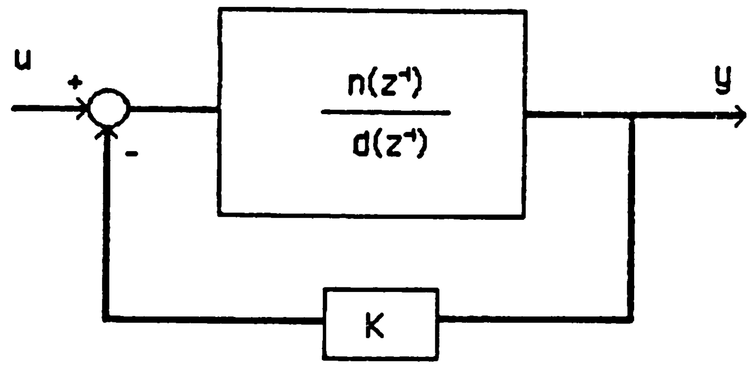

Fig. 3-2.2

The closed loop transfer function is of the form (3-2.1) $\dfrac{n(z)}{d(z)+kn(z)}$.

(3) Systems with some known poles and zeros. Consider the system of Fig.3-2.3, with unknown plant but known actuator and sensor dynamics

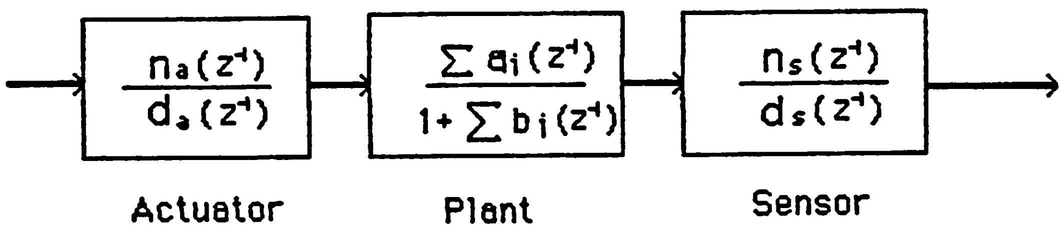

Fig. 3-2.3

The overall transfer function may be written as

$$\frac{\sum a_i n_a n_s z^{-i}}{d_a d_s + \sum b_j d_a d_s z^{-j}}$$

which is of the form (3-2.1) since $n_a n_s z^{-i}$, $d_a d_s z^{-j}$ known.

(4) Classical transfer function model, i.e. a plant of the form

$$\hat{G}(s)=\frac{\alpha_1 s^{m-1}+\ldots+\alpha_m}{s^n+\beta_1 s^{n-1}+\ldots+\beta_n}$$

with $m \le n$ and α_i , β_j unknown, can be stated in terms of the set up of (3-2.1) by choosing

$$n_0(s)=0,\ d_0(s)=s^n$$

$$n_i(s)=s^{m-i} \quad i=1\ldots m$$

$$d_j(s)=-s^{n-j} \quad j=1\ldots n$$

In this section we will discuss both continuous and discrete time systems of the form of (3-2.1). Our methods are extensions of those proposed by Goodwin et al in [24], Narendra et al in [39] and Sastry in [48]. The methods of proof are identical. The novelty of our paper is the set up in which the methods are applied.

3-3 Discrete Time System Identification

3-3-1 Identifiability Condition

In this section ,we will discuss the identifiability condition for some discrete time 'partially known' systems.

First consider a system described by

$$\frac{y(z^{-1})}{u(z^{-1})}=\frac{n_0(z^{-1})+\sum_{i=1}^{l}k_i n_i(z^{-1})}{d_0(z^{-1})-\sum_{i=1}^{l}k_i d_i(z^{-1})}=\frac{n(z^{-1})}{d(z^{-1})} \tag{3-3-1.1}$$

where k_i's are unknown parameters, $d_i(z^{-1})$ and $n_i(z^{-1})$ are known polynomials in the unit delay operator z^{-1}.

$$d_0(z^{-1})=1+d_{01}z^{-1}+...+d_{0n}z^{-n}$$

$$d_i(z^{-1})=d_{i1}z^{-1}+...+d_{in}z^{-n} \quad i=1,...l \tag{3-3-1.2}$$

$$n_i(z^{-1})=n_{i1}z^{-1}+...+n_{im}z^{-m} \quad i=0,1...l \tag{3-3-1.3}$$

Assume $n(z^{-1})$ and $d(z^{-1})$ are coprime. Notice that it is a special case of the form (3-2.1).

Definition 3-3-1.1

A system with some unknown parameters is said to be identifiable if and only if there exist some inputs u(t) such that the unknown parameters can be uniquely determined based on input-output measurements.

Theorem 3-3-1.2

The system (3-3-1.1) is identifiable if and only if the following matrix is full column rank.

$$D=\begin{bmatrix} d_{11} & d_{21} & & d_{l1} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ d_{1n} & d_{2n} & & d_{ln} \\ n_{11} & n_{21} & & n_{l1} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ n_{1m} & n_{2m} & & n_{lm} \end{bmatrix} \tag{3-3-1.4}$$

Proof: Rewrite (3-3-1.1) in the form

$$[d_0(z^{-1})y(z^{-1})-n_0(z^{-1})u(z^{-1})]=\sum_{i=1}^{l}k_i[d_i(z^{-1})y(z^{-1})-n_i(z^{-1})u(z^{-1})]$$

Define

$$z_0(z^{-1})=d_0(z^{-1})y(z^{-1})-n_0(z^{-1})u(z^{-1}) \tag{3-3-1.5}$$

and the regressor vector $\psi(t)$ and parameter vector θ by

$$\psi^T(t)=(y(t-1),...y(t-n),u(t-1),...u(t-m)) \tag{3-3-1.6}$$

$$\theta^T=(k_1,k_2...k_l) \tag{3-3-1.7}$$

Then the system can be written as

$$z_0(t)=\psi^T(t)D\,\theta \tag{3-3-1.8}$$

The necessity may be readily seen, since if D is not full column rank, then any $\theta\in\theta+Null\ D$ will give same transfer function. This situation corresponds intuitively to the case, in which there exists a $\theta=(k_1,...k_l)^T$ such that

$$\sum_{i=1}^{l}k_in_i(z^{-1})\equiv\sum_{i=1}^{l}k_id_i(z^{-1})\equiv 0$$

Now, we give the proof of sufficiency. By assumption, the sufficient richness of the input u(t) implies the persistency of excitation of $\psi(t)$ (see Bai and Sastry [9]) i.e. there exists $\alpha>0$ and $p\in Z_+$, such that

$$\sum_{t=t_0}^{t_0+p-1}\psi(t)\psi^T(t)\geq\alpha I \qquad \text{for } all\ t_0 \tag{3-3-1.9}$$

then following inequality is obtained,

$$\left(\begin{bmatrix}\psi^T(t_0)\\ \vdots\\ \psi^T(t_0+p-1)\end{bmatrix}D\right)^T\begin{bmatrix}\psi^T(t_0)\\ \vdots\\ \psi^T(t_0+p-1)\end{bmatrix}D=D^T\sum_{t=t_0}^{t_0+p-1}\psi(t)\psi^T(t)D>0$$

By the linear algebra, we know that the equation (3-3-1.8) has an unique solution for θ. This completes the proof.

For the general case of the systems described by equation (3-2.1)

$$\frac{y(z^{-1})}{u(z^{-1})}=\frac{n_0(z^{-1})+\sum_{i=1}^{l}\alpha_i n_i(z^{-1})}{d_0(z^{-1})-\sum_{j=1}^{k}\beta_j d_j(z^{-1})}=\frac{n(z^{-1})}{d(z^{-1})} \tag{3-3-1.10}$$

The corollary 3-3-1.3 follows.

Corollary 3-3-1.3

The system (3-3-1.10) is identifiable if and only if the following matrix is full column rank.

$$D=\begin{bmatrix} d_{11} & d_{k1} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ d_{1n} & d_{kn} & 0 & 0 \\ 0 & 0 & n_{11} & n_{l1} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & n_{1m} & n_{lm} \end{bmatrix}$$

Let us now digress a little bit and consider the another type of 'partially known' system of the form

$$y(z^{-1})=(k_1\frac{n_1}{d_1}+...+k_l\frac{n_l}{d_l})u(z^{-1}) \tag{3-3-1.11}$$

where $n_i(z^{-1})$ and $d_i(z^{-1})$ are nonzero known coprime polynomials in the unit delay operator z^{-1},

$$n_i(z^{-1})=n_{i1}z^{-1}+...+n_{im}z^{-m}$$

$$d_i(z^{-1})=1+d_{i1}z^{-1}+...+d_{in}z^{-n}$$

and the $k_i's$ are unknown parameters.

Define

$$h_i(z^{-1})=\frac{n_i}{d_i}u(z^{-1}) \quad i=1,2...l \tag{3-3-1.12}$$

Then it follows that

$$h_i(t)=-d_{i1}h_i(t-1)-...-d_{in}h_i(t-n)+n_{i1}u(t-1)+...+n_{im}u(t-m) \quad i=1...l$$

so that

$$y(t)=(h_1(t),...,h_l(t))\begin{bmatrix}k_1\\ \cdot\\ \cdot\\ k_l\end{bmatrix}=\psi^T(t)D\theta \tag{3-3-1.13}$$

with

$$\psi^T(t)=(h_1(t-1)...h_1(t-n)....h_l(t-1)...h_l(t-n),u(t-1)...u(t-m))$$

$$D=\begin{bmatrix}-d_{11} & 0 & 0\\ \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot\\ -d_{1n} & 0 & 0\\ \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot\\ 0 & 0 & -d_{l1}\\ \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot\\ 0 & 0 & -d_{ln}\\ n_{11} & n_{21} & n_{l1}\\ \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot\\ n_{1m} & n_{2m} & n_{lm}\end{bmatrix} \tag{3-3-1.14}$$

Theorem 3-3-1.4

The necessary condition for system (3-3-1.11) to be identifiable is that the matrix D defined in (3-3-1.14) be of full column rank. The sufficient condition for the system (3-3-1.11) to be identifiable is that

$$rank\begin{bmatrix}\hat{d}_1 & 0 & 0 & \hat{n}_1\\ 0 & \hat{d}_2 & 0 & \hat{n}_2\\ \cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & \hat{d}_l & \hat{n}_l\end{bmatrix}=l \quad \text{for } all\ z \tag{3-3-1.15}$$

where

$$\hat{d}_i=z^n+d_{i1}z^{n-1}+...+d_{in} \tag{3-3-1.16}$$

$$\hat{n}_i=n_{i1}z^{m-1}+...+n_{im} \tag{3-3-1.17}$$

Proof: The proof of the necessary condition is clear. Let us prove the sufficiency. A direct consequence of condition (3-3-1.15) is that sufficient richness of input u(t) implies the persistency of excitation of the regressor vector $\psi(t)$. Since

$$\psi(t)=A\,\psi(t-1)+bu(t) \tag{3-3-1.18}$$

where A and b are similar to those in [9], and the condition (3-3-1.15) guarantees the reachability of the system (3-3-1.18). Then persistency of excitation of $\psi(t)$ follows (see Bai and Sastry [9]). The rest of the proof is similar to the proof of theorem (3-3-1.2).

We now discuss multivariable extensions. Let us restrict ourself to the system described by

$$\begin{bmatrix} y_1(z^{-1}) \\ \cdot \\ \cdot \\ y_p(z^{-1}) \end{bmatrix} = \begin{bmatrix} k_{11}\frac{n_{11}}{d_{11}} & \cdot & k_{1l}\frac{n_{1l}}{d_{1l}} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ k_{p1}\frac{n_{p1}}{d_{p1}} & \cdot & k_{pl}\frac{n_{pl}}{d_{pl}} \end{bmatrix} \begin{bmatrix} u_1(z^{-1}) \\ \cdot \\ \cdot \\ u_l(z^{-1}) \end{bmatrix} \tag{3-3-1.19}$$

where $n_{ij}(z^{-1})$ and $d_{ij}(z^{-1})$ are known nonzero polynomials

$$n_{ij}(z^{-1})=n_{ij}(1)z^{-1}+...+n_{ij}(m)z^{-m} \tag{3-3-1.20}$$

$$d_{ij}(z^{-1})=1+d_{ij}(1)z^{-1}+...+d_{ij}(n)z^{-n} \tag{3-3-1.21}$$

k_{ij}'s are unknown parameters. Define

$$h_{ij}(z^{-1})=\frac{n_{ij}}{d_{ij}}u_j(z^{-1}) \tag{3-3-1.22}$$

$$\begin{aligned}\psi_i^T(t)=(&h_{i1}(t-1)...h_{i1}(t-n)...h_{il}(t-1)...h_{il}(t-n),\\ &u_1(t-1)...u_1(t-m)...u_l(t-1)...u_l(t-m))\end{aligned} \tag{3-3-1.23}$$

and

$$D_i=\begin{bmatrix} -d_{i1}(1) & 0 \\ \vdots & \vdots \\ -d_{i1}(n) & 0 \\ \vdots & \vdots \\ 0 & -d_{il}(1) \\ \vdots & \vdots \\ 0 & -d_{il}(n) \\ n_{i1}(1) & 0 \\ \vdots & \vdots \\ n_{i1}(m) & 0 \\ \vdots & \vdots \\ 0 & n_{il}(1) \\ \vdots & \vdots \\ 0 & n_{il}(m) \end{bmatrix} \tag{3-3-1.24}$$

The system (3-3-1.19) can be written as

$$y_i(t)=\psi_i^T(t)D_i\theta^i \quad i=1...p \tag{3-3-1.25}$$

with

$$\theta^i=(k_{i1},k_{i2}....k_{il})^T$$

Note that (3-3-1.25) is of the form of (3-3-1.13), so that we have the following theorem resembling that of (3-3-1.4).

Theorem 3-3-1.5

The sufficient condition for the system (3-3-1.19) to be identifiable is that

$$rank\begin{bmatrix} \hat{d}_{i1} & 0 & 0 & \hat{n}_{i1} & 0 & 0 \\ 0 & \hat{d}_{i2} & 0 & 0 & \hat{n}_{i2} & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \hat{d}_{il} & 0 & 0 & \hat{n}_{il} \end{bmatrix}=l \quad \text{for } all\ z \text{ and } all\ i \tag{3-3-1.26}$$

where

$$\hat{d}_{ij}=z^n+d_{ij}(1)z^{n-1}+...+d_{ij}(n) \tag{3-3-1.27}$$

$$\hat{n}_{ij}=n_{ij}(1)z^{m-1}+...+n_{ij}(m) \tag{3-3-1.28}$$

Remarks; If $\hat{d}_{ij}$ and $\hat{n}_{ij}$ are coprime for all ij, then the condition (3-3-1.26) is satisfied.

3-3-2 Algorithms and The Rate of Convergence

In this section we will present identification algorithms with their convergence analysis for identifying partially known discrte time systems of the form (3-2.1) that is

$$\frac{y(t)}{u(t)}=\frac{n_0(z^{-1})+\sum_{i=1}^{l}\alpha_i n_i(z^{-1})}{d_0(z^{-1})-\sum_{j=1}^{k}\beta_j d_j(z^{-1})}=\frac{n(z^{-1})}{d(z^{-1})} \tag{3-3-2.1}$$

where α_i, β_j's are unknown parameters, d_j, n_i are known *nth* and *mth* order polynomials in the unit delay operator z^{-1},

$$d_0(z^{-1})=1+d_{01}z^{-1}+\ldots+d_{0n}z^{-n} \tag{3-3-2.2}$$

$$d_j(z^{-1})=d_{j1}z^{-1}+\ldots+d_{jn}z^{-n} \quad j=1,\ldots,k \tag{3-3-2.3}$$

$$n_i(z^{-1})=n_{i1}z^{-1}+\ldots+n_{im}z^{-m} \quad i=0,1,\ldots,l \tag{3-3-2.4}$$

The identification problem is to identify β_j, α_i from input-output measurements of the system. Rearranging equation (3-3-2.1), we get

$$(d_0(z^{-1}))y(t)-n_0(z^{-1})\bar{u}(t)$$

$$=\sum_{1}^{k}\beta_j d_j(z^{-1})y(t)+\sum_{1}^{l}\alpha_i n_i(z^{-1})u(t) \tag{3-3-2.5}$$

Define the following signal vectors

$$z_0(t)=d_0(z^{-1})y(t)-n_0(z^{-1})u(t) \tag{3-3-2.6}$$

$$W^T(t-1)=(d_1(z^{-1})y(t),\ldots,d_k(z^{-1})y(t),n_1(z^{-1})u(t),\ldots,n_l(z^{-1})u(t)) \tag{3-3-2.7}$$

$$\theta_1{}^T=(-d_{01},\ldots,-d_{0n},n_{01},\ldots,n_{0m}) \tag{3-3-2.8}$$

$$\theta^T=(\beta_1,\ldots,\beta_k,\alpha_1,\ldots,\alpha_l) \tag{3-3-2.9}$$

$$\psi^T(t-1)=(y(t-1),\ldots,y(t-n),u(t-1),\ldots,u(t-m)) \tag{3-3-2.10}$$

Then it follows that

$$z_0(t)=\theta^T W(t-1)=\theta^T D^T \psi(t) \qquad (3\text{-}3\text{-}2.11)$$

Let $\hat{\theta}(t)$ denote the parameter estimate at time t. Then since $z_0(t)$ and $W(t-1)$ are obtainable from the input and output , we may construct the equation error

$$e(t)=\hat{\theta}(t-1)W(t-1)-z_0(t) \qquad (3\text{-}3\text{-}2.12)$$

with $\phi(t)=\hat{\theta}(t)-\theta_0$ denoting the parameter error, we see that

$$e(t)=\phi^T(t-1)W(t-1) \qquad (3\text{-}3\text{-}2.13)$$

Equation (3-3-2.13) is linear in the parameter error, so that any one of a number of standard techniques for parameter update (see [25]) may be used. We only discuss two of them, which will be used in the next section.

A. The Projection Type Algorithm

The update law

$$\hat{\theta}(t)=\hat{\theta}(t-1)+\frac{W(t-1)}{1+W^T(t-1)W(t-1)}(z_0(t)-\hat{\theta}(t-1)W(t-1)) \qquad (3\text{-}3\text{-}2.14)$$

is referred to as projection type law. It is well known [25] that this algorithm has following properties.

$$||\hat{\theta}(t)-\theta_0|| \le ||\hat{\theta}(t-1)-\theta_0|| \le ||\hat{\theta}(0)-\theta_0|| \quad \text{for } any\ t \ge 1 \qquad (3\text{-}3\text{-}2.15)$$

$$\lim_{t\to\infty} \frac{(\phi^T(t-1)W(t-1))^2}{1+W^T(t-1)W(t-1)}=0 \qquad (3\text{-}3\text{-}2.16)$$

As mentioned in [25] that nothing has been said about $\hat{\theta}(t)$ necessarily converging to θ. In fact, $\hat{\theta}(t)$ may not converge at all. However the properties have been derived under very week assumption and in many cases this is not a significant limitation and the performance of the algorithm, in a particular application, can be inferred from theses properties. Further it is well known [ ] that the projection type algorithm (3-3-2.14) has exponential convergence when $W(t)$ is persistently exciting.

B. The Least Squares Type Algorithm

The least squares type algorithm is given by

$$\hat{\theta}(t)=\hat{\theta}(t-1)+\frac{P(t-2)W(t-1)}{1+W^T(t-1)P(t-2)W(t-1)}(z_0(t)-W^T(t-1)\hat{\theta}(t-1)) \qquad (3\text{-}3\text{-}2.17a)$$

$$P(t-1)=P(t-2)-\frac{P(t-2)W(t-1)W^T(t-1)P(t-2)}{1+W^T(t-1)P(t-2)W(t-1)} \qquad P(-1)=\alpha I>0 \qquad (3\text{-}3\text{-}2.17b)$$

The least square algorithm generally has very fast initial convergence rate which is much faster than the projection type algorithm. But the rate reduces dramatically after a few iteration. One variant of this algorithm is that covariance $P(t)$ is reset at various times. This is called the least squares type algorithm with covariance resetting.

C. The Least Square Type Algorithm with Covariance Resetting

The least squares type algorithm with covariance resetting is given by

$$\hat{\theta}(t)=\hat{\theta}(t-1)+\frac{P(t-2)W(t-1)}{1+W^T(t-1)P(t-2)W(t-1)}(z_0(t)-W^T(t-1)\hat{\theta}(t-1)) \qquad (3\text{-}3\text{-}2.18a)$$

$$P(t-1)=\begin{cases} P(t-2)-\dfrac{P(t-2)W(t-1)W^T(t-1)P(t-2)}{1+W^T(t-1)P(t-2)W(t-1)} & \text{if } t\neq 0,t_1,t_2,\ldots \\ \alpha I & \text{if } t=0,t_1,t_2,\ldots \end{cases} \qquad (3\text{-}3\text{-}2.18b)$$

in (3-3-2.18b) covariance resetting occurs at $\{0,t_1,t_2,\ldots\}$.

It is pointed out in [30], by a scalar example, that the convergence rate of the least squares type algorithm (without covariance resetting) is 1/t, if W(t) is persistently exciting. We will show, in the following theorem, that this is true for general cases.

Theorem 3-3-2.1

For the algorithm (3-3-2.17) subject to the condition of persistency of excitation, i.e.

$$\alpha_1' I \leq \sum_{i=\tau+1}^{\tau+T} W(i)W^T(i) \leq \alpha_2' I \qquad \text{for all } \tau\in Z_+ \qquad (3\text{-}3\text{-}2.19)$$

for some $T\in Z_+$ and some $\alpha_1', \alpha_2' > 0$. Then it follows that

$$\frac{\beta_1}{t^2}\leq\frac{||\phi(t+1)||^2}{||\phi(0)||^2}\leq\frac{\beta_2}{t^2} \qquad (3\text{-}3\text{-}2.20)$$

for some $\beta_1, \beta_2 > 0$, and for all $t \geq T$. Where $\phi(t)=\hat{\theta}(t)-\theta$ denotes the parameter error.

Proof: For any $t \geq T$, there exists a $k \in Z_+$ such that $kT \leq t \leq (k+1)T$ and by assumption (3-3-2.19), it follows that

$$\frac{1}{t}\sum_{1}^{t} W(i)W^T(i) \geq \frac{1}{t}k\alpha_1' I \geq \frac{k}{k+1}\frac{\alpha_1'}{T} I \geq \frac{1}{2}\alpha_1 I \tag{3-3-2.21}$$

$$\frac{1}{t}\sum_{1}^{t} W(i)W^T(i) \leq \frac{1}{t}(k+1)\alpha_2' I \leq \frac{k+1}{k}\frac{\alpha_2'}{T} I \leq 2\alpha_2 I \tag{3-3-2.22}$$

with

$$\alpha_1=\alpha_1'/T, \quad \alpha_2=\alpha_2'/T \tag{3-3-2.23}$$

This implies

$$\frac{1}{2}\alpha_1 tI \leq \sum_{1}^{t} W(i)W^T(i) \leq 2\alpha_2 tI \qquad \text{for } all \ t \geq T \tag{3-3-2.24}$$

From equation (3-3-2.17b), we have

$$P^{-1}(t)=P^{-1}(-1)+\sum_{i=0}^{t} W(i)W^T(i)$$

and consequently

$$\frac{1}{2}t\alpha_1 I \leq P^{-1}(t) \leq P^{-1}(-1)+2\alpha_2 tI \tag{3-3-2.25}$$

Since $P^{-1}(-1)=\frac{1}{\alpha}I$. Then, it follows immediately

$$\frac{1}{2}t\alpha_1 I \leq P^{-1}(t) \leq (\frac{1}{\alpha T}+2\alpha_2)tI \qquad t \geq T \tag{3-3-2.26}$$

If we define

$$S(t)=\frac{1}{t}P^{-1}(t) \tag{3-3-2.27}$$

It is obvious that S(t) is uniformly ($t \geq T$) bounded both below and above by $\alpha_4=\frac{1}{2}\alpha_1$, $\alpha_3=(\frac{1}{\alpha T}+2\alpha_2)$ respectively. On the other hand, the parameter error may be written as

$$\phi(t) = P(t-1)P^{-1}(t-2)\phi(t-1)$$

by successively substituting, we have

$$\phi(t+1)=\frac{1}{t}S^{-1}(t)P^{-1}(-1)\phi(0)=\frac{1}{\alpha t}S^{-1}(t)\phi(0) \tag{3-3-2.28}$$

where $P^{-1}(-1)=\frac{1}{\alpha}I$ is the initial value of P^{-1}. Since $S^{-1}(t)$ is bounded above and below, this implies that the error $\phi(t)\rightarrow 0$ at the rate of $1/t$. More precisely, let M(t) be an orthonormal matrix such that

$$M^T(t)S(t)M(t)=\Lambda(t) \tag{3-3-2.29}$$

where $\Lambda(t)$ is a diagonal matrix. Then it follows from (3-3-2.28) that

$$||\phi(t+1)||^2=\frac{1}{\alpha^2 t^2}\phi(0)^T M(t)\Lambda^{-2}(t)M^T(t)\phi(0) \tag{3-3-2.30}$$

Note that

$$\frac{1}{\alpha_3}\le\sigma_{\min}(\Lambda^{-1}(t))\le\sigma_{\max}(\Lambda^{-1}(t))\le\frac{1}{\alpha_4}$$

and M(t) is orthonormal, the conclusion follows

$$\frac{\beta_1}{t^2}\le\frac{||\phi(t+1)||^2}{||\phi(0)||^2}\le\frac{\beta_2}{t^2} \tag{3-3-2.31}$$

with

$$\beta_1=\frac{1}{\alpha^2\alpha_3{}^2}\ ,\ \beta_2=\frac{1}{\alpha^2\alpha_4{}^2} \tag{3-3-2.32}$$

This completes the proof.

The theorem above shows that the convergence rate of the least square algorithm is 1/t after a few iterations (T iterations), since the matrix P gets smaller and smaller. However, for the algorithm with covariance matrix resetting (3-3-2.18), the rate of convergence is exponential as stated in the following theorem.

Theorem 3-3-2.2

Consider the least squares type algorithm with covariance resetting (3-3-2.18). Suppose that $W(t)$ is persistently exciting as in (3-3-2.19) and the resetting time is defined by $t_i = t_{i-1} + kT - 1$ for some fixed k. Then it follows that

$$\left(\frac{1}{(1+k\alpha\alpha_2')^2}\right)^i \leq \frac{||\phi(t_i)||^2}{||\phi(0)||^2} \leq \left(\frac{1}{(1+k\alpha\alpha_1')^2}\right)^i \tag{3-3-2.33}$$

Proof: Let $P^{-1}(t_i^- - 1)$ and $P^{-1}(t_i^+ - 1)$ denote the covariance matrix right before and after resetting. Then similar to the proof of theorem 3-3-2.1, we have

$$P^{-1}(t_i^- - 1) = P^{-1}(t_{i-1}^+ - 1) + \sum_{j=t_{i-1}}^{t_i - 1} W(j) W^T(j)$$

$$= \frac{1}{\alpha} I + \sum_{j=t_{i-1}}^{t_i - 1} W(j) W^T(j) \tag{3-3-2.34}$$

This implies that

$$\left(\frac{1}{\alpha} + k\alpha_1'\right) I \leq P^{-1}(t_i^- - 1) \leq \left(\frac{1}{\alpha} + k\alpha_2'\right) I \tag{3-3-2.35}$$

and notice that

$$\phi(t_i) = P(t_i^- - 1) P^{-1}(t_i - 2)\phi(t_i - 1) = P(t_i^- - 1) P^{-1}(t_{i-1}^+ - 1)\phi(t_{i-1})$$

$$= \frac{1}{\alpha} P(t_i^- - 1)\phi(t_{i-1}) \tag{3-3-2.36}$$

Then, the parameter error is given by

$$\phi(t_i) = \frac{1}{\alpha} P(t_i^- - 1) \ldots\ldots \frac{1}{\alpha} P(t_1^- - 1)\phi(0) \tag{3-3-2.37}$$

Combing equations (3-3-2.35) and (3-3-2.37), the conclusion follows

$$\left(\frac{1}{(1+k\alpha\alpha_2')^2}\right)^i \leq \frac{||\phi(t_i)||^2}{||\phi(0)||^2} \leq \left(\frac{1}{(1+k\alpha\alpha_1')^2}\right)^i \tag{3-3-2.38}$$

This completes the proof.

3-4 Continuous Time System Identification

In this section, we will deal with continuous time 'partially known' systems. Unlike the discrete time case, we will use rational functions instead of polynominals for our analysis. The framework of rerpresenting transfer functions as the ratio of proper, stable rational functions, proposed for example by Desoer et al [17] and Francis & Vidyasagar [20] has proven useful in H^∞ approach to linear control systems design and it has payoff in our context as well in studying the effect of unmodeled dynamics on the identification scheme.

We consider the problem of identifying a class of 'partially known' systems described by equation (3-2.1) i.e.

$$\hat{G}(s)=\frac{g_0(s)+\sum_{i=1}^{m}\beta_i g_i(s)}{f_0(s)-\sum_{j=1}^{n}\alpha_j f_j(s)} \tag{3-4.1}$$

where the g_i's and f_j's are known, proper, stable rational functions in s and the β_i , α_j's are unknown, real parameters.

The identification problem is to identify β_i , α_j from input-output measurements of the system.

Remark:

All the models described in section 3-2 can be parametrized in the form of (3-4.1) e.g Classical transfer function identification, i.e. identification of a stable plant of the form

$$\hat{G}(s)=n(s)/d(s)=\frac{\beta_1 s^{m-1}+\ldots+\beta_m}{s^n+\alpha_1 s^{n-1}+\ldots+\alpha_n} \tag{3-4.2}$$

with $m \le n$ and α_i and β_j unknown, can be stated in terms of the set up of (3-4.1) by choosing

$$g_0(s)=0,\quad f_0(s)=\frac{s^n}{(s+\alpha)^n}$$

$$g_i(s)=\frac{s^{m-i}}{(s+\alpha)^n} \quad i=1,..m$$

and

$$f_j(s)=\frac{s^{n-j}}{(s+\alpha)^n} \quad j=1,...n$$

with $\alpha>0$ a positive, real number. Also if m is not known, we may set it equal to n.

let y(s),u(s) denote the input and output to the plant of equation (3-4.1). (The initial conditions of the plant represent exponentially decaying terms which do not change any of the following discussions, as is well understood in the literature.) Then, after some rearrangements, we get

$$f_0 y-g_0 u=\sum_{j=1}^{n}\alpha_j f_j\, y+\sum_{i=1}^{m}\beta_i\, g_i\, u \tag{3-4.3}$$

Defining

$$z_0(s)=f_0(s)y(s)-g_0(s)u(s) \tag{3-4.4a}$$

$$h_j(s)=f_j(s)\hat{G}(s) \quad j=1,...n \tag{3-4.4b}$$

$$h_{n+i}(s)=g_i(s) \quad i=1,...m \tag{3-4.4c}$$

and the unknown parameter vector θ by

$$\theta^T=(\alpha_1,...,\alpha_n,\beta_1,\ldots,\beta_m)$$

we get

$$z_0(s)=\theta^T\begin{bmatrix} h_1(s) \\ \vdots \\ h_{n+m}(s)\end{bmatrix}u(s) \tag{3-4.5}$$

The vector of signal $(h_1(s),...,h_{n+m}(s))^T$ is denoted $z(s)$ and its Laplace inverse $z(t)$ so that in time domain (3-4.5) reads (again modulo decaying initial condition terms.)

$$z_0(t)=\theta^T(z(t)*u(t)) \tag{3-4.6}$$

(* stands for convolution.) By way of notation, we refer to z(t)*u(t) as W(t).

From the form of equation (3-4.6), it is easy to see how an estimator and equation can be derived. Let $\hat{\theta}(t)$ denotes the parameter estimate at time t. Then since W(t) is a vector of signals obtainable from the input and output by proper stable filtering, as seen from (3-4.4), we can construct the error

$$e(t)=\hat{\theta}(t)^T W(t)-z_0(t) \tag{3-4.7}$$

Using (3-4.7) and with $\phi(t)=\hat{\theta}(t)-\theta$ denoting the parameter error, we see that

$$e(t)=\phi^T(t)W(t) \tag{3-4.8}$$

Similar to discrete time case, we have

A. The least Squares Type Algorithm

The parameter update law is of the form (with $P(t)\in R^{(n+m)\times(n+m)}$)

$$\dot{\hat{\theta}}(t)=-P^{-1}(t)W(t)e(t) \tag{3-4.9}$$

$$\dot{P}(t)=W(t)W(t)^T \quad P(0)=\alpha I>0 \tag{3-4.10}$$

It is well known (see Goodwin) that if W is persistently exciting i.e. there exists $\alpha_1, \delta>0$ such that

$$\int_s^{s+\delta} WW^T\, dt \geq \alpha_1 I \quad \text{for } any\ s \in R_+ \tag{3-4.11}$$

then $\phi \to 0$ as $t \to \infty$. Of course, since W(t) is bounded we in fact have

$$\alpha_2 I \geq \int_s^{s+\delta} WW^T\, dt \geq \alpha_1 I \quad \text{for } any\ s \in R_+ \tag{3-4.12}$$

The result of Chapter 2 may be used to give frequency domain condition on u(t) to guarantee (3-4.12). First, we need the following assumption on the identifiability condition.

Assumption on the Identifiability Condition (AIC)

Consider the system (3-4.1), assume that for every choice of distinct (n+m) frequencies $\nu_1, \ldots, \nu_{n+m}$ the vectors $\hat{W}(j\nu_i)\in C^{n+m} (i=1,\ldots,n+m)$ are linearly independent

Remarks:

(1) From (3-4.5), it follows that if an input having (n+m) spectral lines were applied to the system, we would get

$$[z_0(j\nu_1),\ldots,z_0(j\nu_{n+m})]$$

$$=\theta_0^T[z(j\nu_1),\ldots,z(j\nu_{n+m})]diag(u(j\nu_1),\ldots,u(j\nu_{n+m})) \tag{3-4.13}$$

In turn, AIC implies that (3-4.13) has a unique solution for θ.

(2) It is difficult to give a more concrete characterization of identifiability since the component of $\hat{W}(s)$ are proper, stable rational functions of different orders. An exception is the case of classical identification discussed in remark 2 (equation 3-4.2) in which case it has been shown in Boyd [12] that the identifiability condition holds if n(s) and d(s) are coprime polynominals. However, by similar arguments as in chapter 2, we may conclude that AIC holds for almost any choice of $(\nu_1,\ldots,\nu_{n+m})$.

Using the assumption on the identifiability condition , we state the following fact easily derived from [ ];

Under the identifiability assumption AIC, W is persistently exciting, i.e. it satisfies (3-4.12) if the spectral measure of u is not concentrated on less than n+m points.

Thus, if there are at least as many frequencies in the input as there are unknown parameters, the parameter errors converge to zero. Of course, the least squares type algorithm (3-4.9), (3-4.10) shows rapid initial convergence with asymptotically slow adaptation (as P(t) gets large). Some form of resetting of P(t) or forgetting is introduced (as in Goodwin and Sin pg.62), for example

$$\dot{P}(t)=-\lambda P(t)+W(t)W(t)^T \qquad P(0)=\alpha I>0 \tag{3-4.14}$$

It is then easy to show that the convergence of the parameter error is exponential. It is important to note that forgetting is not used when W is not persistently exciting to keep P from going singular.

B. Projection Type Algorithm

The update laws

$$\dot{\theta}(t)=-W(t)e(t) \tag{3-4.15}$$

or

$$\dot{\hat{\theta}}(t) = \frac{W(t)e(t)}{1+W(t)^T W(t)} \tag{3-4.16}$$

are referred to as projection type algorithms. They also yields exponential convergence when the input is sufficiently rich in the sense discussed above and the assumption AIC holds.

To illustrate the methods of this section, consider the following example in Fig. 3-4.1

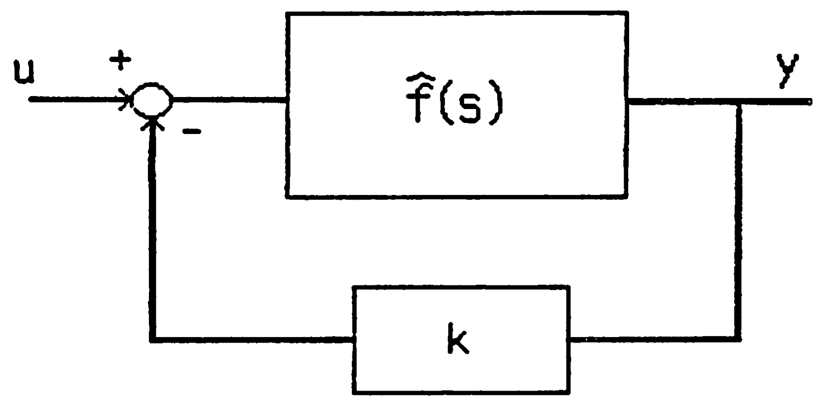

Fig. 3-4.1

In Fig. 3-4.1, f(s) is known (assumed to be $\frac{2s+2}{3s+5}$ for the simulations of Figures 3-4.2 and 3-4.3). The form of the closed loop transfer function is

$$\frac{f(s)}{1+kf(s)} = \frac{as+b}{s+c}$$

With the parameter k to be estimated, Figure 3-4.2 shows the parameter error for the algorithm with the projection type update law (the true value of k=1) and input u(t)=5 (only one spectral line is needed for identification).

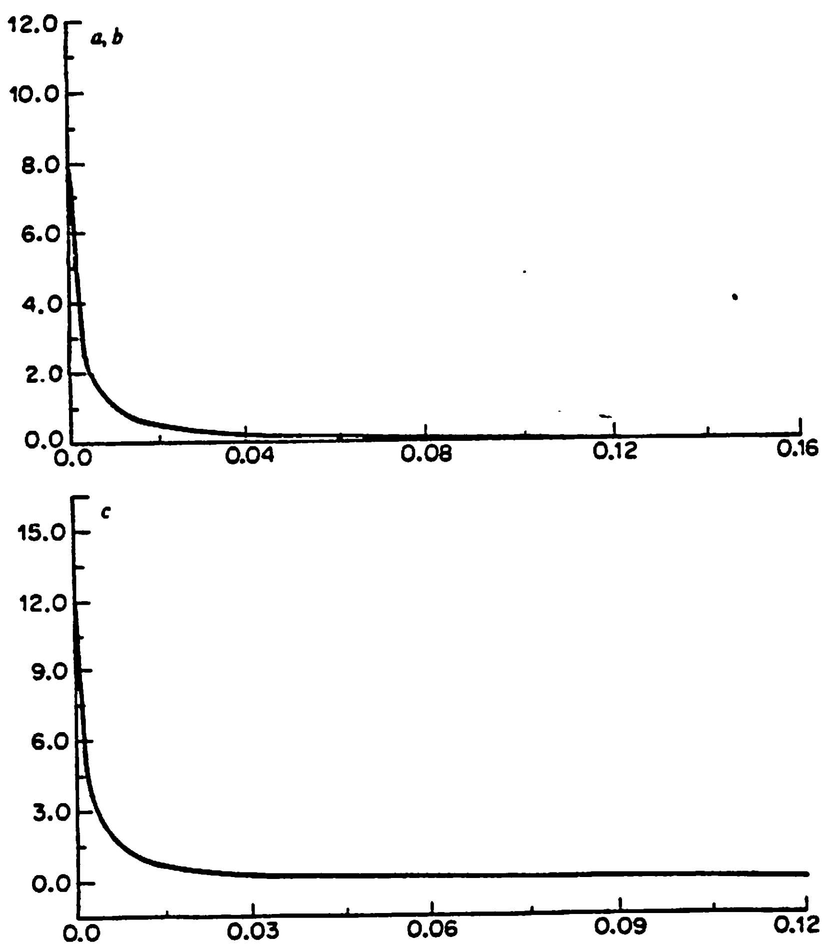

Fig. 3-4.2

Estimation errors of parameters a and b (top) and c (bottom) using prior information.

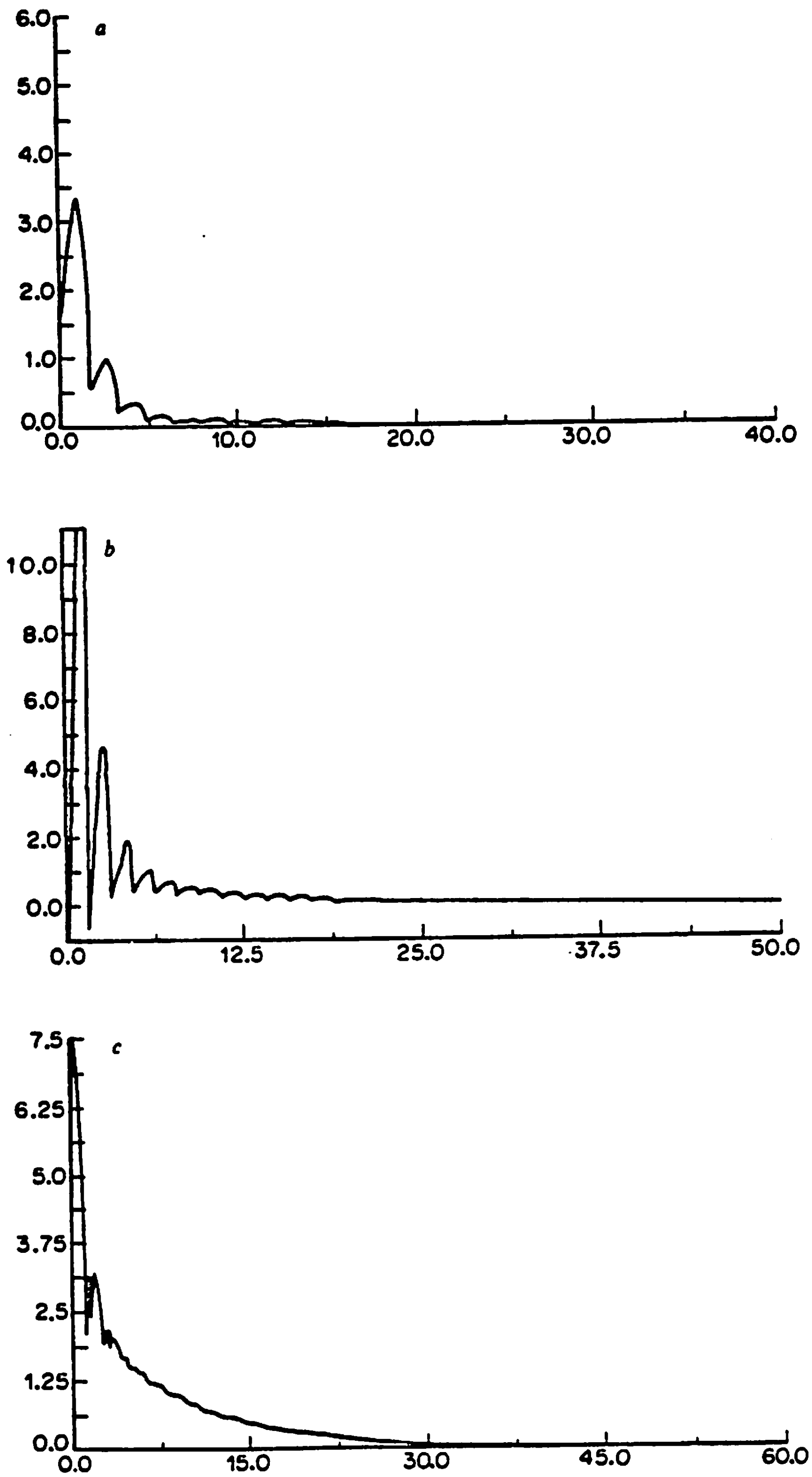

Fig. 3-4.3

Estimation errors of parameters *a* (top), *b* (middle) and *c* (bottom) without using prior information.

Identification of the closed loop plant without utilizing the structure of the system requires the estimation of three parameters a,b and c. Figure 3-4.3 shows the parameter errors for a, b and c using the input u(t)=3+4sin(4t). Note that the two inputs for figures 3-4.2 and 3-4.3 have the same energy. The input for Figure 3-4.3 is richer than that for Figure 3-4.2. However, the rate of convergence is much slower (by a factor of approximately 500) in Figure 3-4.3. In the following, we will see that the scheme using prior information also has a larger robustness margin.

Though the assumption AIC guarantees that the parameter errors converge to zero if and only if the support of the spectrum of input u has at least n+m points, it does not provide much insight into the connection between the spectral content of the input and the convergence rate. We will use averaging techniques to facilitate this analysis.

First consider the projection type algorithm (3-4.15) with slow update law (modeled by adaptation gain ε, a small positive number).

$$\dot{\phi}=-\varepsilon WW^T\phi \qquad (3\text{-}4.17)$$

Defining the averaged value of WW^T to be $R_W(0)$ (see [22]) given by

$$R_W(0)=\lim_{T\to\infty}\frac{1}{T}\int_s^{s+T} WW^T\,dt \qquad \text{for } any\ s\in \bar{R}_+ \qquad (3\text{-}4.18)$$

(provided it exists- this in turn is guaranteed by assuming W to be stationary, see chapter 2 for details.) We see that for ε small enough the dynamics of (3-4.17) (including rate of convergence up to the high order of ε) are approximated by

$$\dot{\phi}_{av}=-\varepsilon R_W(0)\phi_{av} \qquad (3\text{-}4.19)$$

Noting that $R_W(0)$ is the integral of the spectral measure of W, we may rewrite an expression for $R_W(0)$ in terms of the input spectrum and the function z(s) as

$$R_W(0)=\int z(j\nu)S_u(d\nu)z^*(j\nu) \qquad (3\text{-}4.20)$$

where $S_u(d\nu)$ stands for the spectral measure of u. Thus, the convergence rate of (3-4.19) is obtained to lie in an interval $[\underline{\sigma}(R_W(0)),\bar{\sigma}(R_W(0))]$, where $\underline{\sigma}$ ($\bar{\sigma}$) denotes minimum (maximum) eigenvalue. For optimum convergence, the spectrum of the input needs to be in the dominant part of $z(jv)z^*(jv)$. Of course, the expression (3-4.20) involves parameters of the unknown plant so that it is not easily approximated.

For the analysis of the slowed-down least squares algorithm, consider with W assumed to be persistently exciting

$$\dot{\phi}=-\varepsilon P^{-1}WW^T\phi \tag{3-4.21}$$

$$\dot{P}=\varepsilon(WW^T-\lambda P) \tag{3-4.22}$$

As before, we may approximate (3-4.21), (3-4.22) by the averaged system

$$\dot{\phi}_{av}=-\varepsilon P^{-1}{}_{av}R_W(0)\phi_{av} \tag{3-4.23}$$

$$\dot{P}_{av}=\varepsilon(R_W(0)-\lambda P_{av}) \tag{3-4.24}$$

Equation (3-4.24) may be explicitly integrated to give

$$P_{av}(t)=(P_{av}(0)-\frac{1}{\lambda}R_W(0))e^{-\varepsilon\lambda t}+\frac{1}{\alpha}R_W(0) \tag{3-4.25}$$

In turn, using this in (3-4.23) and noting that $P_{av}(t)$ converges exponentially to $\frac{1}{\alpha}R_W(0)$, we see that the tail behavior of (3-4.23) is

$$\dot{\phi}_{av}=-\varepsilon\lambda\phi_{av} \tag{3-4.26}$$

so that the tail convergence rate is a function of the forgetting factor λ alone in the 'covariance' equation (3-4.22) and not the input spectrum!

We now consider the effect of unmodeled dynamics on parameter identification. The set up of the above discussion used transfer functions of the form (3-4.1) with the f_i and g_j's known exactly. In practice, the f_i and g_j's will not be known exactly, but only approximately. In fact, the transfer functions used to approximate the f_i and g_j will generally be low order proper, stable rational functions(neglecting high frequency dynamics, and replacing near pole-zero cancellations by exact pole-zero cancellations). Thus, the identifier's model of the plant is of the form

$$\bar{G}(s)=\frac{\bar{g}_0+\sum_{i=1}^{m}\beta_i\,\bar{g}_i}{\bar{f}_0-\sum_{j=1}^{n}\alpha_j\,\bar{f}_j} \tag{3-4.27}$$

where $\bar{G}$ is a proper,stable transfer function and

$$|(\bar{g}_i-g_i)(jw)|<\varepsilon \quad \text{for } all\ w \quad i=0,...m \tag{3-4.28a}$$

$$|(\bar{f}_j-f_j)(jw)|<\varepsilon \quad \text{for } all\ w \quad j=0,...n \tag{3-4.28b}$$

We refer to $\bar{g}_i-g_i$ as Δg_i in the sequel, similarly for Δf_j. For example, g_i may be of the form

$$\bar{g}_i \frac{1}{v(s)} \frac{q(s)}{p(s)} \tag{3-4.29}$$

where $\frac{1}{v(s)}$ represents stable high frequency dynamics and $\frac{q(s)}{p(s)}$ represents near (stable) pole-zero cancellations.

The identifier uses the form (3-4.27) to derive the identifier for the true plant $\hat{G}(s)$ which is accurately described by (3-4.1). Consequently the transfer functions of (3-4.4) are replaced by

$$\bar{z}_0(s)=\bar{f}_0\hat{G}u-\bar{g}_0u \tag{3-4.30a}$$

$$\bar{h}_j(s)=\bar{f}_j\hat{G} \quad j=1,...,n \tag{3-4.30b}$$

$$\bar{h}_{n+i}(s)=\bar{g}_i \quad i=1,..m \tag{3-4.30c}$$

It is important to note that $\bar{z}_0$ does not satisfy an equation of the form (3-4.5) i.e. it is not true that

$$\bar{z}_0(s)=\theta^T \begin{bmatrix} \bar{h}_1(s) \\ \vdots \\ \bar{h}_{n+m}(s) \end{bmatrix} u(s)$$

Equation (3-4.5) is ,of course, still valid. The update law (least squares type) is now of the form (with $\bar{W}_i(t)=\bar{h}_i(t)*u(t), i=1,...n+m$)

$$\dot{\hat{\theta}}=-P^{-1}\bar{W}(t)(\hat{\theta}^T(t)\bar{W}(t)-\bar{z}_0(t)) \tag{3-4.31}$$

$$\dot{P}=\bar{W}\bar{W}^T-\lambda P \quad P(0)=\alpha I>0 \tag{3-4.32}$$

We need an expression for $\bar{z}_0$ in order to study this algorithm. For this purpose we note that

$$\bar{z}_0(s)=z_0(s)+(\bar{f}_0(s)-f_0(s))\hat{G}(s)u(s)-(\bar{g}_0(s)-g_0(s))u(s) \tag{3-4.33}$$

$$=z_0(s)+\Delta f_0\hat{G}(s)u(s)-\Delta g_0u(s)$$

Also, we have

$$W(s)=\bar{W}(s)-\begin{bmatrix}\Delta h_1\\ \vdots\\ \Delta h_{n+m}\end{bmatrix}u(s) \qquad (3\text{-}4.34)$$

Using (3-4.33) and (3-4.34) we see that equation (3-4.31) may be rewritten as

$$\dot{\hat{\theta}}=-P^{-1}\bar{W}\bar{W}^T(\hat{\theta}(t)-\theta)-P^{-1}\bar{W}(t)\delta(t) \qquad (3\text{-}4.35)$$

where $\delta(t)$ is the Laplace inverse of

$$\theta^T\begin{bmatrix}\Delta h_1(s)\\ \vdots\\ \Delta h_{n+m}(s)\end{bmatrix}u(s)-\Delta f_0(s)\hat{G}(s)u(s)+\Delta g_0(s)u(s)$$

With $\hat{\theta}(t)-\theta=\phi(t)$, the parameter error, the error dynamics are given by

$$\dot{\phi}=-P^{-1}\bar{W}\bar{W}^T\phi-P^{-1}\bar{W}\delta(t) \qquad (3\text{-}4.36)$$

$$\dot{P}=\bar{W}\bar{W}^T-\lambda P \qquad P(0)=\alpha I>0 \qquad (3\text{-}4.37)$$

The last term in equation (3-4.36) may be considered as a (state-dependent) driving term. If the undriven system is exponentially stable, then using the results of [53], the driven system is stable as well. In turn, the undriven system is exponentially stable if and only if $\bar{W}$ is persistently exciting i.e. (3-4.12) holds for $\bar{W}$. We will give conditions, using the following two lemmas, on the persistent excitation of $\bar{W}$ in the case when ε is small enough.

Lemma 3-4.1

Suppose that $W\in R^{n+m}$ is persistently exciting i.e.

$$\alpha_2\geq\int_s^{s+\delta}WW^T\,dt\geq\alpha_1$$

for some $\alpha_1,\alpha_2,\delta>0$ and for all s>0. Then, $W+\Delta W$ is also persistently exciting provided that

$$\|\Delta W(\cdot)\|<(\alpha_1/\delta)^{1/2} \qquad (3\text{-}4.38)$$

Proof: $W+\Delta W$ is persistently exciting if for any $x \in R^{n+m}$ of unit norm

$$\alpha_2' \geq \int_s^{s+\delta} |x^T(W+\Delta W)|^2 \, dt \geq \alpha_1' \tag{3-4.39}$$

The upper bound on the integral in (3-4.39) is automatic for some α_2' simply because ΔW is bounded. For the lower bound, we use the Minkowski inequality to get

$$\left(\int_s^{s+\delta} |x^T(W+\Delta W)|^2 \, dt\right)^{1/2} \geq \left(\int_s^{s+\delta} |x^T W|^2 \, dt\right)^{1/2} - \left(\int_s^{s+\delta} |x^T \Delta W|^2 \, dt\right)^{1/2}$$

$$\geq \alpha_1^{1/2} - \left(\int_s^{s+\delta} |\Delta W|^2 \, dt\right)^{1/2}$$

$$\geq \alpha_1^{1/2} - \delta^{1/2} sup |\Delta W(\cdot)| \tag{3-4.40}$$

The conclusion follows from (3-4.38).

To establish the norm bounded on error, we need the following lemma due to Doyle-Gohberg [18].

Lemma 3-4.2

If $\hat{G}(s)$ is a proper, n-th order stable rational function with Laplace inverse g(t), then

$$\int_0^{\infty} |g(t)| \, dt \leq 2\, n \; sup |\hat{G}(jw)| \tag{3-4.41}$$

Remark:

Lemmas 3-4.1 and 3-4.2 are to be interpreted as follows:

(1) Let $W(t)$ and $\Delta W(t)$ be the Laplace inverse of $(f_1\hat{G}, \ldots, f_n\hat{G}, g_1, \ldots, g_m)^T u(s)$ and $(\Delta f_1\hat{G}, \ldots, \Delta f_n\hat{G}, \Delta g_1, \ldots, \Delta g_m)^T u(s)$ respectively. From (3-4.33) it follows that

$$\tilde{W} = W + \Delta W \tag{3-4.42}$$

If we assume that the true system in (3-4.1) satisfies the assumption AIC, then sufficient richness of the input u (in the sense of before) guarantees that $\tilde{W}$ is persistently exciting, provided that ε in equation (3-4.28) is small enough.

(2) In practice, f_i and g_j are unknown. We may assume that the nominal plant $\bar{G}$ satisfies the identifiability condition. In such a case, equation (3-4.42) still holds with $W(t)$ and $\Delta W(t)$ given by Laplace inverse of $(\bar{f}_1\bar{G},\ldots,\bar{f}_n\bar{G},\bar{g}_1,\ldots,\bar{g}_n)^T u(s)$ and $(\bar{f}_1(\hat{G}-\bar{G}),\ldots,\bar{f}_n(\hat{G}-\bar{G}),0,\ldots,0)^T u(s)$ respectively. Then, we get same result as in remark (1) above.

(3) The classical identification can be thought of the special case of that in remark (2) as follows

$$\hat{G}(s)=\frac{\beta_1 s^{m-1}+\ldots+\beta_m}{s^n+\alpha_1 s^{n-1}+\ldots+\alpha_n}\cdot\frac{1}{v(s)}\,\frac{q(s)}{p(s)}$$

$$=\bar{G}(s)\frac{1}{v(s)}\,\frac{q(s)}{p(s)} \tag{3-4.43}$$

As in (3-4.29), $1/v(s)$ represents stable high frequency dynamics and $q(s)/p(s)$ represents near stable pole-zero cancellations. Then

$$g_i(s)=\frac{s^{m-i}}{(s+\alpha)^n}\,\frac{1}{v(s)}\,\frac{q(s)}{p(s)} \qquad i=1,\ldots,m$$

$$f_0=\frac{s^n}{(s+\alpha)^n}\,,\; f_j(s)=\frac{s^{n-j}}{(s+\alpha)^n} \qquad j=1,\ldots,n$$

For the identifier, both $v(s)$ and $q(s)/p(s)$ are neglected and we have

$$\bar{g}_i(s)=\frac{s^{m-i}}{(s+\alpha)^n} \qquad i=1,\ldots m$$

$$\bar{f}_j(s)=f_j(s) \qquad j=0,1,\ldots n$$

$|\Delta g_i|<\varepsilon$ provided that cancellations are almost perfect and unmodeled dynamics occur at high enough frequencies.

From the form of $\delta(t)$ in (3-4.35) and lemma 3-4.1, it follows that there exists a $K(\hat{m},G)$ depending only on $sup\,|\hat{G}(jw)|$ and m:=maximum order of Δf_i, Δg_j such that

$$sup\,|\delta(\cdot)|\le\varepsilon K(m,\hat{G})sup\,|u(\cdot)| \tag{3-4.44}$$

Under the condition that $\bar{W}$ is persistently exciting, it follows that the parameter errors in (3-4.36) converge to a ball with radius of order ε(see [53]).

To end this section, let us consider the same example discussed before with $f(s)=\frac{2s+2}{s+3}\,\frac{s+5}{s+5.5}$ and $\bar{f}(s)=\frac{2s+2}{s+3}$. The true closed loop transfer function is

$$\hat{G}(s)=\frac{s+5}{s+5.1}\,\frac{0.667s+0.667}{s+1.73}$$

With the parameter k to be estimated, Figure 3-4.4 shows the parameter errors (for the projection type algorithm and input u(t)=5 same as in the no unmodeled dynamics case). It takes about 1 second to converge and the resulting closed loop transfer function is

$$\hat{P}(s)=\frac{0.605s+0.605}{s+1.605}$$

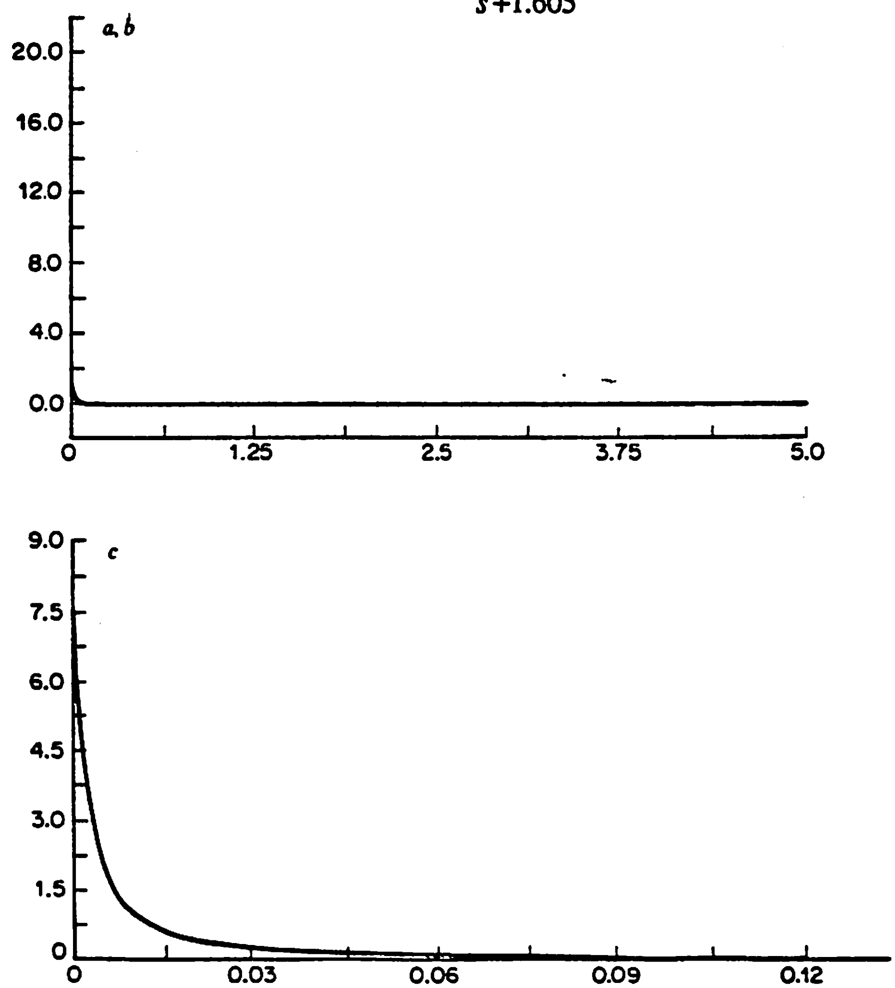

Fig. 3-4.4

Estimation errors of parameters a and b (top) and c (bottom) using prior information in the presence of unmodeled dynamics.

For the identification of the closed loop transfer function without utilizing the structure of the system, we have used input u(t)=3+4sin(4t). After 5000 seconds of simulation, the system does not converge. Figure 3-4.5 shows the estimation error of the parameter c.

The Bode plots of $\hat{G}(s)$ and $\hat{P}(s)$ are compared in Fig.3-4.6

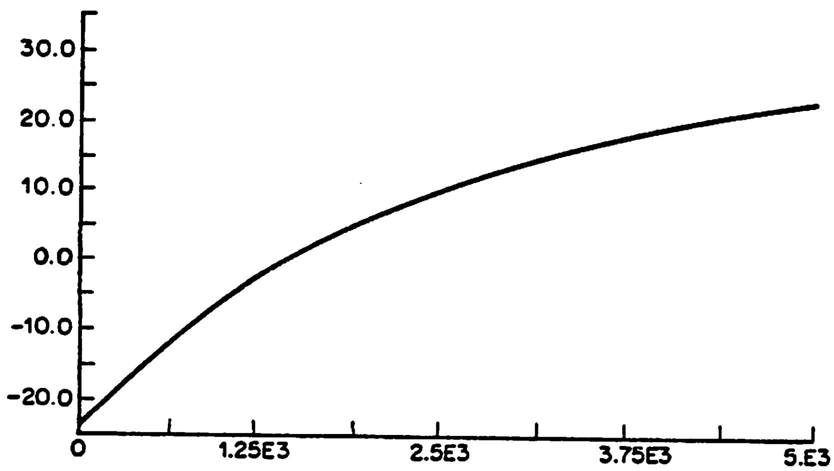

Fig. 3-4.5

Estimation error of parameter c without using prior information in the presence of unmodeled dynamics.

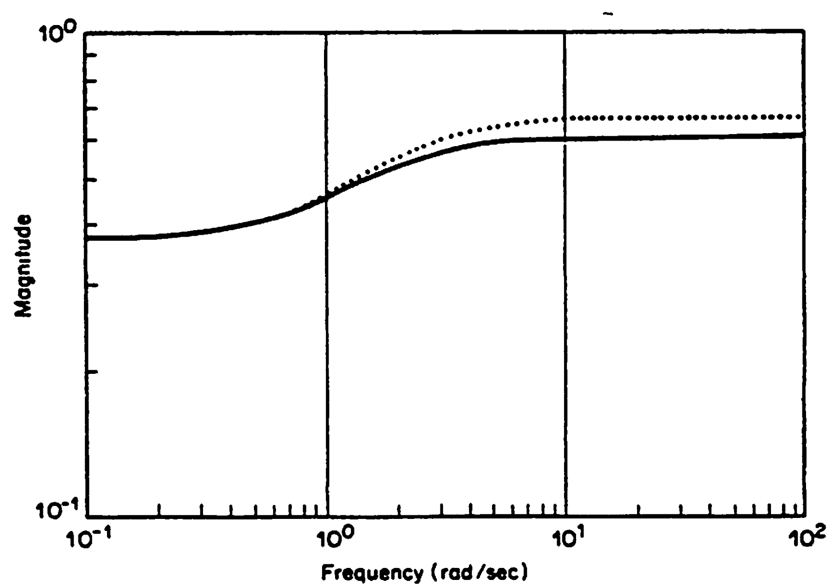

Fig. 3-4.6

The Bode plots of $\hat{G}(s)$ (· · · ·) and $\hat{P}(s)$ (——).

3-5 Adaptive Control of Discrete Time 'Partially Known' Systems

Consider the system of the form (3-3-2.1). An adaptive control law is to be designed to stabilize this system and to cause the output y(t) to track a given reference sequence $y^*(t)$ i.e. we require y(t) and u(t) to be bounded and

$$\lim_{t\to\infty}(y(t)-y^*(t))=0$$

The following assumptions will be made about the system (3-3-2.1)

(1) $n_{01}+\alpha_1 n_{11}+\ldots+\alpha_{l1}n_{l1}\neq 0$ This implies that the pure delay in the transfer function (3-3-2.1) is known and equal to 1. This is for simplicity alone in our analysis, the extension to the case where pure delay is greater than 1 (but known) follows readily.

(2) $n(z^{-1})$has all zeros strictly inside the closed unit disk i.e. the system is inverse stable.

(3) $y^*(t)$ is known a priori and bounded.

Control Algorithm Using Projection Type Identification Law

From equation (3-3-2.5),(3-3-2.8) and (3-3-2.11), we have

$$y(t+1)=\theta_1{}^T\psi(t)+\theta^T W(t)$$

$$=\theta_1{}^T\psi(t)+\theta^T D^T\psi(t) \tag{3-5.1}$$

We choose the projection type estimation law (3-4.14) and a control law specified implicitly by

$$y^*(t+1)=\theta_1{}^T\psi(t)+\hat{\theta}(t)^T D^T\psi(t) \tag{3-5.2}$$

(A minor modification is necessary to ensure that the coefficient of u(t) in (3-5.2) is nonzero. This can be achieved in the same way as in [ ] and does not affect the current analysis.) Then, we have

Theorem 3-5.1 (Convergence Theorem)

Subject to assumptions 1), 2) and 3), consider the control law (3-5.2), together with the projection type estimation law (3-3-2.14), applied to the system (3-3-2.1). Then, y(t) and u(t) are bounded and

$$\lim_{t\to\infty}(y(t)-y^*(t))=0 \qquad (3\text{-}5.3)$$

Proof: Define the output error by

$$e_y(t)=y(t)-y^*(t) \qquad (3\text{-}5.4)$$

It follows from (3-5.1) and (3-5.2) that

$$e_y(t)=-\phi^T(t-1)W(t-1) \qquad (3\text{-}5.5)$$

Now using equation (3-3-2.16), we have

$$\lim_{t\to\infty}\frac{e^2{}_y(t)}{1+W^T(t-1)W(t-1)}=0 \qquad (3\text{-}5.6)$$

Note that

$$\frac{e^2{}_y(t)}{1+W^T(t-1)W(t-1)}\geq\frac{e^2{}_y(t)}{1+\sigma_{\max}(DD^T)\psi^T(t-1)\psi(t-1)} \qquad (3\text{-}5.7)$$

By assumptions 2) and 3), we have as in [ ] that

$$||\psi(t-1)||\leq c_1+c_2\max_{1\leq\tau\leq t}|e_y(\tau)| \qquad (3\text{-}5.8)$$

for some $0\leq c_1<\infty$, $0<c_2<\infty$. The conclusion now follows from equation (3-5.7) and (3-5.8) using the key technical lemma in [24] and by noting that boundedness of $||\psi(\cdot)||$ ensures boundedness of y(t) and u(t).

Control Algorithm Using Least Squares Type Identification Law (with Covariance Resetting)

If the least squares type estimation law (3-3-2.18) is used , then we get the same result.

Theorem 3-5.2

Subject to assumptions 1), 2) and 3), consider the control law (3-5.2), together with the least squares type estimation law (3-3-2.18), applied to the system (3-3-2.1). Then y(t) and u(t) are bounded and

$$\lim_{t\to\infty}(y(t)-y^*(t))=0$$

Proof: The proof proceeds by an argument similar to that in [ ]. Define

$$e_y(t)=y(t)-y^*(t)$$

Then from [ ], we have

$$\lim_{t\to\infty}\frac{e^2{}_y(t)}{1+\alpha W^T(t-1)W(t-1)}=0 \tag{3-5.10}$$

The remainder of the proof is same as that of theorem 3-5.1.

We have shown the global stability of two adaptive control algorithms. Note that nothing has been said about the convergence rate of the output and the parameter convergence. However, if $n(z^{-1})$ and $d(z^{-1})$ in (3-3-2.1) are coprime and the D matrix in (3-3-2.4) has full column rank, then the persistency of excitation of W(t) follows from the sufficient richness of input u(t) (i.e. u(t) has sufficient spectral content, see [9]). This implies that the control algorithm, with either projection type or least squares type with covariance resetting parameter update, has exponential convergence rate both for the output error and parameter error.

To illustrate the methods of last section, consider the following example in Fig 3-5.1

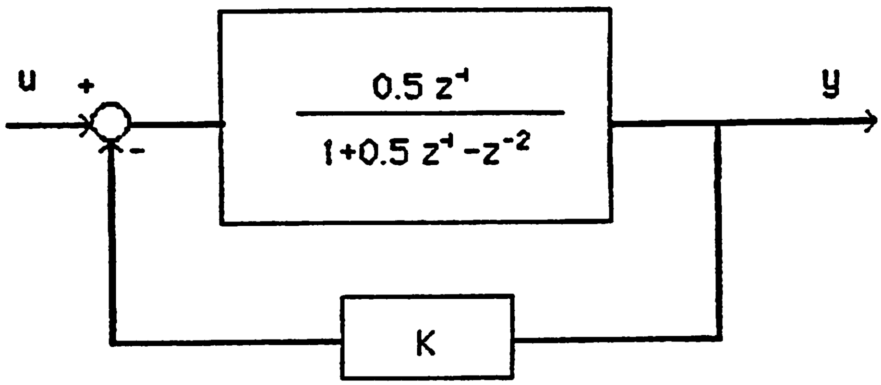

Fig. 3-5.1

where k is unknown. The closed loop transfer function is

$$\frac{0.5z^{-1}}{(1+0.5z^{-1}-z^{-2})+k\,0.5z^{-1}}$$

Fig.3-5.2 shows the plant output under the adaptive control algorithm of (3-5.2) with projection type update law and the plant output under the adaptive control algorithm without using prior information respectively. (for the simulation, k=1, and $y^*(t)\equiv 1$). The algorithm using prior information has faster convergence rate and better transient performance.

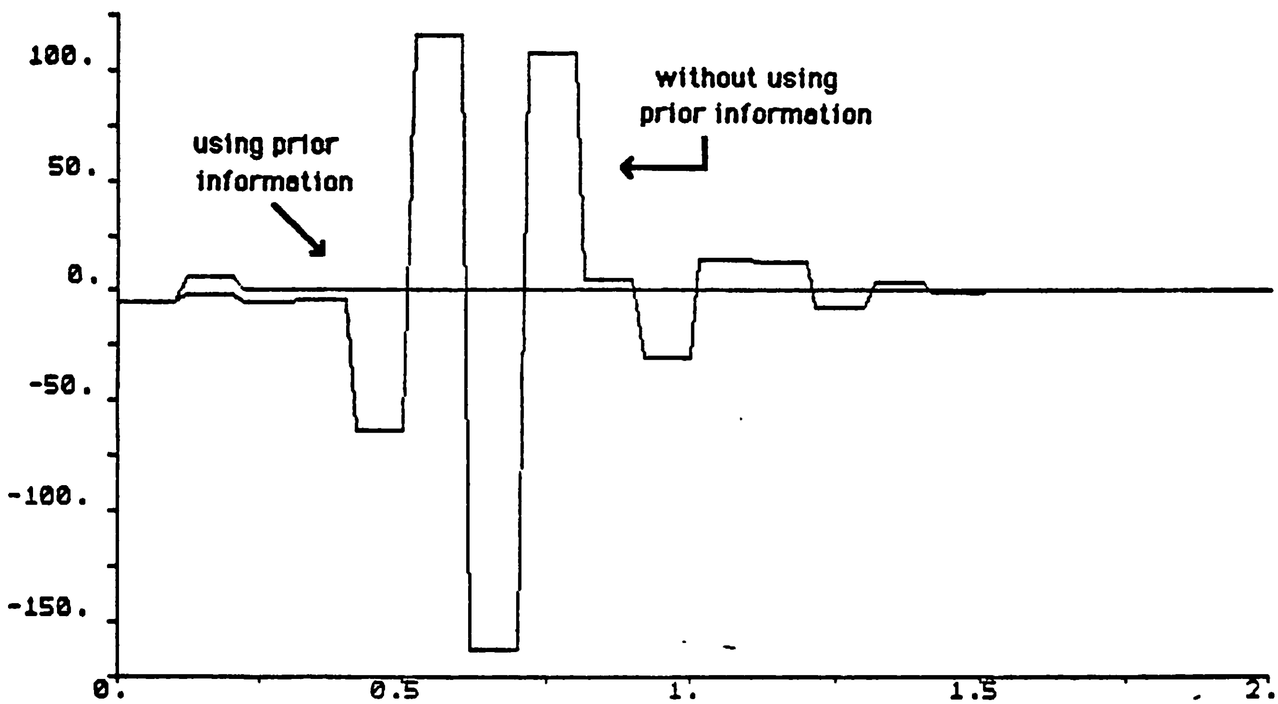
100.
50.
0.
-50.
-100.
-150.
0.
0.5
1.
1.5
2.
using prior information
without using prior information

Fig. 3-5.2

3-6 Concluding Remarks

In this chapter, we have presented algorithms, which utilize prior information about the plant, for adaptive control and identification of linear time invariant systems. If the plant is completely unknown, the algorithms are identical to the standard ones in the literature. However, the algorithms presented here have faster convergence rate and better transient performance when the system is partially known.

In the section 3-4, we have applied the technique of using rational function instead of polynominal to parameter estimation of continuous time systems. We feel that the framework is particularly amenable to the study of sensitivity of the schemes to the presence of unmodeled dynamics. This will prove to be partically impotant when we devise robust adaptive control schemes.

References

(1) Anderson, B.D.O., "Exponential Stability of Linear Equations Arising in Adaptive Identification," *IEEE Trans. on Automatic Control*, Vol. AC-22, No. 1, pp. 83-88, 1977

(2) Anderson,B.,Bitmead,R.,Johnson,C.Jr.,Kokotovic,P.,Kosut,R.,Mareels,I.,Praly,L. and Riedle,B. 'Stability of Adaptive System: Passivity and Averaging Analysis' The MIT Press Cambridge, 1986

(3) Anderson,B.D.O. and Johnstone,R.M. 'Global Adaptive Pole Positioning' *IEEE Trans. on Automatic Control* Vol.30, pp11-22, 1985

(4) Astrom, K.J and Wittenmark,B. 'On Self Tuning Regulators' *Automatica*, Vol. 9, pp. 185-199, 1973

(5) Bai,E.W. and Sastry,S.S. ' Discrete Time Adaptive control Utilizing Prior Information' *IEEE Trans. on Automatic Control* Vol.31 No.8, pp779-782, 1986

(6) Bai,E.W. and Sastry,S.S. 'Global Stability Proofs for Continuous Time Indirect Adaptive Control Schemes' *IEEE Trans. on Automatic Control* Vol.32 No.6, pp537-543, 1987

(7) Bai,E.W. and Sastry,S.S. ' Adaptive Stabilization of Sampled Systems' *Systems and Control Letters* Vol.8 No.5, pp389-395, 1987

(8) Bai,E.W. and Sastry,S.S. 'Parameter Identification Using Prior Information' *Int.J. Control* Vol.44, pp455-473, 1986

(9) Bai,E.W. and Sastry,S.S. 'Persistency of Excitation, Sufficient Richness and Parameter Convergence in Discrete Time Adaptive Control' *Systems and Control Letters* Vol.6 No.3 pp153-163, 1985

(10) Bai,E.W., Fu,L.C. and Sastry,S.S. 'Averaging Analysis for Discrete Time and Sampled Data Adaptive Systems' accepted and to appear in *IEEE Trans. on Circuits and Systems*

(11) Boyd,S and Sastry,S. 'On Parameter Convergence in Adaptive Control' *Systems and Control Letters* Vol.3 pp311-319, 1983

(12) Boyd, S. and Sastry, S.S. 'Necessary and Sufficient Conditions for Parameter Convergence in Adaptive Control' *Proc. of A.C.C.* San Diego, pp1584-1588, June,1984

(13) Clary,J.P. 'Robust Algorithms in Adaptive Control' Ph.D Thesis June Information Systems Lab. Stanford University, 1984

(14) Dasgupta, S. 'Adaptive Identification and Control' Ph.D. Thesis Australia National University, Canberra, Australia, 1984

(15) Dasgupta,S., Anderson B.D.O. and Kaye,R.J. 'Output Error Identification of Partially Known Systems' *Int.J.Control* Vol.43 pp.171-193, 1986

(16) Desoer,C.A. and Kuh,E. 'Basic Circuit Theory' McGraw-Hill Book Company New York, 1969

(17) Desoer, C.A., Liu, R.W., Murray, J. and Saeks, R. 'Feedback System Design: The Fractional Approach to Analysis and Synthesis' *IEEE Trans. on Automatic Control* Vol.AC-25 pp399-412, 1980

(18) Doyle,J. 'Advance in Multivariable Control' *Notes of the Honeywell Workshop on Multivariable Control* Minneapolis,MN. Oct.1984

(19) Elliot,H. Cristi,R. and Das, M. 'Global Stability of Adaptive Pole Placement Algorithms' *IEEE Trans. on Automatic Control* Vol.30, pp348-356, 1985

(20) Francis, B.A. and Vidyasagar, M. 'Algebraic and Topological Aspects of the Regulator Problem for Lumped Linear System' *Automatic* Vol.19 No.1 pp87-90, 1983

(21) Fu,L. and Sastry,S. 'Frequency Domain Synthesis of Optimal Inputs for Adaptive Identification and Control' Proc. 6th American Control Conf. Minneapolis, MN, 1987

(22) Fu, L.C., Bodson, M. and Sastry, S.S. 'New Stability Theorem for Averaging and Their Applications to the Convergence Analysis of Adaptive Identification and Control Schemes' Mem. UCB/ERL M85/21 University of California, Berkeley

(23) Goodwin,G.C. and Elliot,H. 'Deterministic Convergence of A Self-Tuning Regulator with Covariance Resetting' *IEE Proc.* Vol.130 Pt.D No.1 pp6-8 1983.

(24) Goodwin, G.C., Ramadge,P.J. and Caines,P.E. 'Discrete Time Multivariable Adaptive Control' *IEEE Trans. on Automatic Control* Vol.AC-25. pp449-456, 1980

(25) Goodwin, G.C. and Sin, K.S. 'Adaptive Filtering, Prediction and Control' Prentice-Hall, Englewood Cliffs, NJ 1984

(26) Hill, D.J., Middleton, R.H. and Goodwin, G.C. 'A Class of Robust Adaptive Control Algorithms' Proc. of 2nd IFAC Workshop on Adaptive Systems in Control and Signal Processing, Lund Sweden, pp25-30, 1986

(27) Hirsch,M. and Smale,S. 'On Algorithms for Solving f(x)=0' *Communications on Pure and Applied Mathematics* Vol.32, pp281-312, 1979

(28) Ioannou, P.A. and Kokotovic,P.V. 'Robust Redesign of Adaptive Control' *IEEE Trans. on Automatic Control*, Vol. AC-29, No. 3, pp. 202-211, 1984.

(29) Johnstone,R.M. and Anderson,B.D.O. 'Global Adaptive Pole Placement: Detailed Analysis of A First Order System' *IEEE Trans. on Automatic Control* Vol.28, pp852-855, 1983

(30) Johnstone,R.M.,Johnson,C.R.Jr.,Bitmead,R.R. and Andson, B.O.D. 'Exponential Convergence of Recursive Least Squares with Exponential Forgetting Factor' *Systems and Control Letters* Vol.2 No.2 pp77-82, 1982

(31) Kardashov,A.A. and Karnyushin,L. 'Determing SystemParameters from Experimental Frequency Characteristics' *Automation and Remote Control* Vol.19, No.4, 1958

(32) Kosut, R. and Johnson,R. 'An Input-Output View of Robustness in Adaptive Control' *Automatica*, Vol. 20, No. 5, pp. 569-581, 1984.

(33) Kokotovic,P.V.,Riedle,B. and Praly,L. 'On a Stability Criterion for Slow Adaption' *Systems and Control Letters* Vol.6 pp7-14, 1985

(34) Kreisselmeier, G. 'Adaptive Observers with Exponential Rate of Convergence' *IEEE Trans. on Automatic Control* Vol.AC-22, pp2-8, 1977

(35) Kreisselmeier,G. 'An Approach to Stable Indirect Adaptive Control' *Automatica* Vol.21, pp425-431, 1985

(36) Kreisselmeier,G 'A Robust Indirect Adaptive Control Approach' *IntJ. Control* Vol.43, pp161-175, 1986

(37) Levy,E.C. 'Complex Curve Fitting' *IRE Trans. on Automatic Control* Vol.AC-4, pp37-44,1959

(38) Lion, P.M. 'Rapid Identification of Linear and Nonlinear Systems' *AIAA Journal* Vol.5, pp1835-1842, 1967

(39) Luder,G. and Narendra,K.S. 'An Adaptive Observer and Identifier for a Linear Syatem' *IEEE Trans. on Automatic Control* Vol.AC-18, pp496-499 ,1973

(40) Ma,C.C.H. and Vidyasagar,M. 'Stability of Reduced Order Adaptive Control System' Preprint, University of Waterloo, Canada, 1985

(41) Mason,J.,Bai,E.W., Fu,L.C.,Bodson,M. and Sastry,S.S. 'Analysis of Adaptive Identifiers in the Presence of Unmodeled Dynamics: Averaging and Tuned Parameters' Submitted to *IEEE Trans. on AC* for publication

(42) Morgan, A.P. and Narendra,K. 'On the Uniform Asymptotic Stability of Certain Linear Nonautonomous Differential Equations' *SIAM J. Control and Optimization*, Vol. 15, No. 1, pp 5-24, 1977

(43) Nordstrom,N. and Sastry,S.S. 'Persistency of Excitation in Possibly Unstable Continuous Time Systems and Parameter Convergence in Adaptive Identification' Proc. of 2nd IFAC Workshop on Adaptive Systems in Control and Signal Processing, Lund Sweden 1986, pp193-198,

(44) Payne,P.A. 'An Improved Technique for Transfer Function Synthesis from Frequency Response Data' *IEEE Trans. on Automatic Control* Vol.15 pp480-483, 1970

(45) Polak,E. Salcudean,S. and Mayne,D.Q. 'A Sequential Optimal Redesign Procedure for Linear Feedback System' *IEEE Trans. on Automatic ControlR Vol.32, pp388-396, 1987*

(46) Rabkin,G.L.,Mitrofanov,B.A. and Shterenberg,Y.O. 'On the Determination of the Numerical Values of the Coefficients of the Transfer Function of Linearized Links and Systems from Experimental Frequency Characteristics' *Automation and Remote Control* Vol.16 No.5 1955

(47) Sanathanan, C.K. and Koemer,J. "Transfer Function Synthesis as a Ratio of Two Complex Polynomials," *IEEE Trans. on Automatic Control*, Vol. AC-8, No.1, pp56-58, 1963

(48) Sastry, S., "Model Reference Adaptive Control--Stability, Parameter Convergence, and Robustness," *IMA Journal of Mathematical Control & Information*, Vol. 1, pp 27-66, 1984

(49) Sondhi, M. and Metra,D. "New Results on the Performance of a Well-Known Class of Adaptive Filters," *Proc. IEEE*, Vol. 64, No.11, Nov. 1976

(50) Stahl, H., "Transfer Function Synthesis Using Frequency Response Data," *Int. J. Control*, Vol.39, No.3, pp541-550, 1984

(51) Vidyasagar, M. 'Control System Synthesis: A Factorization Approach' MIT Press 1984

(52) Vidyasagar,M. 'Nonlinear System Analysis ' Prentice-Hall, Englewood Cliffs, N.J., 1978

(53) Vidyasagar,M and Vannelli,A. 'New Relationships between Input-Output and Lyapunov Stability' *IEEE Trans. on Automatic Control* Vol.AC-27, pp481-483, 1982

(54) Zames,G. and Francis,B.A. 'A New Approach to Classical Frequency Methods: Feedback and Minimax Sensitivity' *IEEE Trans. on Automatic Control* Vol.28, pp585-601, 1983

(55) Zhang,Z. and Bai,E.W. 'Optimal Model-Reduction of Multivariable Systems with Sine-Function Input' *Control Theory and Applications* (in Chinese with English translations) Vol.1, pp79-86, 1984