ADAPTIVE IDENTIFICATION
AND
CONTROL

by

Soura Dasgupta
B.E. (Elec) (Hons. I) (Qld)

A thesis submitted for the degree of

Doctor of Philosophy
in
Systems Engineering
at

The Australian National University
Canberra, Australia

August, 1984

Chapters two to six of this thesis describe original research. Of these, sections 2.3, 3.3, 3.4, 4.1.5, 4.2, 4.3.1.1, 4.3.1.3, 4.3.3, 4.4, 5.2 and 6.1 contain work that is entirely mine. Sections 3.1, 3.5, 4.3.2, 5.1 and 6.2 are almost entirely and the rest substantially mine.

Soura Dasgupta

Contents

1. Introduction 1

1.1 Problem statement 1

1.2 Survey of recent adaptive identification and control literature 3

1.2.1 Adaptive identification 3

1.2.2 Adaptive control 7

1.3 Contributions of this thesis 9

1.4 Notation of this thesis 12

Figure for Chapter 1 13

References for Chapter 1 14

2. Realization theory 19

2.1 State variable realizations 21

2.1.1 RLC circuits with no pathologies 27

2.1.2 RLC circuits with inductor cutsets and capacitor loops 31

2.2 Transfer functions for SISO systems 37

2.3 A multi input multi output extension of the transfer function result 48

Figures for Chapter 2 53

References for Chapter 2 56

3. Persistence of Excitation 57

3.1 Identification of stable plants 60

3.2 Identification of unstable plants inside an overall stable system 68

3.3 Persistence of Excitation for unstable plants with possibly unbounded signals 77
3.4 Persistence of Excitation of slowly varying plants with bounded signals 88
3.5 Model Reference Control: Known Gain 90
3.6 Model Reference Control: Unknown Gain 96
Appendices for Chapter 3 103
Figures for Chapter 3 134
References for Chapter 3 142

4. Equation Error Identification 145
4.1 Parameter Adjustment Laws 147
4.1.1 The system and some notations 147
4.1.2 Least squares two step algorithm 152
4.1.3 An interpretation of the least squares two step algorithm 156
4.1.4 General two step algorithm 159
4.1.5 Single step algorithm 160
4.2 Persistence of Excitation 162
4.2.1 A PE condition on H(t) 162
4.2.2 A PE condition on u(t) 165
4.3 Convergence analysis 173
4.3.1 Convergence of least squares two step algorithm 173
4.3.1.1 Local and semiglobal stability 175
4.3.1.2 Assured uniform asymptotic convergence 178
4.3.1.3 A modified least squares two step algorithm 182

4.3.2 Convergence of the generalized two step algorithm 184
4.3.3 Convergence of the single step law 188
4.4 Some notions related to p.e. and parameter identifiability 193
4.5 Simulation results 199
Appendices for Chapter 4 205
Figures for Chapter 4 216
References for Chapter 4 233

5. Output Error Identification 235
5.1 A gradient descent output error algorithm 236
5.2 A recursive least squares formulation 250
Appendix for Chapter 5 261
References for Chapter 5 263

6. Adaptive Control 264
6.1 Adaptive control using modified least squares identifier 265
6.2 Adaptive control using a gradient descent identifier 281
6.3 Simulation results 295
Figures for Chapter 6 297
References for Chapter 6 300

7. Conclusions 301
7.1 Areas of further work 303
Reference for Chapter 7 304

Acknowledgements

A great many people have helped me negotiate this important phase of my training and it is a pleasure to formally record my gratitude.

First and foremost I must thank my supervisor Professor Brian Anderson and co-supervisor Dr. John Kaye for their invaluable contributions to my graduate training. Exposure to their considerable technical skills and friendly dispositions made my days as a graduate student both academically and personally rewarding.

Brian helped me select a research area which was appropriate to my aptitude and training. Especially beneficial was his ability to judge when to guide and direct and when to leave me to my own devices. John joined this department in only the second year of my training. But in this relatively short time he influenced me a great deal even beyond my academic endeavours. His healthy skepticism and piercing questions aided me to develop a more questioning attitude towards research. He also motivated me to address a range of ethical and moral issues which Systems Theorists can ill-afford to ignore.

I must thank Dr. Bob Bitmead for his ready access-ability and a willingness to be trapped in time consuming (for him) discussions. Professor John Moore through his keen interest and infectious enthusiasm was a continued source of encouragement.

I have benefitted from numerous discussions with Professor Michel Gevers and Dr. Rene Boel. Many short term visitors including, Dr. Rick Johnson, Dr. Gerhard Kreisselmeier Professor Lennart Ljung, Dr. Ray Rink and Dr. Ah Chung Tsoi have made suggestions for improving this thesis.

I must thank my fellow graduate students - Richard Johnstone, Hui Min Hong, Philip Parker, Geoff Latham, Tony Hotz and Michael Green - from whom too I have learnt a lot and whose friendship I continue to look forward to.

The typing of this thesis turned out to be far more formidable a task than I had expected. The following people did a splendid job with the manuscript, despite the unreasonably short notice I had given them: June Wilson, Susan Watson, Kerrie White, Deborah Spencer, Rosemary Drury and Alice Duncanson.

A special expression of gratitude is reserved for my friends outside the department. Their warmth, wit and companionship helped preserve my sanity and have sustained me through my sojourn in Canberra. Among them I must mention Anne, Aswath, Jan, Kerry, Marion, Mary Alice, Mathew, Philip, Richard and the three Steves.

I welcome this opportunity to thank the three people who have most influenced my earlier academic development; Mr. Jagannath Prasad Srivastava, who introduced me to the simplicity of scientific reasoning; Dr. Louis Westphal, who taught me the rudiments of Systems Theory; and Dr. Prabhakar Murthy who motivated me to study control theory.

Finally, a great debt remains to my parents, Gita and Sugata, and my sister Tanuka who braving the "tyranny of distance", have inspired me in this and every phase of my life. It is to them, more than any one else, that credit must go for any worthy aspects of this venture.

Abstract

This thesis considers the identification and control of linear time-invariant continuous time systems whose unknown parameters have direct physical relevance. In many such systems the transfer functions are shown to be ratios of two polynomials multilinear in the unknown parameters. Accordingly the algorithms proposed exploit this multilinearity.

For identification, several equation and output error algorithms are formulated. Barring one exception, all of these conform to a two step structure. The first, generates an unconstrained estimate of the parameter vector, by ignoring the inherent multilinearity. The second obtains a constrained estimate which is in some sense the nearest to the unconstrained estimate. In the presence of unideal plant behaviour, simulations show that this second step improves upon the accuracy of the estimates obtained in the first. The remaining identification algorithm essentially combines these two steps into one by employing a penalty function term.

One of the equation error algorithms, called the least squares two step algorithm, is uniformly asymptotically stable (u.a.s.) whenever it is implementable and its parameter estimates are initialized to zero. Implementability, however, is conditional on a persistence of excitation (p.e.) condition on the system inputs. The other algorithms are always implementable but are u.a.s. only when theis pe condition is met and when information about the parameter magnitude bounds is available. The latter knowledge is reasonable in view of the physical significance of the unknown parameters.

Also formulated are two indirect adaptive controllers. Both employ a general controller but differ in the identifier used. When excited by pe reference inputs, the first algorithm is globally stable, with uniform asymptotic parameter convergence, as long as the plant is completely controllable and observable. For the second law, the knowledge of a convex region, containing the true parameter value, is assumed. This region has the added property that the frozen closed loop system is asymptotically stable whenever both the plant and the controller are conditioned on the same parameter value belonging to this region. Subject to this assumption, uniform asymptotic parameter estimate convergence and signal boundedness follows due to pe reference inputs.

As a means of establishing input only p.e. conditions, several general tools are derived here. These are applicable not just to the parametrizations of this thesis alone but also to more conventional parametrization found in the literature. Furthermore the p.e. results are not restricted to stable time-invariant systems but apply also to unstable time-invariant plants and slowly time varying plants with bounded signals

§1. Introduction

1.1 Problem Statement:

This thesis considers the adaptive identification and control of partially known continuous time systems. The systems considered here are linear, time invariant, single input - single output and of known finite order.

In general linear, time invariant systems can be described by ordinary differential equations of the form

$$y^{n}(t) + \sum_{i=1}^{n-1} a_i \, y^{(i)}(t) = \sum_{j=0}^{m} b_j \, u^{(j)}(t) \qquad (1.1)$$

If the coefficients $\{a_i\}$ and $\{b_i\}$ are known and if the system satisfies stabilizability and detectability conditions, then the task of designing stable controllers is straightforward. Many of the coefficients may, however, be unknown and even slowly time varying. In such cases adaptive control constitutes one attractive approach to controller design involving a two step process. In the first unknown system parameters are estimated on line and are updated progressively as more and more data become available. In the second, these parameter estimates are used at each time instant to synthesize appropriate control signals.

Clearly such schemes depend heavily on the parameterisations selected. The easiest to handle and the most commonly used parametrisation has the $\{a_i\}$ and $\{b_i\}$ in (1.1) as the unknown parameters. Thereafter under the assumption of a complete lack of knowledge of the $\{a_i\}$ and $\{b_j\}$, barring possibly that of the order and the relative order, the system is identified and controlled on-line.

In practice, many of these parameters are known *a priori* as also are certain relationships, albeit nonlinear, which exists between them. The premise of this thesis is that the exploitation of this additional knowledge should lead to more efficient adaptive algorithms. Usually the unknownness in a system relates to certain physical parameter values. Thus all parts of a mechanical system may be known *a priori* except perhaps the values of a moment of inertia, a frictional coefficient or the like. Accordingly in the parametrization that we consider the unknown parameters have direct physical significance. Such a parametrization also has the following added attraction : the physics of the system allows us to make assumptions on the parameter magnitude bounds and in most cases on the knowledge of their signs. Of the various algorithms formulated in this thesis the stability analyses of some, but *not all*, exploit the knowledge of these assumed magnitude bounds.

Adaptive algorithms are usually designed and analysed under certain idealizing assumptions. It is thus commonly assumed that the system has no noise or other spurious disturbances, is time invariant and lies within an assumed model set. In all likelihood none of these ever hold. The algorithms designed should, thus not only work in the ideal case, but should be robust enough to withstand reasonable departures from these assumptions. A pre-condition for such robustness is that parameter convergence occur uniformly asymptotically when the idealizing assumptions hold [1,2] . Thus, as uniformly asymptotically stable algorithms are totally stable [3,107-108], the algorithms are equipped to overcome moderate deviations from ideality.

In summary, the object of this thesis is to formulate robust identification and control algorithms for systems where the unknown parameters are directly related to physical element values. In all the algorithms presented, we shall demand that parameter convergence occur in a uniform asymptotic manner as a pre-requisite to robust behaviour.

1.2 Survey of recent adaptive identification and control literature

Since this thesis is primarily concerned with continuous time systems this survey will mostly restrict itself to continuous time algorithms. In conducting this survey we shall treat the identification and control literature separately.

1.2.1. Adaptive Identification

Adaptive Identifiers in the literature can broadly be classified into two categories : equation error and output error. Consider the system defined by equation (1.1). In equation error an error signal e(t), defined below, is formed where the $\{\alpha_i\}$ and $\{\beta_j\}$ are respectively the estimates of $\{a_i\}$ and $\{b_j\}$

$$e(t) = y_n(t) + \sum_{i=0}^{n-1} \alpha_i(t)\, y_i(t) - \sum_{j=0}^{m} \beta_j(t)\, u_j(t) \qquad (1.2)$$

Here

$$y_i(t) \triangleq \frac{p^i}{\alpha(p)} y(t) ,$$
$$u_i(t) \triangleq \frac{p^i}{\alpha(p)} u(t) , \qquad (1.3)$$

p is the differential operator and $\alpha(p)$ is a Hurwitz polynomial. The signal $e(t)$ is then used to progressively update the $\{\alpha_i(t)\}$ and $\{\beta_j(t)\}$, with the object that they approach $\{a_i\}$ and $\{b_j\}$, respectively. The notion of using filtered versions of the derivatives of the system inputs and outputs is commonly known as state variable filtering. It was perhaps first introduced by Rucker [4] and is used primarily to avoid explicit differentiation of the system signals.

Output error algorithms on the other hand form an adjustable model

$$\hat{y}_n(t) + \sum_{i=0}^{n-1} \alpha_i(t)\hat{y}_i(t) = \sum_{j=0}^{m} \beta_j(t)\ u_j(t) \qquad (1.4)$$

with $\hat{y}_i(t)$ obviously defined. The output error $\hat{y}_o(t) - y(t)$ is then used to adjust the parameter estimates $\{\alpha_i(t)\}$ and $\{\beta_j(t)\}$. The difference between equation and output error algorithms is best understood through Figure 1.1. It is essentially a question of what constitute the inputs to the adjustable model. In equation error the exogenous inputs to this model are three : the unknown system input, the unknown system output and the difference between the outputs of the unknown system and the adjustable model. In output error, however, the unknown system output enters the adjustable model through the output error $\hat{y}_o(t) - y(t)$, only.

The main disadvantage of equation error algorithms is that they yield biased parameter estimates in the presence of unbiased measurement noise. Output error algorithms do not have this drawback, but their convergence is conditional on a certain transfer function being strictly positive real (SPR). Unfortunately, this transfer function depends on the system parameters and hence this condition for convergence cannot always be checked a priori.

Equation error algorithms in their simplest form are typified by those presented by Young [5] and Lion [6] (Lion calls equation error algorithms using state variable filtering as "generalized equation error" schemes). A further level of complexity is introduced in the schemes of Narendra and Kudva [7] , Lüders and Narendra [8] , Parks [9] and Caroll and Lindorff [10] . Their schemes lead to the use of fewer integrators and involve the use of positive real transfer functions. Unlike the output error situation, however, these transfer functions are independent of the unknown parameters and are thus not difficult to design. Anderson in [11] considers the multivariable extension of these schemes and demonstrates how all of the above [5 - 10] can be unified within the general framework of two prototype structures. As we shall demonstrate in Chapter 3 these schemes may fail to converge to the right parameter values in the presence of unbounded signals. Schemes suggested by Kreisselmeier [12 - 13] on the other hand are capable of tackling unbounded signals as well.

Conditions for the exponential convergence of the schemes in [5 - 13] have been derived variously by Morgan and Narendra [14,15] , Kreisselmeier [12, 13] , Sondhi and Mitra [16] and Anderson [11,17] . In direct terms their's is a persistently spanning condition on certain regression vectors involving system inputs and outputs. For a system with n+m-1 unknowns, for example, the condition requires that the regression vectors span the entire R^{n+m-1} space with time. Intuitively, this translates to a persistence of excitation condition on the system input, even though none of the above results have formalized this assertion.

Persistence of excitation can be viewed as a condition on system identifiability with the proviso that such identifiability should not be lost asymptotically. This in turn requires that the system be excited by inputs which are sufficiently rich in frequencies. For example, a system with two unknown parameters cannot be identified if the input is a d.c. signal. On the other hand a sinusoidal input should suffice as it carries with it two pieces of information namely its magnitude and phase.

As we have stated no precise connection between the persistently spanning conditions on the regression vectors and a persistently exciting condition on the system inputs emerges from [14 - 17] . Moreover, the former conditions, with their explicit dependence on the system outputs, are ill-suited to *a priori* input design. But based on them Yuan and Wonham [18] have considered the synthesis of almost periodic input signals which result in persistently spanning regression vectors. More recently [19] (see also Chapter 3 of this thesis), has presented persistence of excitation conditions on the inputs directly. Similar results, using the technique of Generalized Harmonic analysis have been derived by Boyd et.al. [20,21] . However, whereas the results in this thesis deal also with unstable systems, system stability is crucial to the derivation in [20,21]. Discrete time analogues of these results can be found in [22] .

Discussions of output error algorithms can be found in [23] . Although several discrete time proofs of the exponential convergence of output error algorithms exist [24, 22] we were unable to find any complete analysis of such convergence in continuous time. In [17,25] error models similar to those arising in output error algorithms are analysed under the implicit assumption of bounded signals. However, since, in principle the parameters of the adjustable model in (1.4) can vary arbitrarily, such an assumption seems difficult to justify. In Chapter 5 of this thesis complete analyses of the output error algorithms is presented.

1.2.2. Adaptive Control

Adaptive controllers can in general be classed into two categories : those employing the indirect and direct approaches. The former involve the explicit estimation of the system parameters which are then used to design the controller parameters. In the direct approach,on the other hand,the first phase is sidestepped and the controller parameters are directly estimated.

The analyses of adaptive controllers proved to be much more difficult than their identification counterparts. This stemmed from the feedback configuration which gave rise to nonlinear, time varying differential equations. The first important contribution in the direct control area was a model reference scheme for minimum phase plants, proposed by Monopoli [26], who used an augmented error signal to sidestep a positive real condition otherwise implicit in the analysis. These ideas were further developed by Feuer and Morse [27] , Narendra and Valvani [28], Narendra, Lin and Valvani [29] and Morse [30] . The last two in particular showed global asymptotic convergence of the output tracking error to zero. Their analysis, however, did not show parameter convergence to the correct values, without which, as we have already asserted, robustness may not be forthcoming. In [19] it has been shown that with persistently exciting reference inputs, and with known high frequency gain, parameter convergence for [30] not only occurs, but does so exponentially. Without the knowledge of the high frequency gain, however, exponential stability will not be obtainable. Similar results for the algorithm in [29] have been derived by Boyd et.al. [20,21] .

Åström and co-workers [31 - 33] , Kreisselmeier [34-36], and Elliott and Wolovich [37] in their work have created indirect algorithms which are globally stable, in that irrespective of the initial parameter estimates the system signals are always bounded. Egardt and Samson [38] in their work considered algorithms having a specific controller but a general identifier satisfying certain assumptions. Similarly, Kreisselmeier [36] considered a specific identifier coupled to a general controller. In [36] the minimum phase requirement is substituted by the assumption that the extent of a convex region containing the plant parameters, in which the plant is stabilizable and detectable is known.

Recent work by Rohrs et.al.[39] and Åström and Wittenmark [40] have thrown light upon the behaviour of adaptive controllers in the presence of unmodelled high frequency modes and bounded disturbances. Ioannou [41] and Ioannou and Kokotovic [42] have applied the singular perturbation method to show that the algorithm in [28] retains local stability in face of very high frequency dynamics. Moreover, Narendra and Peterson [43] , Kreisselmeier and Narendra [44] and Sastry [45] have considered the introduction of dead zones in adaptive controllers to tackle bounded disturbances. In [46 - 48] error models have been developed, based on which many adaptive controllers have been shown to retain local stability, even in face of departures from some of the idealizing assumptions, as long as the inputs are persistently exciting.

§1.3 Contributions of this thesis.

As stated earlier the primary objective is to formulate adaptive algorithms for the robust control of systems where the unknown parameters have direct physical relevance. In chapter 2 we motivate the parametrisation of interest . It is shown that the unknown physical element values of most linear time invariant electrical circuits affect the numerator and denominator polynomial coefficients in a multilinear fashion. Thus when two such parameters are unknown, the transfer function becomes

$$T(s,k_1,k_2) = \frac{p_o(s)+k_1p_1(s)+k_2p_2(s)+k_1k_2p_{12}(s)}{q_o(s)+k_1q_1(s)+k_2q_2(s)+k_1k_2q_{12}(s)} \qquad (1.5)$$

where the k_i are the unknown parameters and the polynomials $p_i(s)$ and $q_i(s)$ are known a priori . This result extends to mechanical and chemical analogues but excludes physical elements such as mutual inductors, which permit cross-coupling between energy storage devices and elements such as gyrators.

The algorithms we devise thus exploit the intrinsic multilinearity outlined above. As stated above, their robust behaviour would require that the input signals be persistently exciting (p.e.). However, even in the simple parametrisation of (1.1) , the p.e. conditions involve system outputs as well. In chapter 3 we develop a set of general tools for translating these to input only conditions. The systems for which these tools are applicable include ones which maybe unstable and those which are slowly time varying but have bounded system signals. These results are appealed to in specialized forms in establishing convergence conditions in the later chapters.

In chapter 4 three equation error algorithms are proposed for the identification of the systems in question.

Two of these involve two step methods which we illustrate through a two parameter example. Assume k_1 and k_2 are the unknown parameters and consider the vector $K \triangleq [k_1, k_2, k_1 k_2]^T$. The first steps ignore the dependence of the third element of K on the first two and generate an "unconstrained" estimate $K_u \triangleq [K_{u1}, K_{u2}, K_{u12}]^T$. The second step, which is common to both algorithms then finds $\hat{k}_1$ and $\hat{k}_2$ such that $[\hat{k}_1, \hat{k}_2, \hat{k}_1 \hat{k}_2]^T$ is the "closest" to K_u. The third algorithm, on the other hand, combines these two steps into one by using penalty functions ideas.

Chapter 5 presents two output error algorithms, based on the first two methods outlined above. These are analysed for uniform asymptotic convergence under the assumption of known parameter magnitude bounds. Chapter 6 presents two indirect adaptive controllers, both of which employ the same, general controller, but differ in the identifiers used. One is shown to be globally stable while stability of the second is established under assumptions similar to those in [3 6]. In all algorithms of chapters 4 - 6, persistence of excitation conditions, which yield global uniform asymptotic parameter convergence, are presented.

Chapter 7 presents the concluding remarks and indicates areas of further investigation.

§1.4 Notation of this thesis:

In this thesis, for the sake of clarity, notation shall be abused on several counts. To begin with quantities like

$$v(t) = \frac{a(s)}{b(s)} u(t)$$

shall refer to the solutions of the differential equation

$$b(p)v(t) = a(p)u(t) ,$$

$p \triangleq d/dt$, with arbitrary but finite initial conditions. In vectors such as

$$V(t) \triangleq [\frac{y(t)}{(s+\alpha)^n}, \frac{sy(t)}{(s+\alpha)^n} \cdots \frac{s^{n-1}y(t)}{(s+\alpha)^n}, \frac{u(t)}{(s+\alpha)^n}, \cdots \frac{s^m u(t)}{(s+\alpha)^n}]^T$$

or

$$W(t) \triangleq [u(t) , \frac{u(t)}{s+\beta} , \ldots , \frac{u(t)}{(s+\beta)^{n+m}}]^T$$

the initial conditions shall be assumed to be zero. Also $v(s)$ will refer to the Laplace transform of $v(t)$. We shall often use sets as subscripts for denoting elements of a vector. For example, the elements of a vector K_u shall be denoted by K_{ur} where r is a set. Thus if $r = \{1,2\}$ K_{ur} will K_{u12} .

The symbol "$\equiv$" shall denote "identically equal to". Thus $v(x) \equiv u(x)$ is the same as $v(x) = u(x)$ for all x belonging to the domain of x .

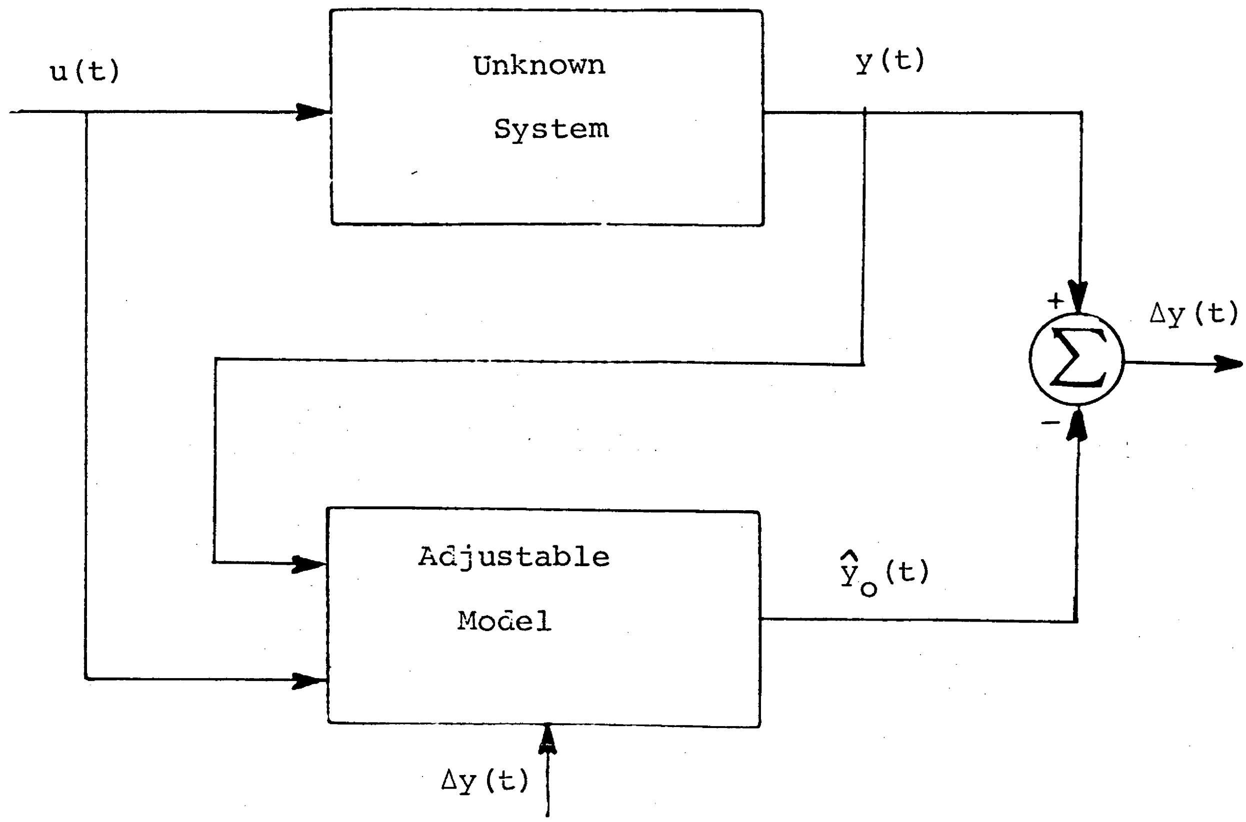


Figure 1.1(a) Equation error configuration.

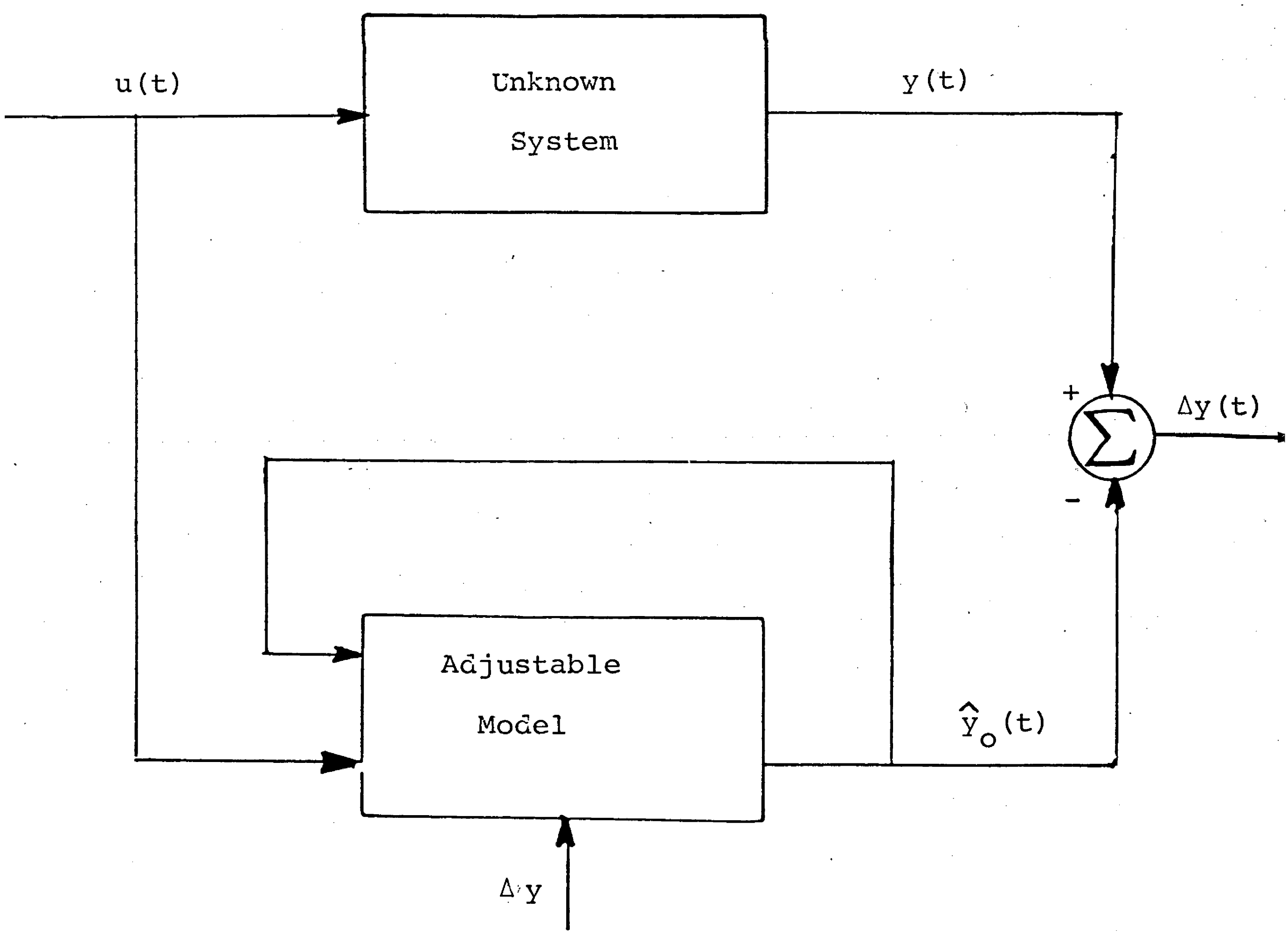


Figure 1.1(b) Output error configuration.

References for Chapter 1

1. B.D.O. Anderson, "Adaptive systems, lack of persistence of excitation and bursting phenomena". Submitted to Automatica.

2. B.D.O. Anderson and R.M. Johnstone, "when Will adaptive systems really adapt? The robustness issue",Proc. 2nd Conf. Control Engg, IE Australia, 1982, pp 59-66

3. W. Hahn, Stability of motion, Berlin, Springer Verlag, 1967.

4. R.A. Rucker,"Real time system identification in the presence of noise". Proc. Western Electronics Conv,Paper 2-3,Aug, 1963.

5. P.C. Young,"Discussion on 'In flight dynamic checkout' by J.E. Walker", IEEE Trans. Aerospace, AS-2, No. 3, pp 1106-1111.

6. P.M. Lion, "Rapid identification of linear and nonlinear systems", AIAA Journal,vol 5, Oct 1967, pp 1835- 1842.

7. K.S. Narendra and P. Kudva, "Stable adaptive schemes for system identification and control - parts I and II",IEEE Trans Systems Man Cyber, vol SMC-4, pp 542-560, Nov 1974.

8. G. Luders and K.S. Narendra, "Stable adaptive schemes for estimation and identification of linear systems", IEEE Trans Auto Contr, AC-19, pp 841-847, Dec 1974.

9. P.C. Parks, "Lyapunov redesign of model reference adaptive control systems",IEEE Trans Aut Contr, AC-11, No. 3,pp 362-367, 1966.

10. R.L. Carroll and D.P. Lindorff, "An adaptive observer for single input single output linear systems", IEEE Trans Auto Contr., AC-18, NO. 5, pp 428-435, 1973.

11. B.D.O. Anderson, "An approach to multivariable system identification", Automatica, vol 13, pp 401-408,1977.

12. G. Kreisselmeier, "Adaptive observers with arbitrary exponential rates of convergence", IEEE Trans Aut Contr, AC-21, Jan 1977 pp 2-8.

13. G. Kreisselmeier, "Algebraic separation in realizing a linear state feedback control law by means of adaptive observers", IEEE Trans Auto Cont, AC-25, April 1980, pp 238-243.

14. A.P. Morgan and K.S. Narendra, "On the uniform asymptotic stability of certain linear nonautonomous differential equations", SIAM J Contr, vol 15, Jan 1977, pp 5-24.

15. A.P. Morgan and K.S. Narendra, "on the stability of non-autonomous differential equations $\dot{x}=(A+B(t))x$ with skew-symmetric matrix $B(t)$", SIAM J Cont Opt, vol 15, Jan 1977 pp 163-176.

16. M.M. Sondhi and D. Mitra, "New results on the performance of a well known class of adaptive filters", Proc. IEEE, vol 64, Nov 76, pp 1583-1597.

17. B.D.O. Anderson, "Exponential stability of linear equations arising in adaptive identification", IEEE Trans Auto Cont, AC-22, Feb 1977, pp 83-88.

18. J.S.C. Yuan and W.M. Wonham, "Probing signals for model reference identification", IEEE Trans Auto Cont, AC-22, Aug 1977, pp 530-538.

19. S. Dasgupta, B.D.O. Anderson and A.C. Tsoi, "Input conditions for continuous time adaptive system problems", Proc 22nd CDC, San Antonio, Texas, Dec 1983. Submitted to IEEE Trans Auto Contr.

20. S. Boyd and S. Sastry, "Necessary and sufficient conditions for parameter convergence in adaptive control", Tech. Report, University of California, Berkeley, UCB/ERL M84/25, 1984.

21. S. Boyd and S. Sastry, "On parameter convergence in adaptive control", Systems and Control Letters, pp 311-319, Dec 1983

22. B.D.O. Anderson and C.R. Johnson Jr., "Exponential convergence of adaptive identification and control algorithms", Automaatica, vol 18, Jan 1982, pp 1-13.

23. Y.D. Landau, Adaptive control: the model reference approach Marcel Dekker, Inc., 1979.

24. C.R. Johnson, Jr.," A convergence proof for a hyperstable adaptive recursive filter", IEEE Trans Info Thy, vol IT-25 Nov 1979, pp 746-749.

25. K.S. Narendra and L.S. Valvani, "A comparison of Lyapunov and hyperstability approaches to adaptive control of continuous systems", April, 1980, pp 243-247, AC-25.

26. R.V. Monopoli, "Model reference adaptive control with an augmented error signal", IEEE Trans Auto Cont, AC-19,pp 474-484, October 1974.

27. A. Feuer and A.S. Morse, "Adaptive control of single input single output linear systems", IEEE Trans Auto Contr, AC-23 pp 557-570, Aug 1978.

28. K.S. Narendra and L.S. Valvani, "Stable adaptive controller design- direct control", IEEE Trans Auto Contr, AC-23,pp 570 - 583, Aug 1978.

29. K.S. Narendra, Y.H. Lin and L.S. Valavani, "Stable adaptive controller design - part II :Proof of stability",AC-25, pp 440-448, June 1980.

30. A.S. Morse, "Global stability of parameter adaptive control systems", IEEE Trans Aut Contr, AC-25, pp 433-440, June 1980.

31. K.J. Astrom and B. Wittenmark, "On self-tuning regulator", Automatica,No. 8, pp 185-199, 1973.

32. K.J. Astrom, U. Borisson, L. Ljung and B. Wittenmark,

"Theory and applications of self tuning regulators", Automatica, vol 13, pp 457-476, 1977.

33. K.J. Astrom, B.Westerberg, B. Wittenmark,"Self - tuning controllers based on pole placement design", CODEN: LUTFD2/(tfrt-3148)/1-52/ Department of Automatic Control Lund Inst. of Tech, Sweden, 1978.

34. G. Kreisselmeier, "Adaptive control via adaptive observation and asymptotic feedback matrix synthesis," IEEE Trans Auto Contr,AC-25, pp 712-722, Aug 1980.

35. G. Kreisselmeier, "On Adaptive state regulation", IEEE Trans Auto Contr, AC-27, pp 3-16, Feb, 1982.

36. G. Kreisselmeier, "An approach to stable indirect adaptive control", Submitted for publication.

37. H. Elliot and W.A. Wolovich, "Parameter adaptive identification and control", IEEE Trans Auto Contr, AC-24, pp 592-599, Aug 1979.

38. B. Egardt and C. Samson,"Stable adaptive control of non-mimimum phase systems", Systems And Control Letters, vol 2 pp 137-144, Oct. 1982.

39. C. Rohrs, L.S. Valavani and M. Athans, "Convergence studies of adaptive control algorithms, part I: Analysis," Proc CDC, Albuquerque, New Mexico, 1980, pp 1138-1141.

40. K.J. Astrom, "Analysis of Rohrs counterexamples to adaptive control", Proc. 22nd CDC, San Antonio, Texas, vol 2,pp 982-987, Dec 1983.

41. P. Ioannou, Robustness of model reference adaptive schemes with respect to modelling errors, Ph.D. Thesis, University of Illinois at Urbana-Champaign, August 1982.

42. P. Ioannou and P. Kokotovic, "Singular perturbations on robust redesign of adaptive control", IFAC workshop on

singular perturbations and robustness of control systems, Lake Ohrid, Yugoslavia, July 1982.

43. B.B. Peterson and K.S. Narendra, "Bounded error adaptive control", IEEE Trans Auto Contr, AC-27 1982 pp 1161-1168.

44. G. Kreisselmeier and K.S. Narendra, "Stable model reference adaptive control in the presence of bounded disturbances", IEEE Trans Auto Contr ,ac-27, 1982, pp 1169-1176.

45. S. Sastry,"Model reference adaptive control- stability, parameter convergence and robustness", Tech. Report, University of California at Berkeley, No. UCB/ERL M83/59, 1983.

46. R.L. Kosut, C.R.Johnson, Jr. and B.D.O. Anderson, "Robustness of reduced-order adaptive model following", Workshop on applications of adaptive systems theory, Yale University, 1983.

47. R.L. Kosut, C.R. Johnson,Jr. and B.D.O. Anderson, "Conditions for local stability and robustness of adaptive control systems", Proc. 22nd CDC, San Antonio 1983, pp 972-976.

48. C.R. Johnson, Jr., B.D.O. Anderson and R.R. Bitmead, "A model reference adaptive controller formulated for local optimality in non-ideal use", Presented at the 18th Conf. on Info. Sciences and Sys., Princeton NJ, 1984.

§2. Realization Theory

The purpose of this chapter is to motivate the form of models to which the adaptive algorithms of this thesis are applicable. As stated earlier the object is to arrive at parameterizations which involve quantities with direct physical relevance. Accordingly, this chapter analyses the way in which physical element values of lumped linear electric circuits appear in state variable realizations and transfer function descriptions. Extensions to mechanical and chemical analogues are then immediate.

The primary goal is to show that when most parts of a linear, time-invariant, finite-dimensional system are known,but certain parameters associated with physical components of the system are unknown, then the transfer function of the system can be viewed as a ratio of two polynomials,with the polynomial coefficients multilinear in the unknown parameters. For example, with three unknown parameters, the transfer function of a single input single output (SISO) system is of the form

$$T(s_1 k_1, k_2, k_3) = \frac{P(s,k_1,k_2,k_3)}{Q(s,k_1,k_2,k_3)} \; ; \qquad (2.1)$$

$$P(s_1,k_1,k_2,k_3) = p_o(s)+k_1p_1(s)+k_2p_2(s)+k_3p_3(s)+k_1k_2p_{12}(s)$$
$$+ k_1k_3p_{13}(s) + k_2k_3p_{23}(s)+k_1k_2k_3p_{123}(s)$$

and

$$Q(s,k_1,k_2,k_3)=q_o(s)+k_1q_1(s)+k_2q_2(s)+k_3q_3(s)+k_1k_2q_{12}(s)$$
$$+ k_1k_3q_{13}(s)+k_2k_3q_{23}(s)+k_1k_2k_3q_{123}(s)$$

Here, k_1, k_2, k_3 are the unknown parameters, the polynomials $p_i(s)$ and $q_i(s)$ are known a priori and T is proper.

As a means of establishing this property we shall first examine the manner in which the state variable realizations of the systems in question are affected by these unknown parameters. In particular it will be shown that such parameters appear in state variable realizations in a rank-1 fashion, the definition of which is given below.

Definition 2.1

A state variable realization described by the quadruple $\{F,G,H,J\}$ has a rank-1 dependence on N parameters $k_1,\ldots,k_N$ if for all $i \in \{1,\ldots,N\}$ $\exists$ $a_i, b_i \in \mathbb{R}$, independent of k_i, such that if we define α_i by either

$$\alpha_i = \frac{1}{a_i + k_i b_i}$$

or

$$\alpha_i = \frac{k_i}{a_i + k_i b_i}$$

then the following hold :

(i) The elements of F,G,H and J are multilinear in the α_i.

(ii) The matrices

$$\frac{\partial}{\partial \alpha_i}\begin{bmatrix} sI - F & G \\ -H & J \end{bmatrix}$$

have rank no greater than 1, $\forall\, i \in \{1,\ldots,N\}$.

Remark:

(2.1) The a_i and b_i defined above may depend on k_j $\forall\, i \neq j$.

In section 2.1 we shall establish this rank-1 property for electrical circuits. This will be done by examining in turn electric circuits having :

(i) resistor, inductor and capacitor (RLC) elements only, and no capacitor loops or inductor cutsets.

(ii) the above elements and possibly also pathologies such as inductor cutsets and capacitor loops.

In section 2.2 the corresponding transfer function result for the SISO case will be given while, in section 2.3 we shall also show that most electric circuits will retain this transfer function property even if one is unable to make definite statements about the state variable realizations. This result will pre-suppose the existence of certain input-output descriptions.

The contents of this chapter are the subject of [1] .

2.1 State variable realizations:

Much of the background material for this section, namely the construction of state variable realizations for electric circuits, is contained in [2,pp156-209]. We shall show that by following the general construction procedure outlined in [2] we arrive at state variable realizations which obey the rank-1 property.

To understand how electric circuit elements generally appear, consider a resistor R appearing in an n-port circuit. Clearly, the resistor can be extracted from the rest of the circuit in a manner depicted in figure 2.1.

Now, suppose U is a vector of inputs having a set of voltages or currents at the n-ports as its elements. Similarly suppose Y is an output vector, containing a disjoint set of voltages and currents as its elements. Suppose also that u_1 and y_1 are v_R and i_R, respectively (v_R and i_R are the respective voltage across and current through the resistor R). Then, if the hybrid matrix relating $[U^T, u_1]^T$ to $[Y^T, y_1]^T$ exists, the input-output description typified by figure 2.2 exists. In figure 2.2, $k_1 = R$. Sometimes, when the hybrid description relating $[U^T, u_1]^T$ to $[Y^T, y_1]^T$ does not exist, one may need to replace k_1 by $1/k_1$, with $u_1 = i_R$ and $y_1 = v_R$. Similarly other element values can also be extracted, in most circuits, in a manner typified by figure 2.1, though u_1 and y_1 may not necessarily represent voltages and currents. The exceptions to this rule arise from element values like mutual inductors, which allow crosscoupling to occur between different energy storage elements or from elements such as gyrators. In this section, however, we are only interested in extracting resistor values, in a manner depicted in figure 2.2. The following lemma indicates the special way in which the parameter k_1 extractable as in figure 2.2, affects input-output description relating U and Y. Here, as in the rest of the thesis we shall abuse notation by denoting U(s), for example, to be the Laplace transform of U(t). Similarly H(s)U(t) will denote the inverse Laplace transform of H(s)U(s).

Lemma 2.1

Consider a multiinput-multioutput system with input vector $U(t)$ and output vector $Y(t)$ and a parameter k_1. Suppose that the parameter k_1 can be extracted as in figure 2.2, with $u_1(t)$ and $y_1(t)$ scalar signals and the input-output description relating $[U^T(t),u_1(t)]^T$ to $[Y^T(t),y_1(t)]^T$ exists and is independent of k_1. Suppose $H(s,k_1)$ is the rational transfer function from $U(s)$ to $Y(s)$, i.e.

$$Y(s) = H(s,k_1)U(s).$$

Then there exist rational matrices $M(s)$ and $N(s)$ and scalar rationals $c(s)$ and $d(s)$, all independent of k_1 such that

$$H(s,k_1) = \frac{M(s) + k_1N(s)}{c(s) + k_1d(s)} \tag{2.2}$$

$$\frac{\partial H(s,k_1)}{\partial k_1} = \frac{c(s)N(s) - d(s)M(s)}{\{c(s)+k_1d(s)\}^2} \tag{2.3}$$

$$\text{rank } \{c(s)N(s) - d(s)M(s)\} \leq 1 \quad \forall\, s \in C^1 .$$

Also if $d(s) \not\equiv 0$ then there exists

$$\alpha(s,k_1) = 1/\{c(s)+k_1d(s)\} \tag{2.3a}$$

such that

$$H(s,k_1) = \bar{H}(s)+\alpha(s,k_1)\, \bar{M}(s) \tag{2.4}$$

with $\overline{H}(\cdot)$ and $\overline{M}(\cdot)$ independent of $\alpha(s,k_1)$ and rank $\overline{M}(\cdot) \leq 1$. Similarly if $c(s) \not\equiv 0$ then there exists

$$\alpha(s,k_1) = k_1/(c(s)+k_1 d(s))$$

such that

$$H(s,k_1) = \hat{H}(s) + \alpha(s,k_1)\hat{M}(s) \qquad (2.5)$$

$\hat{H}(s)$, $\hat{M}(s)$ independent of $\alpha(s,k_1)$ and rank $\hat{M}(\cdot) \leq 1$.

Proof.

Suppose

$$\begin{bmatrix} Y(s) \\ y_1(s) \end{bmatrix} = \begin{bmatrix} H_{11}(s) & h_{12}(s) \\ h_{21}^T(s) & h_{22}(s) \end{bmatrix} \begin{bmatrix} U(s) \\ y_1(s) \end{bmatrix} \qquad (2.6)$$

where $H_{11}(s)$ is a matrix, $h_{12}(s)$ and $h_{21}(s)$ vectors and $h_{22}(s)$ a scalar. Thus

$$y_1(s) = h_{21}^T(s)\ U(s) - k_1 h_{22}(s)\ y_1(s)$$

whence

$$Y(s) = [H_{11}(s) + \frac{k_1 h_{12}(s) h_{21}^T(s)}{1 + k_1 h_{22}(s)}]\ U(s) \qquad (2.7)$$

from which (2.2) and (2.3) follow with $c(s) = 1$ and $d(s) = h_{22}(s)$. Further,

$$\frac{\partial H(s,k_1)}{\partial k_1} = \frac{h_{22}(s)h_{12}(s)h_{21}^T(s)}{1 + k_1 h_{22}(s)}$$

with α defined as in (2.3a) . Thus (2.4) follows. Similarly (2.5) also follows. ∇∇∇

<u>Remarks</u>

(2.2) Replacing k_1 by $1/k_1$ does not alter the conclusions of the lemma, though (2.4) and (2.5) will be interchanged.

(2.3) If k_1 is a resistor value then $\alpha(s,k_1)$ is independent of s .

Consider now a circuit having N parameter values, $k_1,\ldots,k_N$ such that each k_i can be extracted so that input-output description of the form in figure 2.3 exist, with U and Y containing appropriate currents and voltages as their elements. Then the following lemma extends the result of lemma 2.1 to this case.

<u>Lemma 2.2</u>

Consider a system with input vector $U(t)$ and output vector $Y(t)$ and N parameters $k_1,\ldots,k_N$. Suppose each parameter can be extracted so that input-output descriptions of the form of figure 2.3 exist. Define $k \triangleq [k_1,\ldots,k_N]^T$ and $k^{(i)} = [k_1,\ldots k_{i-1},k_{i+1},\ldots k_N]^T$. Then $\forall i \; \exists \; c_i(s,k^{(i)})$ and $d_i(s,k^{(i)})$, rational scalars in s and elements of $k^{(i)}$ such that every element of $H(s,k)$, defined by

$$Y(s) = H(s,k)U(s)$$

is multilinear in the $\{\alpha_i\}$ with α_i defined as $1/(c_i(s,k^{(i)}) + k_i d_i(s,k^{(i)}))$ or

or $k_i/(c_i(s,k^{(i)}) + k_i d_i(s,k^{(i)}))$.

Remark:

(2.4) The α_i correspond to the α_1 in lemma 2.1. the first definition of α_i will not apply if $d_i \equiv 0$ and the second if $c_i \equiv 0$. If neither c_i nor d_i is zero then the lemma holds for α_i defined either way.

Proof

Consider an arbitrary element $h_{pq}(s,k)$ of $H(s,k)$. Then from (2.4) or (2.5) of lemma 2.1, we have that

$$h_{pq}(s,k) = \bar{h}^{(i)}_{pq}(s,k^{(i)}) + \alpha_i \, m^{(i)}_{pq}(s,k^{(i)}) \qquad (2.8)$$

where $\bar{h}^{(i)}_{pq}$ and $m^{(i)}_{pq}$ are independent of α_i . Consider

$$h_{pq}(s,k) = \bar{h}^{\ell}_{pq}(s,k^{(\ell)}) + \alpha_\ell \, m^{\ell}_{pq}(s,k^{(\ell)}) \qquad (2.9)$$

where $\bar{h}^{\ell}_{pq}$ and m^{ℓ}_{pq} are independent of α_ℓ but not necessarily of α_i . From (2.8) , h_{pq} is affine in α_i. Thus a simple argument shows that $\bar{h}_{pq}^{(\ell)}$ and $m_{pq}^{(\ell)}$ are also affine in α_i and that h_{pq} is bilinear in α_i and α_ℓ for all i and ℓ, $i \neq \ell$. Proceeding along these lines the result follows.

∇∇∇

2.1.1. <u>RLC circuits with no pathologies</u>:

Consider an n-port RLC circuit having n_i inputs and n_o outputs. Suppose that all input and output quantities are either port currents or port voltages. Denote u to be the n_i dimensional input vector and y to be the n_o dimensional output vector and assume that the elements of u and y do not overlap. Augment u and y to form the n-dimensional input and output vectors U and Y respectively, in the following manner. Assign all elements of u and y to U and Y respectively. Suppose that the j^{th} port current is an input. Then assign the j^{th} port voltage to Y. Similarly, if the j^{th} port current is an output then assign the corresponding voltage to U. If for a particular port neither the current nor the voltage are in either of u or y, then assign one of these arbitrarily to U and the other to Y. In this way for every port either the voltage or the current appears in U and the other appears in Y.

Consider next the following reactance extraction procedure illustrated in figure 2.4. Suppose there are n_L inductors and n_C capacitors in the circuit in question. Form the vectors U_C, U_L, Y_C and Y_L in the following way. Assign the voltage across the j^{th} capacitor to the j^{th} element of U_C and the corresponding current to the j^{th} element of Y_C. Likewise, assign the j^{th}

inductor current to the j^{th} element of U_L and the corresponding voltage to the j^{th} element of Y_L . Now, suppose all the capacitor and inductor connections were open-circuited. Then we are left with an $n + n_L + n_c$ circuit containing resistors only. Then denoting $[U^T, U_c^T, U_L^T]^T$ and $[Y^T, Y_c^T, Y_L^T]^T$ as input and output vectors respectively, define M

$$M \triangleq \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix}$$

as the hybrid matrix relating the two, i.e.

$$\begin{bmatrix} Y \\ Y_c \\ Y_L \end{bmatrix} = \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix} \begin{bmatrix} U \\ U_c \\ U_L \end{bmatrix} \qquad (2.10)$$

Since, the network relating the two vectors is resistive, M is non-dynamic. Moreover, it has been shown in [2] that with Y, Y_c, Y_L, U, U_c and U_L selected as above M exists and is unique, whenever the RLC circuit we started with is free of inductor cutsets and capacitor loops. Indeed the existence and uniqueness of this M is the standing assumption for this sub-section.

Assumption 2.1

With Y, Y_c, Y_L, U, U_c, Y_L and M defined as above, M exists and is unique.

Theorem 2.1

Consider an RLC circuit having n_L inductors and n_C capacitors, with quantities y, Y, Y_C, Y_L, u, U, U_C, U_L and M defined as above. Suppose, assumption 2.1 holds. Then the (n_L+n_C)-dimensional state variable representation having u and y as input and output and $[U_C^T, U_L^T]^T$ as the state vector, has a rank-1 dependence on all the inductor, capacitor and resistor values appearing in the circuit.

Proof:

We first show that the state variable representation having U, Y and $[U_C^T, U_L^T]^T$ as the input, output and state vectors respectively has a rank-1 dependence on all the elements. From this the conclusion of the theorem will directly follow.

Suppose R_i is a resistance appearing in the reactance extracted network. Then as argued before an input-output description of the form in figure 2.5 is possible. Then by lemma 2.1 there exist a_i, b_i R such that with $\alpha_i = R_i/(a_i+R_ib_i)$ or $\alpha_i = {}^1/(a_i+R_ib_i)$, $\partial M/\partial\alpha_i$, M defined in (2.10), has rank-1. Also by lemma 2.2 the elements of M are multilinear in α_i. Let Λ_c and Λ_L be diagonal matrices having all the capacitor and inductor values respectively. Then provided that elements of Λ_c and Λ_L are appropriately ordered it follows that

$$Y_c = \Lambda_c \dot{U}_c \qquad \text{and} \qquad Y_L = \Lambda_L \dot{U}_L$$

Then with U,Y and $[U_c^T, U_L^T]^T$ as the input, output and state vectors respectively, the following representation results :

$$\begin{aligned} \dot{x} &= Fx + GU \\ Y &= Hx + JU \end{aligned} \qquad (2.11)$$

where

$$F = -\begin{bmatrix} \Lambda_c^{-1} M_{22} & \Lambda_c^{-1} M_{23} \\ \Lambda_L^{-1} M_{32} & \Lambda_L^{-1} M_{33} \end{bmatrix} ; \quad G = -\begin{bmatrix} \Lambda_c^{-1} M_{21} \\ \Lambda_L^{-1} M_{31} \end{bmatrix}$$

$$H = [M_{12} \quad M_{13}] \qquad\qquad J = M_{11}$$

Clearly, if we choose α_{c_i} to be ${}^1/_{C_i}$ for each capacitor value and $\alpha_{L_i} = {}^1/_{L_i}$ for each inductor value then α_i, α_{c_i}, α_{L_i} appear in (2.11) in a rank-1 fashion. Hence the result follows. ∇∇∇

2.1.2 RLC circuits with inductor cutsets and capacitor loops

Suppose inductor cutsets and capacitor loops do appear in the circuits in question. Then from [2] it is clear that assumption 1 need not hold. As it turns out reactance extraction is still possible but U_c, U_L, Y_c and Y_L need to be redefined.

Suppose the reactance extracted $n+n_c+n_L$ port circuit has $(n+n_L+n_c)$-dimenstional input and output vectors $\bar{U}$ and $\bar{Y}$ respectively. Suppose also that all the elements of U, defined in the previous section, are in $\bar{U}$ and all the elements of Y are in $\bar{Y}$. Also if the j^{th} capacitor or inductor current appears in $\bar{U}$ then the corresponding voltage appears in $\bar{Y}$ and vice versa. Then by [2] there always exists a selection of $\bar{U}$ and $\bar{Y}$ such that with $\bar{M}$ defined as

$$\bar{Y} = \bar{M}\,\bar{U}$$

$\bar{M}$ exists and is unique. Clearly $\bar{M}$ is also non-dynamic. Moreover, $\bar{U}$ and $\bar{Y}$ can be partitioned as

$$\bar{U}^T = [U^T, U_{c_1}^T, U_{L_1}^T, U_{c_2}^T, U_{L_2}^T]$$

and

$$\bar{Y}^T = [Y^T, Y_{c_1}^T, Y_{L_1}^T, Y_{c_2}^T, Y_{L_2}^T]$$

where U_{c_1} and U_{L_1} carry, respectively, those of the capacitor voltages and inductor currents which have been assigned to $\bar{U}$. As before, if any element of U_{c_i} or

U_{L_i} is a voltage then the corresponding element of Y_{C_i} or Y_{L_i} is a current and vice versa. Then the following theorem is true.

Theorem 2.2

Consider an RLC circuit with input u and output y and with $U, Y, U_{C_1}, Y_{C_1}, U_{L_1}, Y_{L_1}, U_{C_2}, Y_{C_2}, U_{L_2}, Y_{L_2}$ and $\overline{M}$ defined as above. Suppose n_{C_1} and n_{L_1} are the dimensions of U_{C_1} and U_{L_1} respectively and that $\overline{M}$ exists and is unique. Suppose also that the transfer function relating U to Y is proper. Then the $(n_{C_1}+n_{L_1})$ dimensional state variable realization of the circuit, having, u,y and $[U_{C_1}^T, U_{L_1}^T]^T$ as the input, output and state vectors respectively, has a rank-1 dependence on the elements of the circuit.

Proof:

As in theorem 2.1 we first show that the state variable realization having U, Y and $[U_{C_1}^T, U_{L_1}^T]^T$ as the input, output and state vectors respectively, has a rank-1 dependence on the circuit elements. From this the result will follow. According to [2]

$$\begin{bmatrix} Y \\ Y_{C_1} \\ Y_{L_1} \\ Y_{C_2} \\ Y_{L_2} \end{bmatrix} = \begin{bmatrix} \overline{M}_{11} & \overline{M}_{12} & \overline{M}_{13} & \overline{M}_{14} & \overline{M}_{15} \\ \overline{M}_{21} & \overline{M}_{22} & \overline{M}_{23} & \overline{M}_{24} & \overline{M}_{25} \\ \overline{M}_{31} & \overline{M}_{32} & \overline{M}_{33} & \overline{M}_{34} & \overline{M}_{35} \\ \overline{M}_{41} & \overline{M}_{42} & \overline{M}_{43} & 0 & 0 \\ \overline{M}_{51} & \overline{M}_{52} & \overline{M}_{53} & 0 & 0 \end{bmatrix} \begin{bmatrix} U \\ U_{C_1} \\ U_{L_1} \\ U_{C_2} \\ U_{L_2} \end{bmatrix} \qquad (2.12)$$

where $\overline{M}_{i4} = -\overline{M}^T_{4i}$ and $\overline{M}_{i5} = -\overline{M}^T_{5i}$. Now as in the proof of theorem 2.1 we see that for every resistor R_i there exists α_i such that $\partial\overline{M}/\partial\alpha_i$ has rank-1 with respect to α_i.

Now if the column $\overline{M}_{i5}$ is dependent on α_i then the row $\overline{M}_{5i}$ cannot be as the $\overline{M}_{44}$, $\overline{M}_{55}$, $\overline{M}_{45}$ and $\overline{M}_{54}$ elements are zero. Moreover $\overline{M}_{i5} = -\overline{M}^T_{5i}$. Thus the fourth and fifth rows and columns are independent of all α_i.

Moreover, from [2] we have that the following state space realization is possible, where $x^T = [U^T_{c_1}, U^T_{L_1}]$

$$\dot{x} = D_1^{-1}D_2x + D_1^{-1}(D_3-D_2D_1^{-1}D_4D_5)U$$

$$y = (D_6+D_5^TD_4^TD_1^{-1}D_2)x+[D_7-D_6D_4^{-1}D_3 + D_5^TD_4^TD_1^{-1}(D_3-D_2D_1^{-1}D_4D_5)]U + D_5^T(D_8-D_4^TD_1^{-1}D_4)D_5\dot{U} \tag{2.13}$$

Here

$$D_1 = \begin{bmatrix} \Lambda_{c_1} & 0 \\ 0 & \Lambda_{L_1} \end{bmatrix} + \begin{bmatrix} \overline{M}_{24} & \overline{M}_{25} \\ \overline{M}_{34} & \overline{M}_{35} \end{bmatrix} \begin{bmatrix} \Lambda_{c_2} & 0 \\ 0 & \Lambda_{L_2} \end{bmatrix} \begin{bmatrix} \overline{M}_{24} & \overline{M}_{25} \\ \overline{M}_{34} & \overline{M}_{35} \end{bmatrix}^T$$

with Λ_{c_1}, Λ_{c_2}, Λ_{L_1}, Λ_{L_2} obviously defined;

$$D_2 = -\begin{bmatrix} \overline{M}_{22} & \overline{M}_{23} \\ \overline{M}_{23} & \overline{M}_{33} \end{bmatrix}$$

$$D_6 = \begin{bmatrix} \overline{M}_{12} & \overline{M}_{13} \end{bmatrix}$$

$$D_7 = \overline{M}_{11}$$

and $D_8 - D_4^T D_1^{-1} D_4$ positive definite symmetric. Thus if the system is proper $D_5^T(D_8 - D_4^T D_1^{-1} D_4)D_5$ and hence D_5 must be zero. Thus (2.13) can be rewritten (see [2]) as

$$\dot{x} = D^{-1} D_2 x + D_1^{-1} D_3 U$$
$$y = D_6 x + D_7 U \tag{2.14}$$

Clearly, as $\begin{bmatrix} \overline{M}_{24} & \overline{M}_{25} \\ \overline{M}_{34} & \overline{M}_{35} \end{bmatrix}$ is independent of α_i

the state-variable realization in (2.14) has a rank-1 dependence on all α_i . Consider any capacitor or inductor in set 1, i.e. in the set whose elements are represented in

$$\begin{bmatrix} \Lambda_{c_1} & 0 \\ 0 & \Lambda_{L_1} \end{bmatrix}$$

Let this be k_i . Then

$$D_1 = D_{11} + k_i e_j e_j^T \qquad \text{for some } j$$

where e_j is the unit vector with unity in the j^{th} element, and $\overline{D}_{11}$ is independent of k_i .

Thus

$$D_1^{-1} = D_{11}^{-1} - \frac{k_i D_{11} e_j e_j^T D_{11}^{-1}}{1+k_i e_j^T D_{11} e_j}$$

so that with $\alpha = \dfrac{- k_i}{1+k_i e_j^T D_{11} e_j}$

the rank-1 property is satisfied.

Similarly for an element $\bar{k}_i$ in set 2

$$D_1 = D_{12} + \bar{k}_i \bar{x}\,\bar{x}^T$$

where $\bar{x}$ and D_{12} are independent of $\bar{k}_i$ and $\bar{x}$ is a vector.

Thus

$$D_1^{-1} = D_{12}^{-1} - \frac{\bar{k}_i D_{12}^{-1} \bar{x}\,\bar{x}^T D_{12}^{-1}}{1+\bar{k}_i \bar{x}^T D_{12} \bar{x}}$$

whence with $\alpha = \dfrac{- \bar{k}_i}{1+\bar{k}_i \bar{x}^T D_{12} \bar{x}}$

the rank-1 property holds. ∇∇∇

The above results were derived purely from the standpoint of electric circuits. However, as we have already stated the extension to mechanical and chemical analogues are immediate. Indeed, presented below is a non-electrical example where the state variable realization has a rank-1 dependence on most element values.

Given are the dynamics pertinent to the attitude control of the communications technology satellite, Hermes [3].

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

$$Y = Cx$$

where

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\omega_0 h/I_1 & 0 & 0 & \omega_0 - h/I_2 \\ 0 & 0 & 0 & 1 \\ 0 & h/I_2 - \omega_0 & -\omega_0 h/I_2 & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 & 0 \\ F_1 L_1 G_1 \mathrm{Cos}\alpha / I_1 & 0 \\ 0 & 0 \\ -F_1 L_1 G_1 \mathrm{Sin}\alpha / I_2 & F_2 L_2 G_2 / I_2 \end{bmatrix}$$

$$X = \begin{bmatrix} \phi \\ \dot{\phi} \\ \psi \\ \dot{\psi} \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

Here ϕ is the roll, ψ the yaw, I_1 the moment of inertia about the roll axis, I_2, that about the yaw axis, ω_0 the orbital rate, h the nominal wheel angular momentum, α the offset angle, F_1 and F_2 the offset and yaw thruster levels respectively, L_1 and L_2 the offset and yaw thruster moment arms respectively and G_1 and G_2 the impulse bit factors. The inputs u_1 and u_2 provide a guide for the level of consumed fuel.

It is evident that the parameters $I_1, I_2, F_1, F_2, L_1, L_2, G_1$ and G_2 appear in the state variable representation in a rank-1 fashion. Although

$$\frac{\partial}{\partial\alpha}\left[\begin{array}{c|c} sI-A & B \\ \hline -C & o \end{array}\right]$$

has rank-1, the system matrix is obviously not multilinear in α. Thus α does not quite conform to the definition of rank-1 dependence. The parameters ω_0 and h, on the other hand, clearly do not appear in a rank-1 fashion. But, by definition (one is the orbital rate and the other the wheel angular momentum) one can see that they must allow cross-coupling between energy storage elements. They thus fall in the same category as mutual inductors or gyrators, which as we have emphasized do not appear in state variable realizations in a rank-1 fashion.

2.2 Transfer functions for SISO systems

In this section we show that rank-1 SISO systems have minimal transfer function descriptions typified by (2.1). To begin our analysis we consider first the manner in which a single parameter appearing in a rank-1 system affects the transfer function.

At the outset we introduce the following definition of coprimeness of polynomials in more than one variable:

Definition 2.2

Consider $p(X_1,\ldots X_n,X_{n+1},\ldots X_m)$ and $q(X_1,\ldots X_n, X_{n+1},\ldots X_m)$ which are polynomials in the indeterminates $X_1,\ldots X_m$. Then p and q are coprime with respect to the variables $X_1,\ldots X_n$ if there exist no nontrivial f_1 which is a polynomial in $X_1,\ldots X_n$, but rational in $X_{n+1},\ldots X_m$ such that

$$f_1 f_2 = p$$

and

$$f_1 f_3 = q$$

with f_2 and f_3 , polynomials in $X_1,\ldots X_n$ and rationals in $X_{n+1},\ldots X_m$, as well. The extension to the case where the coprimeness of more than two polynomials is in question is obvious.

The following theorem shows that a system having a state variable realization which has a rank-1 dependence on a single parameter α has a transfer function whose numerator and denominator are affine in α .

Theorem 2.3

An n-dimensional SISO system represented by

$$\dot{x} = F(\alpha_1)x + g(\alpha_1)u$$

$$y = h^T(\alpha_1)x + j(\alpha_1)u$$

with a system matrix of the form

$$\left[\begin{array}{c|c} sI-F(\alpha_1) & g(\alpha_1) \\ \hline -h^T(\alpha_1) & j(\alpha_1) \end{array}\right] = \left[\begin{array}{c|c} sI-F_o & g_o \\ \hline -h_o^T & j_o \end{array}\right] + \alpha_1 \left[\begin{array}{c} h_1 \\ \hline h_2 \end{array}\right] \left[\begin{array}{c|c} g_1^T & g_2 \end{array}\right] \quad (2.15)$$

where $F_o, g_o, h_o, j_o, h_1, h_2, g_1$ and g_2 independent of α, has α transfer function

$$W(s) = \frac{a(s) + \alpha_1 b(s)}{c(s) + \alpha_1 d(s)} \quad (2.16)$$

for every α_1. The polynomials $a(s), b(s), c(s)$ and $d(s)$ are independent of α_1 and obey the following restrictions:

(a) $\delta[c(s)] = n$, $\delta[d(s)] < n$, $\delta[a(s)] \leq n$ and $\delta[b(s)] \leq n$

(b) $a(s)d(s) - b(s)c(s)$ is factorizable into two polynomials of degree not exceeding n.

Conversely, any transfer function of the form (2.16) has a state variable realization of the form (2.15) provided that conditions (a) and (b) hold.

Proof:

(i) From equation (2.15)

$$W(s) = [a(s)c(s) + \alpha_1\{a(s)d(s) + g_2h_2c^2(s) - h_2\gamma(s)c(s) + g_2\beta(s)c(s) - \beta(s)\gamma(s)\}]/\{c^2(s) + \alpha_1 c(s)d(s)\} \quad (2.17)$$

where

$$\left.\begin{aligned}
\frac{a(s)}{c(s)} &= h_o^T(sI - F_o)^{-1} g_o + j_o \\
c(s) &= \det(sI - F_o) \\
\frac{d(s)}{c(s)} &= g_1^T(sI - F_o)^{-1} h_1 \\
\frac{\gamma(s)}{c(s)} &= g_1^T(sI - F_o)^{-1} g_o \\
\frac{\beta(s)}{c(s)} &= h_o^T(sI - F_o)^{-1} h_1
\end{aligned}\right\} \qquad (2.18)$$

We note that

$$\begin{aligned}
&a(s)d(s) + g_2h_2c^2(s) - h_2\gamma(s)c(s) + g_2\beta(s)c(s) - \gamma(s)\beta(s) \\
&\quad = a(s)d(s) - (\gamma(s) - g_2c(s))(\beta(s) + h_2c(s)) \qquad (2.19)
\end{aligned}$$

is divisible by $c(s)$ because

$$\overline{W}(s) = \begin{bmatrix} a(s)/c(s) & \beta(s)/c(s) \\ \gamma(s)/c(s) & d(s)/c(s) \end{bmatrix} = \begin{bmatrix} h_o^T \\ g_1^T \end{bmatrix} [sI-F_o]^{-1} [g_o \quad h_1]$$

Thus $\overline{W}(s)$ has an n^{th} order realization and its Macmillan degree is not greater than n . Thus $ad-\beta\gamma$ is divisible by $c(s)$ whence (2.19) is divisible by $c(s)$. Define $b(s)$ by

$$b(s)c(s) = a(s)d(s) - [\gamma(s) - g_2c(s)][\beta(s) + h_2c(s)] \qquad (2.20)$$

Then (2.17) has the same form as (2.16) and by (2.20) and (2.18) conditions (a) and (b) are satisfied.

(ii) Suppose that (a) and (b) hold for (2.16). Let

$$a(s)d(s) - b(s)c(s) = f_1(s)f_2(s) = (\gamma(s) - g_2c(s))(\beta(s) + h_2c(s))$$

with $\delta[\beta(s)]$ and $\delta[\gamma(s)] < n$. Then $a(s)d(s) - f_1(s)f_2(s)$ is divisible by $c(s)$ whence

$$\hat{W}(s) = \begin{bmatrix} a(s)/c(s) & f_2(s)/c(s) \\ f_1(s)/c(s) & d(s)/c(s) \end{bmatrix}$$

has Macmillan degree not greater than n.
Hence $\hat{W}(s)$ can be expressed as

$$\begin{bmatrix} h_o^T \\ g_1^T \end{bmatrix} \begin{bmatrix} sI-F_o \end{bmatrix}^{-1} \begin{bmatrix} g_o & h_1 \end{bmatrix} + \begin{bmatrix} j_o & h_2 \\ -g_2 & 0 \end{bmatrix}$$

so that by reversing the argument in the first part of this theorem a state variable realization of (2.16) exists in the form typified by (2.15). ∇∇∇

<u>Remark</u>:

(2.5) The reverse implication of theorem 2.3 is interesting. It shows that not all transfer functions whose numerator and denominator polynomials are affine in α_1, have state variable descriptions which have rank-1 dependence on α_1. For example if $c(s) = -b(s) = s^3 + 1$

and $a(s)$ and $d(s)$ are such that

$$a(s)d(s) = 6s^4 - 2s^3 + 11s^2 + 5 .$$

Then

$$a(s)d(s) - b(s)c(s) = (s^2+1)(s^2+2)(s^2+3) .$$

Thus there do not exist $f_1(s)$ and $f_2(s)$ of degree less than or equal to 3 , for which

$$a(s)d(s) - b(s)c(s) = f_1(s)f_2(s)$$

is true. Thus

$$W(s) = \frac{a(s) + \alpha_1 b(s)}{c(s) + \alpha_1 d(s)}$$

has no state variable realization which has a rank-1 dependence on α . In general one of the following three conditions obtain:

(i) If $\delta[b(s)] < \delta[c(s)] = n$ then $\delta[a(s)d(s) - b(s)c(s)] \leq 2n-1$ (as $\delta[d(s)] < n$). Thus $a(s)d(s) - b(s)c(s)$ is always expressible as a product of two polynomials of degree less than or equal to n .

(ii) If $\delta[b(s)] = n$ and n is even, then again $f_1(s)$ and $f_2(s)$ will have degree no greater than n .

(iii) However, if $\delta[b(s)] = n$ and n is odd then such $f_1(s)$ and $f_2(s)$ may not be found.

Case II: m is dependent on k_1 . Then m is linear in k_1 and p,q are independent of k_1 . Suppose $m(s,k_1,k_2) = r_1(s,k_2) + k_1 r_2(s,k_2)$.

Then
$$a(s,k_2) = r_1(s,k_2)p(s,k_2)$$
$$b(s,k_2) = r_2(s,k_2)p(s,k_2)$$
$$c(s,k_2) = r_1(s,k_2)q(s,k_2)$$
$$d(s,k_2) = r_2(s,k_2)q(s,k_2)$$

whence $ad-bc \equiv 0$, which too contradicts our hypothesis.

∇∇∇

Remark: (2.6) Violation of (ii) implies that W is independent of k_1 .

Lemma 2.4

Suppose that the transfer function $W(s,k_1,-k_N)$ is expressible as

$$W(s,k_1,\ldots k_N) = \frac{a_i(s,K^{(i)}) + k_i b_i(s,K^{(i)})}{c_i(s,K^{(i)}) + k_i d_i(s,K^{(i)})} \quad \forall\, i \in \{1,\ldots N\}$$

where $K^{(i)} \triangleq \{k_1,\ldots k_N\} - \{k_i\}$ and a_i,b_i,c_i and d_i are polynomials in $s,k_1,\ldots k_N$. Suppose $a_i d_i - b_i c_i \not\equiv 0$ and a_i,b_i,c_i,d_i are coprime with respect to s and $K^{(i)}$.
Then

$$W\,(s,k_1,\ldots k_N) = \frac{P(s,k_1,\ldots k_N)}{Q(s,k_1,\ldots k_N)}$$

where P and Q are multilinear in k_i .

Thus for a strictly proper transfer function like (2.16) there always exists a state-variable realization which has a rank-1 dependence on α_1 .

We now extend the first part of Theorem 2.3 to the case where there are N-parameters.

Lemma 2.3

Consider a transfer function

$$W(s,k_1,k_2) = \frac{a(s,k_2) + k_1 b(s,k_2)}{c(s,k_2) + k_1 d(s,k_2)}$$

where a,b,c,d are polynomials in s and k_2 . Suppose

(i) $a(s,k_2)$, $b(s,k_2)$, $c(s,k_2)$ and $d(s,k_2)$ are coprime with respect to s and k_2 .

(ii) $a(s,k_2)d(s,k_2) - b(s,k_2)c(s,k_2) \not\equiv 0$.

Then $a(s,k_2) + k_1 b(s,k_2)$ and $c(s,k_2) + k_1 d(s,k_2)$ are coprime with respect to s,k_1 and k_2 .

Proof

After [4,p36] we have that the ring of polynomials in the variables $s,k_1 k_2$ over the field of real numbers, is a unique factorization domain.

Let $a(s,k_2) + k_1 b(s,k_2) = m(s,k_1,k_2)p(s,k_1,k_2)$

$c(s,k_2) + k_1 d(s,k_2) = m(s,k_1,k_2)q(s,k_1,k_2)$

with m,p,q are polynomials in s,k_1 and k_2 . Consider the following cases:

Case I: m is independent of k_1 . It is immediate that (i) is violated.

Remark: The proof of the above theorem is not trivial as even though $a_1(s,K^{(1)}) + k_1b_1(s,K^{(1)})$ and $a_2(s,K^{(2)}) + k_2b_2(s,K^{(2)})$, for example, are respectively affine in k_1 and k_2, it is not clear that a_1 and b_1 are affine in k_2 or for that matter, that a_2 and b_2 are affine in k_1.

Proof

We shall prove the case when $N=2$. The more general case follows along the same lines. Suppose

$$W(s,k_1,k_2) = \frac{a_1(s,k_2) + k_1b_1(s,k_2)}{c_1(s,k_2) + k_1d_1(s,k_2)}$$

$$= \frac{a_2(s,k_1) + k_2b_2(s,k_1)}{c_2(s,k_1) + k_2d_2(s,k_1)}$$

and suppose that the other hypotheses specialized to $N=2$, hold. Then

$$[a_1(s,k_2) + k_1b_1(s,k_2)][c_2(s,k_1) + k_2d_2(s,k_1)]$$

$$= [a_2(s,k_1) + k_2b_2(s,k_1)][c_1(s,k_2) + k_1d_1(s,k_2)] .$$

By lemma 2.3 $a_1(s,k_2)+k_1b_1(s,k_2)$ and $c_1(s,k_2)+k_1d_1(s,k_2)$ are coprime with respect to s,k_1 and k_2. Thus $a_1(s,k_2)+k_1b_1(s,k_2)$ divides $a_2(s,k_1)+k_2b_2(s,k_1)$. Thus $a_1(s,k_2)$ and $b_1(s,k_2)$ can be at most linear in k_2. Similarly $c_1(s,k_2)$ and $d_1(s,k_2)$ can be at most linear in k_2. Hence the transfer function $W(s,k_1,\ldots k_N)$ can be written as

$$\frac{p_o(s) + k_1p_1(s) + k_2p_2(s) + k_1k_2p_{12}(s)}{q_o(s) + k_1q_1(s) + k_2q_2(s) + k_1k_2l_{12}(s)}$$

Arguing along similar lines, the result follows. ∇∇∇

We are interested in minimality with respect to s alone. That is, we wish to show that if the hypothesis of Lemma 2.4 holds then P and Q have no common factors which are <u>polynomial</u> in s but <u>rational</u> in the k_i. The following lemma which follows from [5], shows that this is indeed the case.

<u>Lemma 2.5</u>

If $P(s,k_1,k_2,...k_N)$ and $Q(s,k_1,...k_N)$, polynomials in $s,k_1,...k_N$, have no common factors which are polynomials in $s,k_1,...k_N$, then they have no common factors which are polynomials in s but rational in $k_1,...k_N$.

<u>Proof</u>

If P and Q have no common factors which are polynomials in $s,k_1,...k_N$ then P and Q are minor coprime as well (see [5] for a definition). Thus by [5] there exist polynomials x,y and ψ, with ψ nontrivial and independent of s for which

$$P(s,k_1,...k_N)x(s,k_1,...k_N) + Q(s,k_1,...k_N)y(s,k_1,...k_N) = \psi(k_1,...k_N) .$$

Thus dividing both sides by ψ the result is immediate.

∇∇∇

Thus theorem 2.3 and lemmata 2.3 - 2.5 together yield the following main result of this chapter.

Theorem 2.4

If the state variable realization of a linear time invariant finite dimensional system has a rank-1 dependence on N parameters $k_1,\ldots k_N$, then it has a minimal transfer function description whose numerator and denominator polynomials are multilinear in the k_i .

Remark (2.7) From (a) of theorem 2.3 it follows that one can find a transformation in k_i (viz by replacing by α_i of definition 2.1) to make the denominator polynomial with coefficients independent of $k_1,\ldots k_N$, have a higher degree than all other polynomials in the denominator. In other words in (2.1) q_o has a higher degree than all other q_r .

(2.8) Suppose we have a transfer function

$$W(s,k_1,\ldots k_N) = \frac{P(s,k_1,\ldots k_N)}{Q(s,k_1,\ldots k_N)}$$

where P and Q are coprime in s and $k_1,\ldots k_N$ and are multilinear in the k_i . Suppose also that

$$W(s,k_1,\ldots k_N) = \frac{a_i(s,k^{(i)}) + k_i b_i(s,k^{(i)})}{c_i(s,k^{(i)}) + k_i d_i(s,k^{(i)})} \qquad \forall\ i \in \{1,\ldots N\}$$

and that the degree of Q with respect to s is n . Here $k^{(i)} = [k_1 \ldots k_{i-1}, k_{i+1}, \ldots k_N]^T$.

Then by Theorem 2.3 a necessary condition for a state variable representation having a rank-1 dependence on all of the k_i is that $\exists$ $f_{1i}(s,k^{(i)})$ and $f_{2i}(s,k^{(i)})$ such that

$$a_i d_i - b_i c_i = f_{1i} f_{2i} \quad \forall\ i \in \{1,\ldots N\}$$

and the degrees of f_{1i} and f_{2i} with respect to s do not exceed n . It is not clear however that this is also a sufficient condition. This is because while there exists at least one state variable realization which is rank-1 with respect to a given k_i it is not obvious that all of these need necessarily be the same.

(2.9) The results of this and the previous section show that all RLC circuits have minimal transfer function descriptions which have both denominator and numerator polynomials multilinear in the circuit element values.

2.3 A Multiinput Multioutput extension of the transfer function result:

In this section we present a multi-input, multi-output (MIMO) extension of theorem 2.4 by showing that there exists a transfer function description of a rank-1 system, whose numerator and denominator are multilinear in the k_i . The particular representation used here will be referred to as a quasi-minimal representation. It is defined as follows.

Definition 2.3

Consider a MIMO system having a rational transfer function matrix $W(s,k_1,\ldots k_N)$ with

$$W_{ij}(s,k_1,\ldots k_N) = a_{ij}(s,k_1,\ldots k_N)/b_{ij}(s,k_1,\ldots k_N)$$

with a_{ij}, b_{ij} coprime in s and $k_1,\ldots k_N$. Then the transfer function description $N(s,k_1,\ldots k_N)/d(s,k_1,\ldots k_N)$, d scalar, is <u>quasi-minimal</u> if d is the l.c.m. of b_{ij}.

<u>Remark</u>: (2.10) It is not difficult to see from [4,p36] that the numerator and denominator of the quasi minimal transfer function matrix are unique to within a scalar constant.

(2.11) By lemma 2.5 a_{ij} and b_{ij} must be coprime in s alone as well.

We then have the following result.

<u>Theorem 2.5</u>

Consider a system having a state variable description which is rank-1 with respect to N parameters k_i. Let α_i be the corresponding parameters defined in Definition 2.1. Then the quasi minimal tranfer function matrix monic in s, is given by

$$W(s,\alpha) = \frac{A_o(s)+\alpha_1 A_1(s)+\ldots\ldots+\alpha_1\alpha_2 A_{12}(s)+\ldots\ldots+\alpha_1\ldots\alpha_N A_{12-N}(s)}{b_o(s)+\alpha_1 b_1(s)+\ldots\ldots\ldots\ldots+\ldots\ldots\ldots\ldots+\alpha_1\ldots\alpha_N b_{12-N}(s)} \tag{2.21}$$

where $\alpha \triangleq [\alpha_1,\ldots\alpha_N]^T$

<u>Proof</u>

By mimicking the proof of theorem 2.3 one can see that the transfer function $W(s,\alpha)$ can be written as

$$W(s,\alpha) = \frac{M_i(s,\alpha^{(i)}) + \alpha_i N_i(s,\alpha^{(i)})}{c_i(s,\alpha^{(i)}) + \alpha_i d_i(s,\alpha^{(i)})} \tag{2.22}$$

where

$$\alpha^{(i)} \triangleq [\alpha_1,\ldots\alpha_{i-1},\alpha_{i+1},\ldots\alpha_N]^T$$

and $\delta_s[c_i(s,\alpha^{(i)})] > \delta_s[d_i(s,\alpha^{(i)})]$ (δ_s refers to degree with respect to s).

Suppose $A(s,\alpha)/b(s,\alpha)$ is the quasi minimal representation of $W(s,\alpha)$. Then

$$b(s,\alpha) \mid \{c_i(s,\alpha^{(i)}) + \alpha_i d_i(s,\alpha^{(i)})\}$$

and one can show that $b(s,\alpha)$ is multilinear in the α_i . Let $p(s,\alpha)$ be such that

$$p(s,\alpha)b(s,\alpha) = c_i(s,\alpha^{(i)}) + \alpha_i d_i(s,\alpha^{(i)}) \text{ .}$$

Then $p(s,\alpha)$ divides every element of $M_i + \alpha_i N_i$. It is then easy to show that $A(s,\alpha)$ is multilinear in the α_i . Here $\delta[b_o(s)] > \delta[b_r(s)]$ for all $r \neq 0$.

Remark (2.12) There clearly also exists an equavalent representation of the quasi-minimal tranfer function which is multilinear with respect to the k_i . However, the property that $\delta[b_o(s)] > \delta[b_r(s)]$ will no longer hold. As this proves useful in proving results in later chapters, henceforth when discussing transfer function descriptions we shall assume that k_i's refer to their transformed versions i.e. to erstwhile α_i's .

Assumption 2.2

Consider a system with N unknown parameters $k_1 \dots k_N$. Then all the parameters can be so extracted as to yield equivalent block diagram representations of the form in fig. 2.3. Sometimes in fig. 2.3, k_i may need to be replaced by $1/k_i$.

We thus have the following lemma and theorem.

Lemma 2.7

Consider a linear time invariant, finite dimensional system with N unknown parameters $k_1, \dots k_N$. Suppose assumption 2.2 holds. Denote $k = [k_1, \dots k_N]^T$ and $k^{(i)} = [k_1, \dots k_{i-1}, k_{i+1}, \dots k_N]^T$. Then for all $i \in \{1, \dots N\}$ there exist matrices $A_i(s,k^{(i)})$, vectors $h_{1i}(s,k^{(i)})$ and $h_{2i}(s,k^{(i)})$ and scalars $b_i(s,k^{(i)})$, $c_i(s,k^{(i)})$ and $d_i(s,k^{(i)})$ all polynomials in s and $\{k_1, \dots k_{i-1}, k_{i+1}, \dots k_N\}$ and independent of k_i such that the transfer function $W(s,k)$ can be expressed as

$$W(s,k) = \frac{A_i(s,k^{(i)})}{b_i(s,k^{(i)})} + \frac{k_i h_{1i}(s,k^{(i)}) h_{2i}^T(s,k^{(i)})}{c_i(s,k^{(i)}) + k_i d_i(s,k^{(i)})} \qquad (2.23)$$

Proof

Proceeds as in lemma 2.1.

∇∇∇

Theorem 2.7

Consider a system which satisfies the hypotheses of lemma 2.7. Then the quasi minimal transfer function has the form of (2.21) with the k_i replacing α_i.

Proof

Proceeds as in theorem 2.6, using lemma 2.7.

∇∇∇

Remark

(2.13) Equation (2.23) acts as a guide to the question of the existence of state variable realizations having a rank-1 dependence on all the k_i , whenever (2.23) is satisfied. It would appear from this that given a transfer function of the form

$$\frac{W_{11} + kW_{12}}{w_{21} + kw_{22}}$$

where W_{11} and W_{12} are matrix polynomials and w_{21} and w_{22} are scalar polynomials, a rank-1 state variable realization with respect to k exists if

$$W_{11}w_{22} - W_{12}w_{21} = h_1 h_2^T$$

where h_1 and h_2 are vectors having elements of degree no greater than $\delta[w_{21}]$. As in the SISO case, the extension to the multiparameter problem is nontrivial.

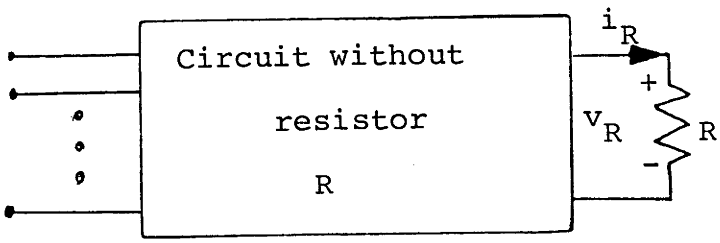


Figure 2.1 A circuit with a resistor.

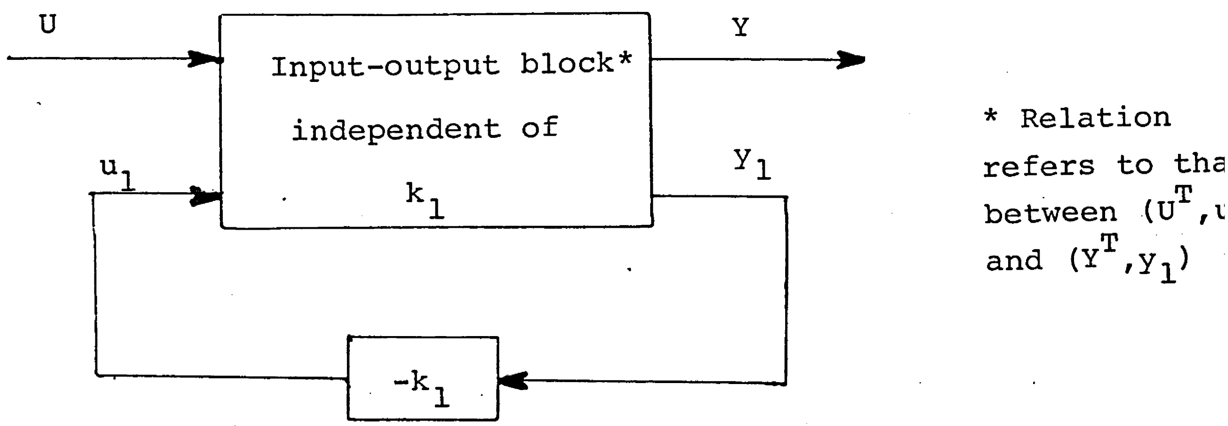


Figure 2.2 A common input-output description for a circuit with a physical element of value k_1.

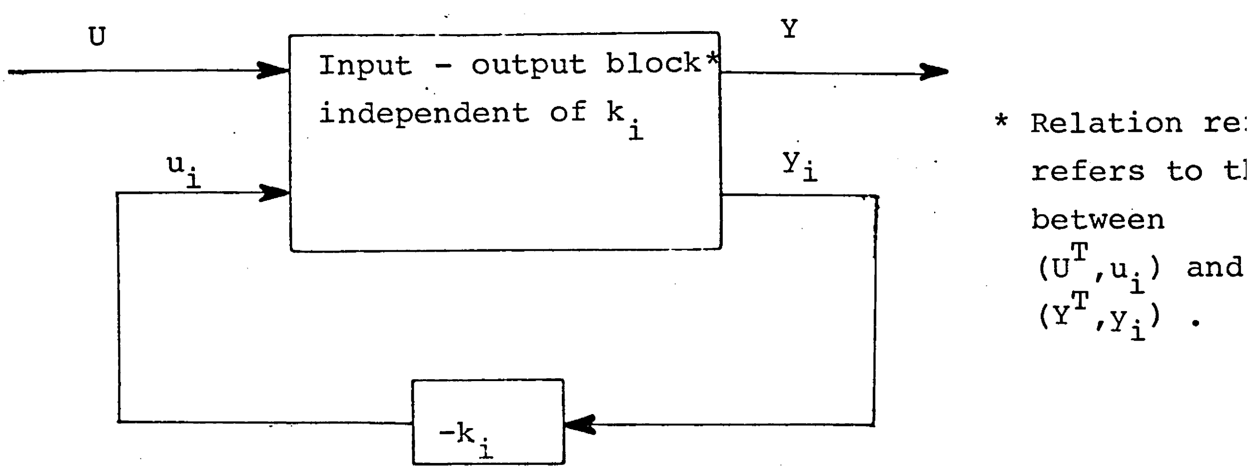


Figure 2.3 Common input-output descriptions for circuits with unknown physical parameters k_i.

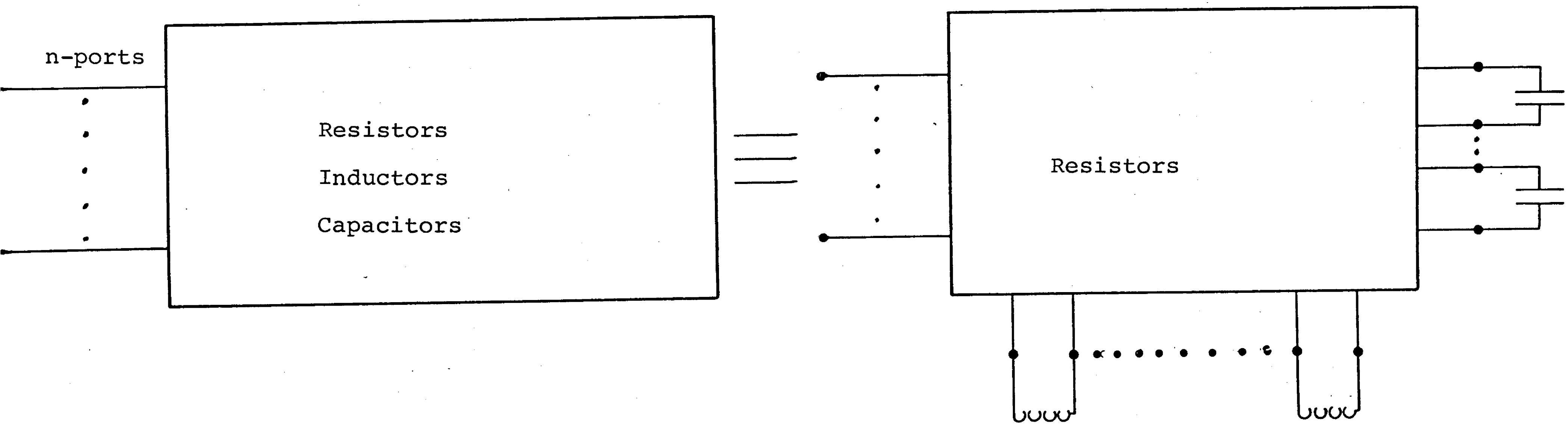
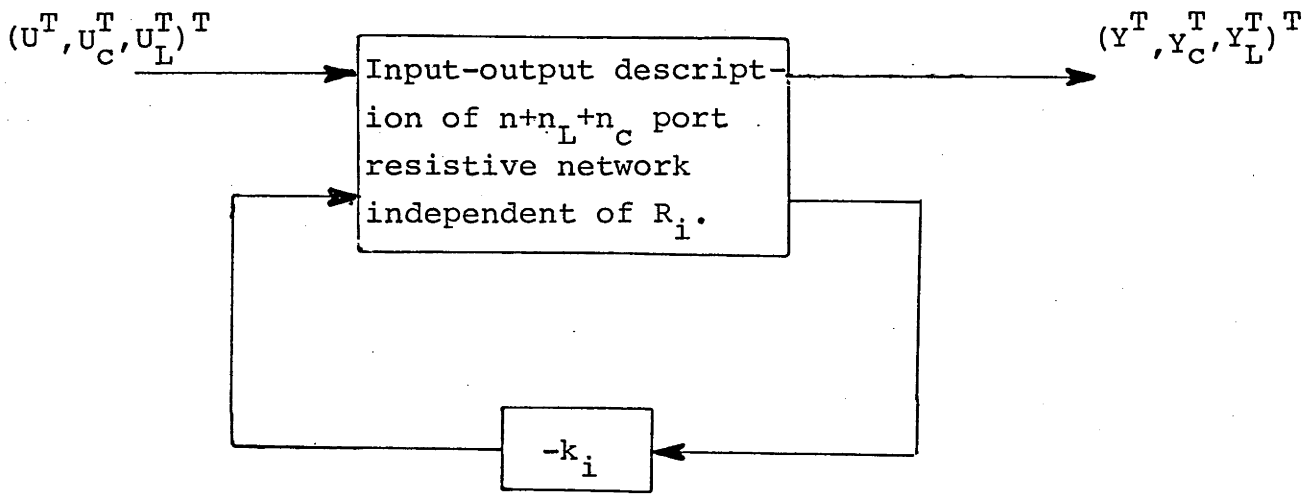


Figure 2.4 Reactance extraction for an n-port circuit.



n_L = number of inductors in the circuit of fig 2.4

n_C = number of capacitors in the circuit of fig 2.4

Figure 2.5 Equivalent input-output description of the network in fig. 2.4

References for Chapter 2

1. B.D.O. Anderson and S. Dasgupta,"Physically based parametrisation for designing adaptive algorithms", Under preparation.
2. B.D.O. Anderson and S. Vongpanitlerd, Network analysis and synthesis, a modern systems theory approach,Prentice Hall, Electrical Engineering Series, 1973.
3. C. Diduch and R. Balasubramanian, "Integrated attitude control of a satellite using a microprocessor", Proceedings IFAC symposium on theory and application of digital control, pp 9-14 New Delhi, Jan 1982.
4. W.V.D. Hodge and D. Pedoe, Methods of Algebraic Geometry, Cambridge University Press, 1953.
5. D.C. Youla and G. Gnavi, "Notes on n-dimensional systems theory" submitted for publication.

§3. PERSISTENCE OF EXCITATION

The purpose of this chapter is to develop a set of general tools for establishing persistence of excitation (p.e.) conditions, guaranteeing the uniform asymptotic (exponential for linear algorithms) convergence of a general class of adaptive identification and control algorithms. The tools developed here will be used in later chapters to establish the uniform asymptotic convergence of algorithms proposed for the special parameterizations considered in this thesis.

Uniform asymptotic convergence, we believe, is important for robustness. Adaptive algorithms without such convergence have been known to display unacceptable behaviour in the presence of modelling inadequacies and noise [1,2]. Uniformly asymptotically convergent algorithms on the other hand are totally stable [3, pp. 107-108], a property which allows them to retain stability in face of modest departures from ideality.

Most results for the exponential convergence of continuous time algorithms [4-10] require the uniform positive definiteness of gramians of certain regression vectors. Unfortunately, apart from the inputs, these vectors involve the system outputs as well, so that the conditions in question are not useful for *a priori* input design. The tools developed in this chapter are aimed, primarily, at translating these conditions to ones involving the system inputs only. Earlier Yuan and Wonham [11] had developed some techniques for input design.

These, however, were based primarily on conditions similar to those established in [4-10] and did not involve conditions independent of system outputs.

The first contribution of this chapter (section 3.1) is to use the criterion given in [10] as the basis for deriving a convergence condition on input signals alone for the identification of a stable plant. This is similar in statement, but not in proof, to a condition derived in [12] for discrete time plants and for inputs which are linear combinations of sinusoids. The proofs of this chapter use a possibly little known but potentially very useful inequality relating derivatives of functions of one variable (See Lemma (3.A.1) in Appendix 3.A).

In section 3.2 the result of section 3.1 is extended to consider plants which are not necessarily stable but are stabilized by feedback. Section 3.3 considers plants with possibly unbounded signals while section 3.4 considers slowly time varying plants with bounded signals. While the results in sections 3.2 and 3.3 have no discrete time parallels, the corresponding result for slowly time varying plants has been presented in [13]. As with the results of section 3.1, however, the technical artifacts used in the continuous and discrete time cases are substantially different. In [14,15] Boyd and Sastry have independently derived input only conditions using generalized harmonic analysis. Their results are restricted, however, to stable, linear time-invariant systems and thus do not encapsulate the results in sections 3.2-3.4.

In sections 3.5 and 3.6 we consider the exponential stability of a model reference adaptive control (MRAC) algorithm studied by Morse [16], by examining in turn the situations where (a) the constant gain of the process transfer function is known and (b) where no such information is available *a priori*. In the former case exponential stability is shown to be conditional on the satisfaction of a p.e. condition on the reference input while in the latter such stability is shown as impossible to achieve; even with a p.e. condition. The negative result uses a stability lemma which is similar to a discrete time result derived in [17].

An MRAC algorithm, similar to [16] and proposed in [18], has been analysed in [14], with conclusions which parallel the ones derived in sections 3.5 and 3.6.

As some of the technical proofs of this chapter tend to be involved without always being very illuminating, they have been relegated to appendices at the end of this chapter. Most of the work contained here is available in the papers [19-20].

In this chapter the exponential convergence of adaptive controllers refers to the rates of convergence of the controller parameters and *not* to any characteristic of the controller itself. In fact the controller in the limit may well approach a control law which is not exponentially stabilizing, though the rate of convergence of the controller itself, to this law must be exponential.

3.1 Identification of Stable Plants:

Consider an n-dimensional asymptotically stable, single-input single-output (SISO) system

$$y^{(n)} + \sum_{i=0}^{n-1} a_i y^{(i)} = \sum_{j=0}^{m} b_j u^{(j)} \quad (m \leqslant n) \quad . \tag{3.1}$$

Define

$$V(t) = [y_{n-1}(t), \ldots, y_0(t), u_m(t), \ldots, u_0(t)]^T \tag{3.1a}$$

with

$$y_i(t) = s^i y(t)/(s+\alpha)^n \text{ and } u_i(t) = s^i u(t)/(s+\alpha)^n .$$

Throughout this thesis notation shall be abused by referring, for example, to $y(s)$ as the Laplace transform of $y(t)$. By the same token

$$y_i(t) = s^i y(t)/(s+\alpha)^n$$

will denote the solution of the differential equation

$$(p+\alpha)^n y_i(t) = p^i y(t)$$

with p defined as the differential operator d/dt and unless otherwise mentioned, with all initial conditions set to zero.

In the sequel, vector time functions such as $V(t)$ will be required to belong to a special set $\Omega_\Delta[0,\infty)$, defined below.

The importance of this set in identification problems was first illustrated in [11].

Definition 3.1:

C_Δ is a set $\{t_i\}$ of points in $[0,\infty)$ for which there exists a Δ such that for any $t_i, t_j \varepsilon C_\Delta$ with $t_i \neq t_j$, one has $|t_i - t_j| \geqslant \Delta$, (i.e. C_Δ comprises points spaced at least Δ apart).

Definition 3.2:

A function $f(\cdot)$ belongs to Ω_Δ if there corresponds some Δ and C_Δ such that

(1) $f(t)$ and $\dot{f}(t)$ are continuous on $\{[0,\infty)-C_\Delta\}$,

(2) there exist constants M_1 and M_2 such that $|f(t)| < M_1 \forall t \varepsilon [0,\infty)$ and $\|\dot{f}(t)\| < M_2 \forall t \varepsilon \{[0,\infty)-C_\Delta\}$.

(3) $\dot{f}(t)$ has finite limits as $t \downarrow t_i$ and $t \uparrow t_i$, for each $t_i \varepsilon C_\Delta$.

In other words functions in Ω_Δ are smooth enough to have bounded continuous derivatives, except that a countable number of switchings are allowed and these do not occur "too" frequently. Note, that conditions (1) and (2) also imply that $f(t)$ has finite one-sided limits everywhere. Anderson's result derived in [10] is now restated. Following the restatement we shall explain the role of $V(t)$ in identification.

Theorem 3.1: Suppose $U(\cdot): R_+ \to R^n$ be such that $U(\cdot) \varepsilon \Omega_\Delta [0,\infty)$. The differential equations listed below are exponentially stable iff there exist positive δ, α_1

and α_2 such that for all $s \in R_+$

$$\alpha_1 I \leq \int_s^{s+\delta} U(t)U^T(t)\,dt \leq \alpha_2 I \tag{3.2}$$

$$\text{(a)} \quad \dot{x} = -UU^T x \tag{3.3}$$

$$\text{(b)} \quad \dot{x} = \begin{bmatrix} 0 & -UB^T \\ BU^T & A \end{bmatrix} x \tag{3.4}$$

$$\text{(c)} \quad \dot{x} = \begin{bmatrix} -D \otimes UU^T & -C^T \otimes U \\ B \otimes U^T & A \end{bmatrix} x \tag{3.5}$$

In (3.4) A is a real constant $n \times n$ matrix, with $A + A^T = -I$ and B is a real constant nonzero n-dimensional vector. In (3.5) $\{A,B,C,D\}$ is a quadruple of constant matrices defining a minimal realization of a transfer function matrix $Z(s) = D + C^T (sI-A)^{-1} B$ with $Z(s-\sigma)$ positive real for some $\sigma > 0$, nonsingular almost everywhere and with $D = D^T$.

Remarks

(3.1) The satisfaction of (3.2) for some $\delta = \delta_0$ implies the same for all $\delta > \delta_0$

(3.2) It can be seen [9,10] that if U is replaced by V, defined in (3.1a) then the differential equations (3.3-3.5) represent most of the better known identification schemes. Thus if $V \in \Omega_\Delta[0,\infty)$ then (3.2) is a necessary and sufficient condition for the exponential convergence of most identification schemes. The vector V is a regression vector required by (3.2) to be persistently spanning (i.e. $V(\cdot)$ should span the entire R^n space over time). The point of this chapter is to assert that this persistently spanning condition is synonymous with the requirement that the inputs to the system be "sufficiently rich".

(3.3) It has been shown in [21] that for the differential equation (3.3)

$$\frac{\|x(t+\delta)\|^2 - \|x(t)\|^2}{\|x(t)\|^2} \leqslant - \frac{2\alpha_1}{(1 + \frac{1}{\sqrt{2}}\alpha_2)^2}$$

Thus the convergence rate of $x(t)$ increases with increasing α_1, and decreasing α_2. Equally, a smaller δ results in faster convergence.

We now state the main result of this section.

Theorem 3.2 Consider a strictly stable n-dimensional SISO system with

$$y^{(n)} + \sum_{i=0}^{n-1} a_i y^{(i)} = \sum_{j=0}^{m} b_j u^{(j)}$$

$m \leqslant n$, the polynomials $A(s) = s^n + \sum_{i=0}^{n-1} a_i s^i$ and $B(s) = \sum_{j=0}^{m} b_j s^j$ coprime and $u(t) \in \Omega_\Delta[0,\infty)$. Define $V(t)$, $W(t)$ as

$$V \triangleq [y_{n-1},\ldots,y_0,\ u_m,\ldots,u_0]^T$$

$$W \triangleq [u, \frac{u}{s+\beta}, \ldots, \frac{u}{(s+\beta)^{n+m}}]^T$$

where $y_i = s^i y/(s+\alpha)^n$ and $u_i = s^i u/(s+\alpha)^n$, $\beta, \alpha > 0$. Suppose there exist some positive α_1, α_2 and δ', such that

$$\alpha_1 I \leqslant \int_{\sigma}^{\sigma+\delta'} W(t)W^T(t)dt \leqslant \alpha_2 I \qquad (3.6)$$

$\forall \sigma \varepsilon R_+$. Then there exist positive α_3, α_4 and a suitably large $\delta > \delta'$ also independent of σ such that $\forall \sigma \varepsilon R_+$

$$\alpha_3 I \leqslant \int_{\sigma}^{\sigma+\delta} V(t)V^T(t)dt \leqslant \alpha_4 I \qquad (3.7)$$

Conversely, if there exist some positive α_5, α_6 and δ'' independent of σ such that

$$\alpha_5 I \leqslant \int_{\sigma}^{\sigma+\delta''} V(t)V^T(t)dt \leqslant \alpha_6 I \qquad (3.8)$$

for all $\sigma \varepsilon R_+$, then there exist positive α_7, α_8 and a suitably large $\delta > \delta''$ independent of σ such that

$$\alpha_7 I \leqslant \int_{\sigma}^{\sigma+\delta} W(t)W^T(t)dt \leqslant \alpha_8 I \qquad (3.9)$$

Proof

While the detailed proof has been given in appendix 3.A, here we sketch a brief outline. That the upper bounds of (3.6) and (3.8) imply those of (3.7) and (3.9), respectively, follows from the boundedness and stability assumptions. As far as the lower bounds are concerned, a contradiction argument is used. It is shown that the violation of the lower bound in (3.7) implies the same for that in (3.6), whence (3.6) must imply (3.7) (the proof relating (3.8) to (3.9) proceeds likewise).

The proof uses the Lemmata 3.A.1 (taken from [22]) and 3.A.2 given in appendix 3.A. Briefly, Lemma 3.A.1 states that if a function and its nth derivative are respectively "small" and bounded over a given interval then the first n-1 derivatives must also be "small" over the same interval.

Lemma 3.A.2 states that if the input to an asymptotically stable system is "small" over an interval then so must be its output over a smaller interval. This reduction in the size of the interval arises because of the need for effects of the boundary conditions at the start of the first interval to decay.

Using Lemma 3.A.1 one can show that if the lower bound in (3.7) is arbitrarily small for some σ then there exists a constant unit vector θ such that $|\theta^T V(t)|$ is also arbitrarily small over the interval $[\sigma,\sigma+\delta]$. Then by using $\theta^T V(t)$ as the input to an appropriately selected stable system, applying Lemma 3.A.2 and the coprimeness of $A(s)$ and $B(s)$, it follows that the lower bound in (3.6) will be arbitrarily small. Similar arguments are employed in proving the implication from (3.8) to (3.9).

Remarks

(3.4) Theorem 3.2 in conjunction with theorem 3.1 and the remarks thereafter demonstrates that the input-only conditions are necessary and sufficient for the exponential convergence of identification schemes typified by (3.3-3.5).

(3.5) As can be seen from the proof in Appendix 3.A, the reduction in the intervals of (3.6) and (3.8) stems from the need for transients associated with nonzero boundary conditions at $t = \sigma$ to die away. The extents of the reductions are functions of the α_i. This remark shall apply to all other such results of this chapter.

(3.6) The system (3.1) is described by a total of $m + n + 1$ parameters once the integers m and n have been specified. Application of a single sinusoid to a stable linear system allows identification of the real and imaginary parts of the transfer function at one frequency. More generally, application of

$$u(t) = \sum_{i=1}^{q} u_i e^{j\omega_i t} \tag{3.10}$$

where the ω_i are real, $\omega_i \neq \omega_j$ for $i \neq j$, $u_i \neq 0$ and $u(t)$ real, allows identification of q real pieces of information about the transfer function as there arise q independent equations of the system parameters. The following result is therefore no surprise and is applicable to the identification problem when $p = m + n + 1$.

Corollary 3.1

Let $\beta > 0$ and define

$$W = [u \; \frac{u}{s+\beta} \; , \ldots , \; \frac{u}{(s+\beta)^{p-1}}]^T \tag{3.11}$$

Suppose that $u(\cdot)$ satisfies (3.10) and the associated conditions. Then

$$\alpha_9 I \leqslant \int_{\sigma}^{\sigma+\delta} W(t)W^T(t)dt < \alpha_{10} I$$

$\forall \sigma \geqslant 0$, $\delta \geqslant \delta_0 > 0$, with α_9, α_{10}, depending on δ, u, ω_i for $i = 1,\ldots,q$ iff $p \leqslant q$.

Proof

Since the derivatives of $u(t)$ are now all continuous and bounded, one can show by an argument similar to that used in proving in theorem 3.2 that the lower boundedness condition on $\int_{\sigma}^{\sigma+\delta} W(t)W^T(t)dt$ in (3.11) is equivalent to a similar lower boundedness condition on $\int_{\sigma}^{\sigma+\bar{\delta}} \bar{W}(t)\bar{W}^T(t)dt$ where

$$\bar{W}(t) \triangleq [u, su, \ldots, s^{p-1}u]^T$$

and $\bar{\delta}$ is independent of σ. Moreover, for the $u(t)$ in question the integral $\int_{\sigma}^{\sigma+\delta} \bar{W}(t)\bar{W}^T(t)dt$ is bounded below, away from zero if and only if $q \geqslant p$. This completes the proof of the corollary.

3.2 Identification of Unstable Plants Inside an Overall Stable System

In this section, we shall consider the plants described by (3.1), but with the stability assumption removed. Instead, the following assumption is in force throughout this section.

Assumption 3.1: The plant is part of an overall system which is stable, and in which all signals are bounded.

This overall stability assumption is by no means unreasonable. In many adaptive control algorithms (eg. [23]) boundedness of signals results even without p.e.

Theorem 3.2 Assume the plant is described by (3.1), with

$$A(s) = s^n + \sum_{i=0}^{n-1} a_i s^i \quad \text{and} \quad B(s) = \sum_{j=0}^{m} b_j s^j \quad \text{coprime,}$$

$u(t) \in \Omega_\Delta[0,\infty)$, and

$$V(t) \triangleq [y_{n-1},\ldots,y_0 \quad u_m,\ldots,u_0]^T \tag{3.12}$$

$$W(t) = [u \; \frac{u}{(s+\beta)} \;,\ldots,\; \frac{u}{(s+\beta)^{n+m}}]^T \tag{3.13}$$

where $y_i = s^i y/(s+\alpha)^n$, $u_i = s^i u/(s+\alpha)^n$, $\beta,\alpha > 0$.
If there exist positive α_1 and δ' such that

$$\alpha_1 I < \int_{\sigma}^{\sigma+\delta'} W(t)W^T(t)dt \tag{3.14}$$

for all $\sigma\varepsilon R_+$, then there exist a positive α_2 and a $\delta > \delta'$, such that $\forall\sigma\varepsilon R_+$,

$$\alpha_2 I < \int_{\sigma}^{\sigma+\delta} V(t)V^T(t)dt \tag{3.15}$$

Proof

This runs similar to the proof of Theorem 3.2. Defining ξ with $\|\xi\| = 1$ such that $|\xi^T V|$ is arbitrarily small on $[\sigma,\sigma+\delta]$ for some σ implies that

$$\left| \frac{\sum_{i=0}^{n-1} \gamma_i s^i [\sum_{k=0}^{m} b_k s^k] u}{(s+\alpha)^n (s^n + \sum_{k=0}^{n-1} a_k s^k)} + \frac{\sum_{j=0}^{m} \theta_j s^j u}{(s+\alpha)^n} + i(t) \right|$$

is also arbitrarily small where $i(t)$ consists of quantities decaying at rates determined by poles of the overall closed-loop system, which is stable by the standing assumption. So for $t > \bar{\delta}_1$, and thus $t > \sigma + \bar{\delta}_1$, $i(t)$ becomes "small". The remainder of the proof goes as before.

∇∇∇

It is the reverse implication of Theorem 3.3 which no longer parallels Theorem 3.2. Let us illustrate the problem with the plant of transfer function $\frac{\omega^2}{s^2+\omega^2}$ which is part of an overall stable system, with ω unknown.

This plant may then have an input-output pair

$$u(t) = 1$$

$$y(t) = 1 - \cos\omega t$$

provided that the overall stable system has an external input of the form $\lambda_0 + \lambda_1\cos(\omega t+\phi)$ for some λ_0, λ_1 and ϕ. It can be checked that $\int_{\sigma}^{\sigma+\delta} V(t)V^T(t)dt$ is positive definite for all δ, while obviously the same is not true of $\int_{\sigma}^{\sigma+\delta} W(t)W^T(t)dt$. The problem, as we shall see, is associated with a pole on the $j\omega$-axis: poles of the plant in $\mathrm{Re}[s] > 0$ cause no problem (so long as the plant is embedded in an overall stabilizing arrangement of course, in accordance with the standing assumption). Analysis of this situation depends on the following Lemma.

<u>Lemma 3.1</u> Let $\dot{x} = Ax + Bu$ be a possibly unstable subsystem of an overall larger linear, time-invariant, finite-dimensional system with bounded input which is asymptotically stable, so that all variables associated with the system are bounded by M. Let

$$A = \begin{bmatrix} A_- & 0 & 0 \\ 0 & A_+ & 0 \\ 0 & 0 & A_0 \end{bmatrix} \qquad B = \begin{bmatrix} B_- \\ B_+ \\ B_0 \end{bmatrix}$$

with $\mathrm{Re}\lambda_i(A_-) < 0$, $\mathrm{Re}\lambda_i(A_+) > 0$ and $\mathrm{Re}\lambda_i(A_0) = 0$ for all i. For some $\varepsilon > 0$ define $\delta_1(\varepsilon)$, $\delta_2(\varepsilon)$ by

$$\|e^{A_-\delta_1(\varepsilon)}\| M = \varepsilon , \quad \|e^{-A_+\delta_2(\varepsilon)}\| M = \varepsilon$$

[Notice that $\delta_i(\varepsilon) \leqslant K_i|\ell n\varepsilon|$ for some K_i]. Suppose that the zero-input component of all the large system variables (i.e. the component due to nonzero initial state) is $0(\varepsilon)$ for $t \geqslant \sigma_0(\varepsilon)$. Let $w(s)$ be a proper stable transfer function and

$$|w(s)u| < \varepsilon \quad \text{on} \quad [\sigma, \sigma+\delta]$$

where $\sigma \geqslant \sigma_0$ and $\delta > \delta_1 + \delta_2$. Then for some complex x_i,

$$|w(s)[x(t) - \sum_{i=1}^{\alpha} x_i e^{j\omega_i t}]| \leqslant 0(\varepsilon) + \delta 0(\varepsilon)$$

on $[\sigma+\delta_1, \sigma+\delta-\delta_2]$, where $j\omega_i$ are the eigenvalues of A_0 and $\Sigma\, x_i e^{j\omega_i t}$ is real.

Remark 3.7

Figure 3.1 illustrates the conclusions of the Lemma.

Proof

Make the definitions $U = w(s)u$, $X = w(s)x$. Also, let $\bar{u},\bar{x}$ denote the u,x which would result with zero initial conditions, with $\bar{U},\bar{X}$ defined obviously. Since for $\sigma > \sigma_0$, $|U-\bar{U}|$ is $0(\varepsilon)$, $|X-\bar{X}|$ is $0(\varepsilon)$, it is enough to prove that $|\bar{U}| < \varepsilon$ on $[\sigma, \sigma+\delta]$ implies that $|\bar{X} - \Sigma\, \bar{X}_i e^{j\omega_i t}| \leqslant 0(\varepsilon) + \delta 0(\varepsilon)$ on $(\sigma+\delta_1, \sigma+\delta-\delta_2]$.

As the figure shows, and recognizing the relevance of zero initial condition,

$$\dot{\bar{X}} = A\bar{X} + B\bar{U}$$

Now with $\bar{X}(t) = [\bar{X}_-^T(t)\bar{X}_+^T(t)\bar{X}_0^T(t)]^T$, we have

$$\bar{X}_-(t) = e^{A_-(t-\sigma)}\bar{X}_-(\sigma) + \int_\sigma^t e^{A_-(t-\tau)} B_-\bar{U}(\tau)d\tau$$

and immediately, $\|\bar{X}_-(t)\|$ is $0(\varepsilon)$ on $(\sigma+\delta_1,\sigma+\delta)$. Similarly, using

$$\bar{X}_+(t) = e^{A_+(t-\sigma-\delta)}\bar{X}_+(\sigma+\delta) - \int_t^{\sigma+\delta} e^{A_+(t-\tau)} B_+\bar{U}(\tau)d\tau$$

we obtain $\|\bar{X}_+(t)\|$ is $0(\varepsilon)$ on $(\sigma,\sigma+\delta-\delta_2)$. Last,

$$\bar{X}_0(t) = e^{A_0(T-\sigma)}\bar{X}_0(\sigma) + \int_\sigma^t e^{A_0(t-\tau)} B_0\bar{U}(\tau)d\tau$$

and this yields

$$\|\bar{X}_0(t) - \Sigma X_{0_i} e^{j\omega_i t}\| = 0(\varepsilon\delta)$$

The result is then immediate.

<u>Theorem 3.4</u> Assume the plant is described by (3.1), with

$$A(s) = s^n + \sum_{i=0}^{n-1} a_i s^i \quad \text{and} \quad B(s) = \sum_{j=0}^{m} b_j s^j \quad \text{coprime,}$$

and $u(t) \in \Omega_\Delta[0,\infty)$. Let $A(s)$ have no more than p zeros with zero real part, and thus at least $n-p$ zeros with nonzero real part. With $V(t)$ as in (3.12) and

$$\bar{W}(t) = [u \ \frac{u}{(s+\beta)} \ \cdots \ \frac{u}{(s+\beta)^{n+m-p}}]^T \tag{3.16}$$

if there exist a positive α_3 and δ' such that

$$\alpha_3 I < \int_\sigma^{\sigma+\delta} V(t)V^T(t)\,dt \tag{3.17}$$

for all σ, then there exist a positive α_4, a σ_0 and suitably large $\delta > \delta'$ such that

$$\alpha_4 I < \int_\sigma^{\sigma+\delta} \bar{W}(t)\bar{W}^T(t)\,dt \tag{3.18}$$

for all $\sigma \geqslant \sigma_0$.

For a proof, see appendix 3.B.

Remarks

(3.8) If nothing is known about the zeros of $A(s)$, one must assume $p = n$ in applying this theorem.

(3.9) Taken together, Theorems 3.3 and 3.4 show that if $u(\cdot)$ is a real linear combination of imaginary exponentials, $m+n+1$ different exponents are sufficient, but only $m+n+1-p$ may be necessary, to identify the system. Nondecaying, nongrowing components of the system output stemming from nonzero initial conditions may make up the remaining information required to give the spanning condition on $V(\cdot)$. Given that the unknown system is part of an overall stable system, such special output components can only be present if the external input to the overall stable system contains a sinusoidal component at each of the relevant frequencies. We return to the question of conditions on the external input below.

This theorem is the best possible in the sense that we cannot necessarily expect to have (3.17) implying (3.18) when $\bar{W}(t)$ in (3.18) is replaced by $\hat{W}(t) = [u, (s+\beta)^{-1} u, \ldots, (s+\beta)^{-(n+m-p+1)} u]$, at least without further assumptions on the poles of the plant.

For example if $B(s) = 1$ (i.e. $m = 0$) and

$$y(t) = \sum_{i=1}^{n} \bar{y}_i e^{j\omega_i t} \quad \bar{y}_i \neq 0, \; \omega_i \neq \omega_j, \quad y(t) \text{ real} \tag{3.19}$$

it is easily verified that (3.17) holds. Now as many as p of the frequencies in (3.19) could coincide with the poles of the plant; and in consequence $u(t)$ need contain as few as $n-p$ frequencies. Whence from corollary 3.1 it is clear that while (3.18) holds, the same cannot be said if $\bar{W}$ were to be replaced by $\hat{W}$.

Theorems 3.3 and 3.4 relate $V(t)$, the vector which is crucial in the adaptive parameter adjustment law to the input u of a subsystem of a larger system, itself possessing an external input, r. It is relevant to ask what conditions on r imply condition (3.14) on W (which in turn implies condition (3.15) on V, guaranteeing exponential convergence of an equation error identifier).

The standing assumption for the section implies that

$$u = \frac{b(s)}{a(s)} r \qquad (3.20)$$

for some $a(s)$ with all zeros in $Re[s] < 0$ and with b/a proper. Note that it may be the case that b,a depend on the parameters of the subsystem, and so may not be known. Suppose that $r(t) = \sum_{i=1}^{N} \bar{r}_i e^{j\omega_i t}$. Then $u(t) = \sum_{k=1}^{M} \bar{u}_k e^{j\omega_k t}$ where $M < N$ is possible if $b(s)$ has purely imaginary zeros, and if the ω_i are such that $j\omega_i$ is a zero of $b(s)$. Notice that if $A(s)$, the denominator of the plant or unknown subsystem transfer function, has $j\omega$-axis zeros, $b(s)$ necessarily has such zeros when the overall setup is stable. Now in order that a condition like (3.14) hold for $u(t)$ above it is necessary and sufficient that $M = m + n + 1$. In the absence of any information about the zeros of $b(s)$, one needs

$$N = m + n + 1 + \deg b(s) . \qquad (3.21)$$

The above argument is developed for sinusoids. The content of the next theorem generalizes the above argument.

Theorem 3.5 Assume that the overall system input $r \varepsilon \Omega_\Delta[0,\infty)$ is related to the input of the plant being identified by (3.20), with b/a proper and $a(s)$ possessing all zeros in $Re[s] < 0$. Suppose that the number of $j\omega$-axis zeros of $b(s)$ is no greater than ρ ($\rho = \deg b(s)$ is a possibility). Define

$$R(t) = [r \; \frac{r}{(s+\gamma)} \; \cdots \; \frac{r}{(s+\gamma)^{n+m+\rho}}]^T \qquad (3.22)$$

with $\gamma > 0$ and suppose that for some positive α_1, α_2 and δ, and for all σ

$$\alpha_1 I < \int_{\sigma}^{\sigma+\delta} R(t)R^T(t)\,dt < \alpha_2 I \tag{3.23}$$

Then there exist positive α_3, α_4 such that for some $\delta' > \delta$, with $W(t)$ defined as in (3.13)

$$\alpha_3 I < \int_{\sigma}^{\sigma+\delta'} W(t)W^T(t)\,dt < \alpha_4 I \tag{3.24}$$

for all σ.

Proof

Suppose that (3.24) fails. Then given arbitrary $\varepsilon > 0, \delta'$, we can find a ξ of unit length such that

$$|\xi^T W(t)| < \varepsilon \quad \text{on} \quad [\sigma, \sigma+\delta']$$

i.e.

$$\left| \sum_{i=0}^{n+m} \frac{\xi_i (s+\beta)^{n+m-i}}{(s+\beta)^{n+m}} u \right| < \varepsilon \quad \text{on} \quad [\sigma, \sigma+\delta']$$

i.e.

$$\left| \frac{q(s)b(s)}{a(s)(s+\beta)^{n+m}} r \right| < \varepsilon \quad \text{on} \quad [\sigma, \sigma+\delta']$$

for some polynomial $q(s)$ of degree $n+m$.

Write $b(s) = b_{\pm}(s)b_0(s)$, where $\deg b_0(s) = \rho$ and all $j\omega$-axis zeros of $b(s)$ are included in b_0. It follows, using arguments contained in appendix 3.A that

$$\left| \frac{a(s)(s+\beta)^{n+m}}{b_{\pm}(s)(s+\gamma)^{n+m+\rho}} \; \frac{q(s)b(s)}{a(s)(s+\beta)^{n+m}} \, r \right| < 0(\varepsilon^{1/\nu})$$

on $(\sigma + \bar{\delta}_1, \sigma + \delta)$. Here, $\nu = 1 + \deg a - \deg b$ i.e.

$$\left| \frac{q(s)b_0(s)}{(s+\gamma)^{n+m+\rho}} \, r \right| < 0[\varepsilon^{1/\nu}]$$

on $[\sigma + \bar{\delta}_1, \sigma + \delta]$. From this, it is easy to obtain a contradiction to (3.23), and the theorem is proved.

Remark (3.10)

Observe that in this theorem the coprimeness of $A(S)$ and $B(S)$ is not required.

3.3 Persistence of Excitation for Unstable Plants with Possibly Unbounded Signals

In this section we consider plants which are possibly unstable. Clearly, one cannot expect the upper bound in (3.7) to hold. What is of interest, however, is whether or not the lower bound condition can hold. One should stress that if the upper bound is violated equations of the form in (3.3-3.5) may not be exponentially convergent even if the lower bound is satisfied. Consider for example the system of equations

$$\dot{\chi}(t) = -[e^{t/2}\ 1]^T [e^{t/2}\ 1]\ \chi(t)$$

Clearly with $V(t) \triangleq [e^{t/2}\ 1]$ the lower bound in (3.2) is satisfied as $e^{t/2}$ and 1 are linearly independent. However, it can be verified that

$$\chi_1 = e^{-\frac{1}{2}t}$$

and

$$\chi_2 = -1 + \tfrac{1}{2}e^{-t}$$

form a solution and clearly χ_2 does not converge to zero. There are, however, adaptive laws, (see for example [7]) where the lower bound alone is guaranteed to assure convergence. Also for some of the algorithms to be studied in the later chapters the lower bound in (3.2) guarantees that all signals eventually become bounded. Whence, the upper bound is also satisfied.

The particular results we shall establish are the following:

(1) Condition for satisfaction of the lower bound in (3.7) given bounded inputs

(2) Condition for satisfaction of the lower bound in (3.7) given possibly unbounded inputs

(3) Condition for satisfaction of the lower bound of a gramian of the form in (3.7) but involving a vector of the output and its filtered versions, viz.

$$Y_k = [y, \frac{y}{s+\gamma}, \ldots. \frac{y}{(s+\gamma)^{k-1}}]^T$$

For the rest of the chapter the subscript k used with any of the vectors Y, W or V will denote the dimension of that vector. The proofs require technical lemmata which are given in Appendix 3.C. It is worth noting that the proofs in this section differ from the ones in the previous sections on several counts. To begin with Lemma 3.A.1 is no longer applicable due to unbounded signals. Thus, our arguments will involve integrals of the form $\int_{\sigma}^{\sigma+\delta} |\theta^T V| dt$ as opposed to quantities like $|\theta^T V|$. Consequently Lemma 3.A.2 is modified.

Suppose the input $v(t)$ to an asymptotically stable system is bounded and $\int_{\sigma}^{\sigma+\delta} |v(t)| dt$ is "small". Then the integral

$$\int_{\sigma+\delta'}^{\sigma+\delta} |w(t)| dt \tag{3.26}$$

is also "small", $w(t)$ being the output of the system and δ' a positive number. However, if $v(t)$ is not bounded for all $t \varepsilon R_+$ then this may not always be true. Consider, however, the case where $v(t)$ is the output of a system having a transfer function

$$\frac{G(s)}{H_-(s)H(s)}$$

where $H_-(s)$ is Hurwitz. Then $w(t)$ given by

$$w = \frac{G_1(s)H(s)}{H_2} v$$

with G_1H/H_2 proper, and H_2 Hurwitz, will be such that the integral in (3.26) will be "small". This is mainly because the unstable components of v would be unobservable at w. The lemmata (3.C.1-3.C.3) used at various points in the proofs of this section are special cases of this result.

Finally, the following lemma, similar to lemma 3.1 will also prove useful.

Lemma 3.2 Consider the output $y(t)$ of a system with a transfer function $(s-a)/(s+b)$, bounded input $v(t)$ and arbitrary finite initial conditions, a and b being complex numbers with positive real part. Define $w(t)$ as the output of any system with transfer function $1/(s+b)$ and input $v(t)$. Suppose $\exists\, M$ such

$$|w(t)| < M \qquad \forall\, t \varepsilon R_+$$

For any $\varepsilon > 0$, define a real $\delta'(\varepsilon)$ by

$$M|e^{-a\delta'(\varepsilon)}| = \varepsilon$$

Suppose $\exists$ a $\delta > \delta'(\varepsilon)$ and a $\sigma > 0$ such that

$$|y(t)| \leqslant \varepsilon \qquad \forall\, t\ \varepsilon[\sigma, \sigma+\delta] \tag{3.27}$$

Then

$$|w(t)| \leqslant 0(\varepsilon) \qquad \forall\, t\, \varepsilon[\sigma, \sigma+\delta-\delta'(\varepsilon)] \tag{3.28}$$

Proof

Since

$$y(t) = \frac{s-a}{s+b} v(t)$$

and

$$w(t) = \frac{1}{s+b} v(t)$$

$$w(t) = \frac{1}{s-a} y(t)$$

$$= w(\sigma+\delta) e^{-a(\sigma+\delta-t)} + \int_{\sigma+\delta}^{t} e^{a(t-\tau)} y(\tau) d\tau$$

$$\forall t \leqslant \sigma + \delta$$

Now, $\delta'(\varepsilon)$ is such that $M|e^{-a\delta'(\varepsilon)}| = \varepsilon$.

Thus $|w(t)| \leqslant 0(\varepsilon) \forall \quad t \varepsilon [\sigma, \sigma + \delta - \delta']$

as from (3.27) $|y(t)| \leqslant \varepsilon \forall \quad t \varepsilon [\sigma + \delta - \delta'(\varepsilon), \sigma+\delta]$

We now prove the main results of this section.

Theorem 3.6 Consider the linear, time invariant SISO, proper system described by

$$A(s) y(t) = B(s) u(t) \tag{3.29}$$

where $A(s)$ and $B(s)$ are coprime polynomials of degrees n and m respectively, $m \leqslant n$ and $u(t) \varepsilon \Omega_{\Delta}[0,\infty)$.
Define

$$V_{n+m+1}(t) = [\ \frac{y(t)}{(s+\alpha)^n}\ ,\ldots,\ \frac{s^{n-1}}{(s+\alpha)^n}\ y(t),\ \frac{1}{(s+\alpha)^n}\ u(t)\ ,\ldots,\ \frac{s^m}{(s+\alpha)^n}\ u(t)\ ]^T$$

and

$$W_{n+m+1}(t) = [u(t),\ \frac{u(t)}{(s+\beta)}\ ,\ldots,\ \frac{u(t)}{(s+\beta)^{n+m}}\ ]$$

$\beta \varepsilon R_+$ and $\alpha \varepsilon R$. Let $\exists\ \alpha_1, \delta' > 0$ such that $\forall\ \sigma \varepsilon R_+$

$$\int_{\sigma}^{\sigma+\delta'} W_{n+m+1}(t)\ W^T_{n+m+1}(t)dt \geqslant \alpha_1 I \tag{3.30}$$

Then $\exists\ \alpha_2$ and $\delta > \delta'$ such that $\forall\ \sigma \varepsilon R_+$

$$\int_{\sigma}^{\sigma+\delta} V_{n+m+1}(t)\ V^T_{n+m+1}(t)dt \geqslant \alpha_2 I \tag{3.31}$$

Proof

See Appendix 3.C.

Remark (3.11)

It is possible to show that in going from (3.31) to (3.30) the dimension of W may be reduced to as small as $n + m + 1 - \mu$, μ being the number of poles of (3.29) which lie in the closed right half plane. This is because in general the system is capable of producing upto μ independent excitations, arising due to initial condition effects, which do not decay to zero.

(3.12)

Notice that α is no longer required to be positive. We now prove a result with bounded inputs which relates p.e. conditions on the input to their counterparts involving outputs. Such a result is important in adaptive control. In particular in chapter 6 the controller-cum-system can be configured as depicted in figure 3.2, where $u(t)$ could well be unbounded. Equivalently, a non-minimal representation of the form in fig. 3.3 also exists. Then the result given below helps to establish conditions on the reference input $r(t)$ which guarantee the p.e. of $u(t)$ and thus indirectly of a relevant regression vector.

Theorem 3.7 Consider the proper time-invariant SISO system described by

$$A(s)y(t) = B(s)u(t) \tag{3.32}$$

with $A(s)$ and $B(s)$ polynomials of degree n and m respectively. Assume $u(t) \varepsilon \Omega_{\Delta}[0,\infty)$ and let z be the number of zeros of $B(s)$ with zero real part.

Define

$$y_{\nu} = [y, \frac{y}{s+\beta}, \ldots, \frac{y}{(s+\beta)^{\nu}}]^T$$

and

$$W_{\nu+z} = [u, \frac{u}{s+\alpha}, \ldots, \frac{u}{(s+\alpha)^{\nu+z}}]^T$$

Let $\alpha \varepsilon R_+$ and $\beta \varepsilon R$. Suppose there exist α_1 and δ', such that $\forall\ \sigma \varepsilon R_+$

$$\int_{\sigma}^{\sigma+\delta} W_{\nu+z}(t) W_{\nu+z}^T(t)\, dt \geqslant \alpha_1 I \ . \tag{3.33}$$

Then there exist $\delta > \delta'$ and $\alpha_2 > 0$ such that $\forall \sigma \varepsilon R_+$

$$\int_{\sigma}^{\sigma+\delta} Y_{\nu}(t) Y_{\nu}^T(t)\, dt \geqslant \alpha_2 I \tag{3.34}$$

Proof

See Appendix 3.C.

Remark (3.13)

The result does not require $A(s)$ and $B(s)$ to be coprime. The reduction in the dimension of Y is due to the fact that if $u(t)$ were a linear combination of $\nu + z$ sinusoids then as many of z of these could correspond to the z imaginary axis zeros of $B(s)$. Had we allowed $u(t)$ to be unbounded then the number z would have equalled the number of zeros of $B(s)$ which belong to the closed right half plane.

We now extend theorem 3.6 by relaxing the boundedness assumption on $u(t)$ and remark that theorem 3.7 can be similarly extended.

Let $\{F,G,H,j\}$ be any minimal state variable realization of the single-input multiple-output transfer function

$$[\frac{1}{(s+\beta)^{n-m}}, \ldots, \frac{1}{(s+\beta)^{2n+1}}]^T .$$

Let $x(t)$ be the corresponding state vector. Clearly

$$W(t) = Hx(t) + ju(t)$$

Define the set $W^{(\sigma,M)}$ to be the set of all trajectories $W(\cdot)$ for which $\|x(\sigma)\| < M$. Then the following is true

Theorem 3.8 Consider the linear, time invariant, proper system

$$A(s)y(t) = B(s)y(t)$$

where $A(s)$ and $B(s)$ are coprime polynomials of degrees n and m respectively.

Define

$$V^T_{n+m+1}(t) = [\frac{y(t)}{(s+\alpha)^n}, \ldots, \frac{s^{n-1}}{(s+\alpha)^n} y(t), \frac{1}{(s+\alpha)^n} u(t), \ldots, \frac{s^m}{(s+\alpha)^n} u(t)]$$

and

$$W^T(t) = [\frac{u(t)}{(s+\beta)^{n-m}}, \ldots, \frac{u(t)}{(s+\beta)^{2n}}]$$

$\beta \epsilon R_+$ and $\alpha \epsilon R$. Suppose $\exists\, \alpha_1, \delta' > 0$ and independent of σ such that $\forall\, \sigma \epsilon R_+$, any finite M and any $W(\cdot) \epsilon W^{(\sigma,M)}$

$$\int_{\sigma}^{\sigma+\delta'} W(t)\, W^T(t)\,dt \geq \alpha_1 I \tag{3.35}$$

where $W^{(\sigma,M)}$ is defined above.

Then $\exists\, \alpha_2,\ \delta > \delta'$ and independent of σ such that $\forall \sigma \varepsilon R_+$

$$\int_{\sigma}^{\sigma+\delta} V_{n+m+1}(t)\ V^T_{n+m+1}(t)\,dt \geq \alpha_2 I\ . \tag{3.36}$$

Proof

See Appendix 3.C.

Remark (3.14)

If the condition on the Gramian in (3.35) holds for a particular M and δ' then it holds for any other M' which is finite and for some δ''. This is because, by definition $\|x(\sigma)\| < M$ in the one case and $\|x(\sigma)\| < M'$ in the other. Thus δ'' is determined by the time required for the boundary conditions to decay, at rates governed by β.

(3.15)

The above remark indicates how one would proceed to check condition (3.35) in practice:

Step I

Pick an arbitrary M

Step II

Choose an arbitrary, minimal realization of $[\frac{1}{(s+\beta)^{n-m}}, \ldots, \frac{1}{(s+\beta)^{2n}}]$ (viz. F,G,H,j), initialize the state to zero and generate $W(t) = Hx(t) + ju(t)$ until $\|x(t)\| \geq M$. Compute the Gramian (3.35) using $W(t)$.

Step III

At any t_i where $\|x(t_i)\| = M$, reinitialize $x(t_i)$ to zero. Repeat step II.

Finally, we note that the results of this section are also useful in proving global exponential convergence of hybrid controllers [24, 25]. Consider for example the configurations of figures 3.2 and 3.3. Let the controller parameters be updated at discrete time instants and let them be constant over intervals of length T_k. Then using the results of this section it is easy to prove that the associated regression vector is p.e. whenever $r(t)$ is p.e. Thus parameter convergence at an exponential rate is immediate. Suppose the controller is designed in a way that when the controller parameters equal the true plant parameters, the stationary closed loop system is stable. Then global, robust stability can be concluded. In [25] a similar analysis has been carried out under the assumption that $r(t)$ is a linear combination of sinusoids. This work, however, deals with a far more general class of reference inputs.

3.4 Persistence of Excitation of Slowly Varying Plants with Bounded Signals

In this section we consider persistence of excitation properties for slowly time varying plants. The main result derived here may have wide applicability in the analysis of a wide class of adaptive controllers. As stated earlier, many adaptive controllers can be configured as in fig. 3.2, where the controller may well be time varying. However, in most adaptive control applications eq. in [23] the variations in the controller parameters approach asymptotically to zero, even without persistence of excitation. Thus the results of this section will give conditions on the reference input which will ensure p.e. of the control input once the initial parameter variations decay. We now state the main result of this section, a discrete time equivalent of which can be found in [13]. The proof is available in Appendix 3.D.

Theorem 3.9 Consider the linear time varying n-dimensional, SISO system

$$\dot{x}(t) = A(t)x(t) + b(t)r(t)$$
$$u(t) = c^T(t)x(t) + d(t)r(t) \qquad (3.37)$$

with $r(t)$ and $x(t)$ $\varepsilon\Omega_\Delta[0,\infty)$. $A(\cdot)$, $b(\cdot)$, $c(\cdot)$ and $d(\cdot)$ are bounded. Suppose $\exists\, m_1$ such that

$$||c(t)|| \; ||b(t)|| + |d(t)| > m_1 \tag{3.38}$$

$\forall t$ except on a set of measure zero. Suppose there also exist M_A, M_B, M_C, M_D and δ such that

$$\begin{aligned} ||A(t+\delta) - A(t)|| &\leq M_A \\ ||b(t+\delta) - b(t)|| &\leq M_B \\ ||c(t+\delta) - c(t)|| &\leq M_C \\ |d(t+\delta) - d(t)| &\leq M_D \end{aligned} \tag{3.38a}$$

$\forall t$. Define

$$R_{\nu+n}(t) = [r(t), \; . \; . \; . \; . \; . \; . \; . \; . \; , \; \frac{r(t)}{(s+\alpha)^{\nu+n}}]^T$$

where $\alpha > 0$ and n is the dimension of $x(\cdot)$. Suppose there exists a unit length vector $\theta^T = [\theta_o, \; \ldots \; , \theta_\nu]$ and $\varepsilon > 0$ such that for some σ

$$\int_{\sigma}^{\sigma+\delta} |\theta^T W_\nu(t)| dt < \varepsilon \tag{3.38b}$$

with

$$W_\nu(t) \triangleq [u(t), \; . \; . \; . \; . \; . \; . \; . \; . \; , \; \frac{u(t)}{(s+\beta)^\nu}]^T$$

and some $\beta > 0$. Then $\exists$ a $\delta' < \delta$ and a γ, $\|\gamma\|$ bounded away from zero for which

$$\int_{\sigma+\delta'}^{\sigma+\delta} |\gamma^T R_{\nu+n}(t)|dt < 0(\varepsilon) + \{K_1(\beta,\alpha)M_A + K_2(\beta,\alpha)M_b + K_3(\beta,\alpha)M_c + K_4(\beta,\alpha)M_d\}M \quad (3.38c)$$

where the K_i are the functions of β and α and M is the bound on $|r(t)|$ and $\|x(t)\|$.

Remark (3.16)

Put differently, the above theorem requires that if

$$\int_{\sigma}^{\sigma+\delta} R_{\nu+n}(t)\ R^T_{\nu+n}(t)dt > \alpha_1 I$$

where α_1 is greater than some quantity determined by the magnitude of the system signals and time variation in the parameters then $\alpha_2 > 0$ such that

$$\int_{\sigma}^{\sigma+\delta} W_\nu(t)\ W^T_\nu(t)dt > \alpha_2 I$$

for some $\delta' < \delta$.

(3.17)

Condition (3.38) ensures that the output signal is not identically zero over an interval for arbitrary input. It is worth noting that in contrast with the previous sections the boundedness of signals is crucial even to get the lower bound of the p.e. condition to hold.

3.5 Model Reference Control : known gain.

In the remainder of this chapter we examine the exponential stability of a Model Reference Adaptive Control algorithm proposed by Morse [16]. Apart from being an interesting problem in its own right this will demonstrate an application of some of the results derived in this chapter. In this section we consider the case where the high frequency gain, g_p, of the plant to be controlled is known *a priori* while in section 3.6 the unknown g_p case is considered.

First we briefly outline the philosophy and nature of Morse's algorithm, adhering closely to the terminology employed in [16], but omitting details irrelevant to the course of our development.

Consider a single-input single-output plant, modelled by a strictly proper transfer function

$$T_p(s) = \frac{g_p \, \alpha_p(s)}{\beta_p(s)} \tag{3.39}$$

having degree and relative degree of n and n^* respectively; g_p is a non zero constant with known signs (assumed positive) and α_p and β_p are monic, coprime polynomials and $\alpha_p(s)$ is strictly stable as well. It is assumed that the output $y(t)$ is required to track a reference trajectory $y_r(t)$, itself the output of a known, stable transfer function T_r, having relative degree no smaller than n^* (as otherwise explicit differentiation would be required) and a reference input $r(t)$.

Consider next the scheme depicted in figure 3.4. Here $1/\beta_r$ is an arbitrary, known, stable, all pole transfer function of degree n^*, $W(s)$ is a stable transfer function, $\theta^T = (\theta_u^T, \theta_y^T, \theta_r)$ is an auxiliary signal vector and $\hat{k} = (\hat{k}_u, \hat{k}_y, \hat{k}_r^T)$ is a parameter estimate vector. The true value k_p of $\hat{k}$ is such that when $\hat{k} = k_p$, the transfer function relating θ_r to y equals $1/\beta_r$, with k_{up}, k_{yp} and $k_{rp} = 1/g_p$ serving to assign the zeros, poles and gain of the plant respectively.

Defining $k = \hat{k} - k_p$ we can redraw figure 3.4 as figure 3.5 with the transfer function representing the system within the dotted box equalling $1/\beta_r(s)$. Accordingly we have the error model of figure 3.6 as

$$e = y - y_r = \frac{1}{\beta_r}(\theta_r + g_p k^T \theta) - \frac{1}{\beta_r}\theta_r = \frac{g_p}{\beta_r} k^T \theta .$$

Now if we were to identify k by performing a steepest descent on e^2, existing results tell us [10] that $1/\beta_r$ will need to be strictly positive real, a condition clearly unattainable for $n^* > 1$. In [16] this problem is circumvented by adding an auxiliary signal so that the error model for this augmented signal e'' approaches that in figure 3.7 asymptotically. This auxiliary signal is $\psi(t)\hat{g}(t)$ in figure 3.8 and the augmented error equals e''.

Clearly $$\psi = -\frac{1}{\beta_r}(\theta^T(k_p + k)) + \phi^T(k_p + k)$$

$$= \phi^T k(t) - \frac{1}{\beta_r}\theta^T(t) k(t)$$

Thus if $\hat{g}(t) \equiv g_p$, $e'' \equiv g_p \phi^T k(t)$. However, in general unless g_p is known a priori, this condition will not hold. Given below are the equations which define the signals in figure 3.8. Let A,b,c be a minimal realization with

$$c^T(sI-A)^{-1}b = 1/\beta_r \tag{3.40}$$

$$\phi^T = c^T H \tag{3.41}$$

$$\dot{H} = AH + b\theta^T \tag{3.42}$$

$$\dot{z} = Az + b\hat{k}^T\theta \tag{3.43}$$

$$\psi = \hat{k}^T\phi - c^T z \tag{3.44}$$

$$\dot{\hat{g}} = -q\,\psi\,\bar{e} \quad , \qquad q > 0 \tag{3.45}$$

$$\dot{\hat{k}} = -Q\,\phi\,\bar{e} \qquad Q > 0 \tag{3.46}$$

with

$$Q = \begin{bmatrix} Q_1 & 0 \\ 0 & q_2 \end{bmatrix} \tag{3.47}$$

and

$$\bar{e} = \frac{1}{\lambda_0 + \phi^T Q\phi}(\hat{g}\psi + e) \quad , \qquad \lambda_0 > 0 \tag{3.48}$$

Observe that $\bar{e}$ in (3.48) is in fact e'' of figure 3.8 divided by $(\lambda_0 + \phi^T Q\phi)$.

The vector $\theta^T \triangleq [\theta_u^T, \theta_y^T, \theta_r]$ is generated by

$$\dot{\theta}_y = A_0\theta_y + b_o y \tag{3.49}$$

$$u = \hat{k}^T \theta \tag{3.50}$$

$$\theta_r = \beta_r(s) T_r(s) r \tag{3.51}$$

$$y_r = T_r(s) r \tag{3.52}$$

$$\dot{\theta}_u = A_0 \theta_u + b_0 u \tag{3.53}$$

where $[A_0, b_0]$ is a completely controllable n-dimensional stable system. If g_p is known a priori then $\hat{g}(t)$ can be set to g_p and (3.45) can be dispensed with. Then by specializing the error model given in [16] we get

$$\dot{x} = \bar{A}x + \bar{b}(\bar{k}^T \theta + r) \tag{3.54}$$

$$\theta = \bar{c}x + \bar{d}r + \bar{\varepsilon}(t) \tag{3.55}$$

$$\dot{H} = AH + b\theta^T \tag{3.56}$$

$$\phi^T = c^T H \tag{3.57}$$

$$\dot{z} = Az + b\bar{k}^T \theta \tag{3.58}$$

$$\dot{\bar{k}} = - Q_1 \bar{\phi} \bar{e} \tag{3.59}$$

$$\bar{e} = (\lambda_0 + \phi^T Q \phi)^{-1} (g_p \bar{k}^T \bar{\phi} + \varepsilon(t)) \tag{3.60}$$

Here $\bar{\varepsilon}(t)$ and $\varepsilon(t)$ are exponentially decaying quantities, $(\bar{A}, \bar{b}, \bar{c}, \bar{d})$ is an unknown but strictly stable system and $\bar{k}$ and $\bar{\phi}$ are given by

$$\bar{k} = \begin{bmatrix} I_{2n} & 0 \\ 0 & 0 \end{bmatrix} k \qquad \bar{\phi} = \begin{bmatrix} I_{2n} & 0 \\ 0 & 0 \end{bmatrix} \phi \tag{3.61}$$

Note that the replacement of k and ϕ by $\bar{k}$ and $\bar{\phi}$ respectively becomes possible since, with known g_p, the last element of k becomes zero. The stability results of [16], applicable for unknown g_p, may in this case be trivially modified to yield the Proposition below. Following the Proposition, we indicate in Theorem 3.10 a development of this stability result to reflect exponential stability under a persistence of excitation condition.

Proposition 3.1: For any time $t > 0$ and bounded, piecewise continuous input $r(t)$, the state response of the adaptive control system defined by (3.54 - 3.61) is uniformly bounded and the quantities $\bar{e}$ and $\dot{\hat{k}}$ decay asymptotically to zero.

Remark:

(3.18) Observe that Proposition 3.1 assures the boundedness of signals without pe. This is precisely the situation we had stated in the beginning of section 3.2, as being typical of adaptive controllers. As matters stand, however, only the tracking error converges to zero. Although $\dot{\hat{k}}$ decays to zero, $\hat{k}$ need not converge to the right parameter values, let alone doing so exponentially. Theorem 3.10 presents a condition on the reference input which guarantees exponential convergence.

<u>Theorem 3.10</u>: For a reference input $r(t) \varepsilon \Omega_{\Delta}[0,\infty)$ and the adaptive control system with known g_p defined by (3.54 - 3.61), the quantities k and $\bar{e}$ approach zero exponentially fast, provided that for some α_{11}, α_{12} and $\Delta>0$ and all $\sigma \varepsilon R_+$ the following relation holds:

$$\alpha_{11} I < \int_{\sigma}^{\sigma+\Delta} R(t) R^T(t) dt < \alpha_{12} I \tag{3.62}$$

where

$$R = \left[r, \frac{r}{s+\bar{\beta}} \cdots \frac{r}{(s+\bar{\beta})^{2n+n_r-1}} \right]$$

$\bar{\beta}$ is any positive number, and n_r is the number of imaginary axis zeros of $T_r(s)$.

For a proof, see Appendix 3.E.

<u>Remark</u>:

(3.19) It is evident from figure 3.6, that the exponential decay of k leads to the exponential decay of the tracking error e . Thus the condition in (3.62) also guarantees the exponential convergence of e . A similar result for the adaptive controller in [18] has been derived in [14-15] .

3.6 Model reference control : Unknown gain.

In this section we examine Morse's algorithm, described in section 3.5, for exponential stability when the high frequency gain is unknown. As shown in [16] , the algorithm is globally stable in the sense that all the signals are bounded. Yet, while the tracking error converges to zero, the parameter estimates do not necessarily converge to their desired values. However, as we have shown, when the high frequency gain is known a priori, the exponential convergence of parameter estimates results whenever the reference input is persistently exciting. We now consider the case where the high frequency gain is unknown. Rather surprisingly it is discovered that in such a case the structure of the algorithm precludes exponential convergence even with sufficiently rich reference inputs. The problem appears to lie in the identification of the gain parameter. The signal central to its identification converges asymptotically to zero and consequently loses persistence of excitation.

Before we proceed to prove our assertion we need the following stability lemma, a discrete time analogue of which is available in [17] .

Lemma 3.2.

Consider the differential equation

$$\dot{\omega} = G(\omega,t)\omega \tag{3.63}$$

with $G(\omega,t)$ continuous in ω with a Lipschitz condition uniform in t and $\omega(\cdot) \varepsilon R^n$. Assume that initial conditions on ω, lie in some arbitrarily large but bounded region. Then the linearization of (3.63) about the zero trajectory is exponentially stable whenever (3.63) is exponentially stable.

Proof

Exponential stability of (3.63) implies the existence of a Lyapunov function $V(\omega,t)$ and $c_1,\ldots,c_4 > 0$ [26,p.86] such that

$$c_1 \|\omega\|^2 < V(\omega,t) < c_2 \|\omega\|^2$$

$$\frac{dV}{dt} = \frac{\partial V}{\partial t} + \sum_{i=0}^{n} \frac{\partial V}{\partial \omega_i} f_i(\omega,t) \leq -c_3 \|\omega\|^2$$

$$\left|\frac{\partial V}{\partial \omega}\right| \leq c_4 \|\omega\| .$$

Consider any two possible initial conditions of (3.63) $\omega_1(0)$ and $\omega(0)$, both bounded in magnitude by K. Now, there exists a $c_5 > 0$ such that

$$\|G(\omega_1(t),t) - G(\omega(t),t)\| \leq c_5 \|\omega_1(t) - \omega(t)\|$$

$$\leq c_5 \{\|\omega_1(t)\| + \|\omega(t)\|\}$$

Since (3.63) is exponentially stable both $\|\omega_1(t)\|$ and $\|\omega(t)\|$ decay exponentially to zero. Thus $\exists$ a T, depending on K, such that $\forall t \geq T$ there holds $\|\omega_1(t)\| + \|\omega(t)\| \leq (1-q')c_3/(nc_4c_5)$ with c_3, c_4 and c_5 defined as above and $0 < q' < 1$.

Thus

$$\| G(\omega_1(t),t) - G(\omega(t),t) \| \leq \frac{(1-q')c_3}{nc_4}, \forall t \geq T \quad (3.64)$$

Now, as shown in [26, p.86] the perturbed system

$\dot{\omega}(t) = G(\omega,t)\omega + R(\omega,t)$ will be exponentially stable if

$$|R_i(\omega(t),t)| \leq \frac{(1-q')c_3 \| \omega \|}{nc_4} \qquad i = 1,2,\ldots,n .$$

Thus using (3.64), for all sufficiently large t

$$\dot{\omega} = G(\omega(t),t)\omega + (G(\omega_1(t),t)\omega - G(\omega(t),t)\omega) = G(\omega_1(t),t)\omega \quad (3.65)$$

is exponentially stable by (3.64) . Now, suppose $\omega_1(0) = 0$. Then clearly $\omega_1(t) = 0 \; \forall t \geq 0$. Thus by (3.65) $\dot{\omega} = G(0,t)\omega$ is exponentially stable as well.

∇∇∇

It has been shown in [16] that the error model reduces to

$$
\begin{aligned}
\dot{x} &= \bar{A}\,x + \bar{b}\,(k^T\theta + r) \\
\theta &= \bar{c}\,x + \bar{d}r + \bar{\varepsilon}(t) \\
\overset{\circ}{H} &= AH + b\theta^T \\
\phi^T &= c^T H \\
\dot{z} &= Az + bk^T\theta \\
\psi &= k^T\phi - c^T z \\
\dot{k} &= -Q\phi\bar{e} \\
\dot{g} &= -q\psi\bar{e} \\
\bar{e} &= \frac{1}{\lambda_0+\phi^T Q\phi}[g_p k^T\phi + g\psi + \varepsilon(t)]
\end{aligned}
\qquad (3.66)
$$

where $[\bar{A},\bar{b},\bar{c},\bar{d}]$ is a stable, unknown system and $\bar{\varepsilon}(t)$ and $\varepsilon(t)$ are linear combinations of exponentially decaying signals.

We have then the following theorem.

<u>Theorem 3.11</u>. Consider equations (3.66) . Define $x_o(t)$, $H_o(t)$ and $\phi_o(t)$ as the values obtained for $x(t)$, $H(t)$ and $\phi(t)$ when $k \equiv 0$, $g = \hat{g}-g_p \equiv 0$ and $\tilde{x}$, $\tilde{H}$ and $\tilde{\phi}$ as $x-x_o$, $H-H_o$ and $\phi-\phi_o$ respectively. Then for bounded and piecewise continuous reference inputs,

$$\omega^T \triangleq [\tilde{x}^T,\ \tilde{\phi}^T,\ \tilde{\tilde{H}}^T,\ z^T, k^T]$$

is not exponentially stable, with $\tilde{\tilde{H}} = [\tilde{h}_1^T, \tilde{h}_2^T, \ldots, \tilde{h}_{2n+1}^T]$, $\tilde{h}_i$ being the ith column of $\tilde{H}$.

Proof.

The error model becomes

$$\dot{\tilde{x}} = \bar{A}\,\tilde{x} + \bar{b}\,k^T\theta$$

$$\tilde{\theta} = \bar{c}\,\tilde{x}$$

$$\dot{\tilde{H}} = A\tilde{H} + b\tilde{\theta}^T$$

$$\dot{z} = Az + bk^T\theta$$

$$\psi = k^T\phi - c^T z$$

$$\dot{k} = -\frac{Q\,\phi}{\alpha(t)}\,(g_p k^T\phi + g\psi)$$

$$\dot{g} = -\frac{q\psi}{\alpha(t)}\,(g_p k^T\phi + g\psi)$$

where $\alpha(t) = (\lambda_0 + \phi^T Q\phi)$. In this we have neglected the exponentially decaying signals $\varepsilon(t)$ and $\bar{\varepsilon}(t)$. Thus we have

$$\dot{\omega} = G(\omega,t)\,\omega \qquad (3.67)$$

with

$$G(\omega,t) = \begin{bmatrix} \bar{A} & 0 & 0 & \bar{b}\theta^T & 0 \\ f(b,c) & A & 0 & 0 & 0 \\ 0 & 0 & A & b\theta^T & 0 \\ 0 & 0 & 0 & -\frac{Q\phi}{\alpha(t)}g_p\phi^T & -\frac{Q\phi}{\alpha(t)}\psi(t) \\ 0 & 0 & 0 & -\frac{q\psi}{\alpha(t)}g_p\phi^T & -\frac{q}{\alpha(t)}\psi^2 \end{bmatrix}$$

for some function $f(b,c)$, bounded for bounded b,c .

Now by [16] for piecewise continuous and bounded $r(t)$ all entries of $G(\cdot,\cdot)$ are bounded and $G(\cdot,\cdot)$ can easily be shown to be Lipschitz. Now the last row corresponding to update law for $\hat{g}$ is zero when $\omega = 0$, since $\psi = 0$ when $\omega = 0$. Thus $\dot{\omega} = G(0,t)\omega$ is not exponentially stable, whence by Lemma 3.2, (3.67) cannot be exponentially stable.

Remark.

(3.20) The intuition behind the lack of exponential stability is the following. The equations governing the update laws for $\hat{k}$ and $\hat{g}$ can be rewritten as

$$\begin{bmatrix} \dot{\hat{k}} \\ \dot{\hat{g}} \end{bmatrix} = - \begin{bmatrix} Q/\alpha(t) & 0 \\ 0 & q/\alpha(t) \end{bmatrix} \begin{bmatrix} \phi \\ \psi \end{bmatrix} \bar{e} =$$

$$- \frac{1}{\alpha(t)} \begin{bmatrix} Q/g_p & 0 \\ 0 & q \end{bmatrix} \begin{bmatrix} \phi \\ \psi \end{bmatrix} \begin{bmatrix} \phi^T & \psi \end{bmatrix} \begin{bmatrix} k \\ g \end{bmatrix} + \varepsilon_2(t) \tag{3.68}$$

where $\varepsilon_2(t)$ is exponentially decaying. Thus by a result in [10] (3.68) will be exponentially stable when $[\phi^T, \psi]$ is persistently exciting. However, in [16], it is shown that $\lim_{t\to\infty} \dot{\hat{k}}(t) = 0$. Thus from figure 3.8 it is clear that $\lim_{t\to\infty} \psi(t) = 0$. Thus persistence of excitation is being asymptotically lost.

(3.21) It is clear from the foregoing that the lack of exponential stability is inherent in the very structure of this algorithm. More precisely the problem lies in the fact that g_p is estimated as both $1/\hat{k}_r$ and $\hat{g}$ and the update law of $\hat{g}$ is structurally deficient in the sense that the associated component of the regression vector approaches zero. The reason for updating $\hat{g}$ in the first place is to avoid the problem of inverting $\hat{k}_r$ when the latter crosses zero. If a lower bound on the magnitude of g_p is known this problem is avoidable [27].

(3.22) The convergence in the $n^* \leq 1$ case is of course trivial as then $1/\beta_r$ can be chosen as positive real and the artifact of figure 3.8 becomes redundant.

(3.23) Whether or not this result applies to other model reference adaptive control algorithms, is in principle an open question. However, in view of the broad similarity of the algorithms [16,18] , for example, and in the light of the results in [15] one would expect the conclusion to be general.

APPENDIX 3.A : Proof of Theorem 3.2

The proof of Theorem 3.2 depends on the following Lemmas, the first of which has been obtained from [22].

Lemma 3.A.1: If $f(.)$ is an n times differentiable function on an interval I of length Δ and if $|f(x)| < M_0$ and $|f^{(n)}(x)| < M_n$ then for $x \varepsilon I$ and for $0 < k < n$

$$|f^{(k)}(x)| < 4e^{2k}\{^nC_k\}^k M_0^{(1-k/n)} M_n'^{(k/n)} \qquad (3.A.1)$$

where $M_n' = \max(M_n, M_0 n! \Delta^{-n})$, and $^nC_k = {}^{n!}/\{(n-k)!k!\}$.

Lemma 3.A.2: For any stable system with a proper transfer function $T(s)$, if the input $u(t)$ is such that there exist M and $\varepsilon > 0$ for which

$$|u(t)| < M \qquad \text{on} \quad [0,T] \qquad (3.A.2)$$

$$|u(t)| < \varepsilon \qquad \text{for all} \quad t > T ,$$

and if the initial state lies in some fixed ball B of radius R then there exists a $\nu(\varepsilon)$ independent of T , such that for $t > \nu + T$, $|y(t)| < 0(\varepsilon)$.

Proof: For any minimal realization $\{F,G,H,J\}$ of $T(s)$ stability and (3.A.2) imply the existence of a K such that

$$|x(t)| < KM + \| e^{Ft} \| R \quad \text{on} \quad [0,T]$$

where $x(t)$ is the state of the system.

Thus, for $t > T$

$$||x(t)|| < ||KMe^{F(t-T)}|| + ||e^{Ft}|| R + 0(\varepsilon)$$

whence,

$$|y(t)| < KM ||He^{F(t-T)}|| + ||H \quad e^{FT}|| R + 0(\varepsilon)$$

since

$$y(t) = Hx(t) + Ju(t) .$$

Thus if ν is selected to make $||e^{F\nu}|| < \varepsilon$ it follows that for $t > \nu + T$

$$|y(t)| < 0(\varepsilon) . \qquad \nabla\nabla\nabla$$

<u>Lemma 3.A.3</u>: If $u(t)\varepsilon\Omega_\Delta[0,\infty)$, then under the assumption of arbitrary finite initial conditions, for any Hurwitz polynomial $D(s)$ and polynomials $N_1(s)$ and $N_2(s)$, such that $\delta[N_1(s)] \leq \delta[D(s)]$ and $\delta[N_2(s)] = 1 + \delta[D(s)]$, the following properties hold:

(i) $\{N_1(s)/D(s)\}u\varepsilon\Omega_\Delta[0,\infty]$

and

(ii) $\{N_2(s)/D(s)\}u$ is continuous and bounded on $\{[0,\infty) - C_\Delta\}$, and has finite limits as $t \downarrow t_i$ and $t \uparrow t_i$, $t_i \varepsilon C_\Delta$.

<u>Proof</u>: The proof follows from a simple modification of the argument presented in the proof of Lemma 3.A.2.

$\nabla\nabla\nabla$

<u>Proof of Theorem 3.2</u>: We note first of all that as $u \varepsilon \Omega_\Delta[0,\infty)$, Lemma 3.A.3 implies that both W and V must belong to this set as well. Thus the equivalence of the upper bounds in (3.6), (3.7), and (3.8) , (3.9) follows from the resultant boundedness of W(t) and V(t) .

Consider now the situation where in (3.7) , no matter what choice of δ is made, the lower bound is violated.

We shall establish a contradiction. The flow of the proof in outline is : Lower bound in (3.7) is violated => $\xi^T V(t)$ is small for some ξ over an interval => (recognizing each entry of V(t) is a linear functional of $u(\tau)$, $\tau < t$) a linear combination of causal functionals of u(.) is small over an interval => $\eta^T W(t)$ is small over an interval for some η .

In the process, we must be concerned with initial conditions (which decay because of the stability assumption) and with showing that $\|\eta\|$ cannot be vanishingly small.

Violation of the lower bound in (3.7) means that for arbitrary $\varepsilon > 0$, and arbitrary δ there exists a particular σ and a particular unit length vector $\xi^T = [\gamma_0,\ldots,\gamma_{n-1},\theta_0,\ldots,\theta_m]$, with σ and ξ both depending on ε,δ , such that

$$f(t) = \int_\sigma^t \{\xi^T V(\tau)\}^2 d\tau < \varepsilon^4 \quad \forall t \varepsilon [\sigma,\sigma+\delta] .$$

Now

$$\dot{f}(t) = \{\xi^T V(t)\}^2$$

and

$$\ddot{f}(t) = \frac{d}{dt}\{\xi^T V(t)\}^2$$

Below, we shall require $\delta > \bar{\delta}_i$ $i=1,2$ where $\bar{\delta}_1$ is a certain constant associated with the decaying of initial condition effects and ε , and $\bar{\delta}_2$ depends on ε, n, m and, ultimately, bounds on V ; details are given below. We shall also require that $\delta > \Delta$.

Now, as $V(t) \varepsilon \Omega_\Delta[0,\infty)$, $|\ddot{f}(t)| < C$ on $[\sigma, \sigma+\delta] - C_\Delta$ for some finite positive constant C . Moreover, if ε^4 is chosen to be small enough to ensure that $2\varepsilon^4 < C\Delta^2$ [or that $M_n' = \max(C, \varepsilon^4 2! \Delta^{-2}) = C$ holds] an application of Lemma 3.A.1 to any subinterval of length Δ in $[\sigma, \sigma+\delta] - C_\Delta$ ensures that on this subinterval,

$$\{\xi^T V\}^2 < 8e^2\varepsilon^2\sqrt{C}$$

Now $\xi^T V$ has one sided limits approaching the points belonging to C_Δ , and so

$$|\xi^T V| < 0(\varepsilon) \text{ on } [\sigma, \sigma+\delta] \quad .$$

Now $y = A^{-1}(s)B(s)u$ + initial condition effects due to any nonzero initial condition at $t = 0$. Hence

$$\left| \frac{\sum_{i=0}^{n-1} \gamma_i s^i [\sum_{k=0}^{m} b_k s^k] u}{(s+\alpha)^n (s^n + \sum_{k=0}^{n-1} a_k s^k)} + \frac{\sum_{j=0}^{m} \theta_j s^j u}{(s+\alpha)^n} + i(t) \right| < 0(\varepsilon)$$

on $[\sigma, \sigma+\delta]$. Here, $i(t)$ denotes initial condition effects.

Let us make the assumption that all initial conditions for the plant lie in some finite ball. Then there exists a $\bar{\delta}_i$ depending on ε, but independent of the γ_i, θ_j choice (note that these constants are bounded in magnitude by 1!), such that $|i(t)| = 0(\varepsilon)$ for all $t > \bar{\delta}_1$ and so for all $t > \sigma + \bar{\delta}_1$, no matter what σ is. Hence

$$\left| \frac{\sum_{i=0}^{n-1} \gamma_i s^i [\sum_{k=0}^{m} b_k s^k] u}{(s+\alpha)^n (s^n + \sum_{k=0}^{n-1} a_k s^k)} + \frac{\sum_{j=0}^{m} \theta_j s^j u}{(s+\alpha)^n} \right| < 0(\varepsilon) \qquad (3.A.3)$$

on $[\sigma + \bar{\delta}_1, \sigma + \delta]$

Now as a consequence of Lemma 3.A.3, it follows that the first $(n-m+1)$ derivatives of the quantity under the modulus sign in (3.A.3) are bounded in magnitude by some constant K_1. This bound can be taken to be independent of ξ, since $\|\xi\| = 1$. Now suppose that

$$\bar{\delta}_2 = \left[\frac{(n-m+1)!K_2\varepsilon}{K_1} \right]^{1/(n-m+1)}$$

where $K_2\varepsilon$ can be used on the right side of (3.A.3). Assume also that $\delta - \bar{\delta}_1 > \bar{\delta}_2$. Then Lemma 3.A.1 and the existence of one-sided limits of u and $\dot{u}$ imply after some manipulation that

$$\left| (s+\alpha)^{n-m} \left\{ \frac{\sum_{i=0}^{n-1} \gamma_i s^i [\sum_{k=0}^{m} b_k s^k] u}{(s+\alpha)^n (s^n + \sum_{k=0}^{n-1} a_k s^k)} + \frac{\sum_{j=0}^{m} \theta_j s^j u}{(s+\alpha)^n} \right\} \right| < 0(\varepsilon^{1/(n-m+1)})$$

(3.A.4)

on $[\sigma + \overline{\delta}_1, \sigma + \delta]$, i.e.

$$\left| \frac{\sum_{i=0}^{n-1} \gamma_i s^i (\sum_{k=0}^{m} b_k s^k) u}{(s+\alpha)^m (s^n + \sum_{k=0}^{n-1} a_k s^k)} + \frac{\sum_{j=0}^{m} \theta_j s^j u}{(s+\alpha)^m} \right| < 0(\varepsilon^{1/(n-m+1)}) \tag{3.A.5}$$

on $[\sigma+\overline{\delta}_1, \sigma+\delta]$.

Now the left side of (3.A.5) is bounded for all time. Hence, by Lemma 3.A.2 , we can postulate the existence of a $\overline{\delta}_3$, independent of σ , such that with $\delta > \overline{\delta}_1 + \overline{\delta}_3$,

$$\left| \frac{(s+\alpha)^m (s^n + \sum_{k=0}^{n-1} a_k s^k)}{(s+\beta)^{n+m}} \times [\text{contents of } |\ | \text{ in } (3.A.5)] \right|$$

$$< 0(\varepsilon^{1/(n-m+1)})$$

on $(\sigma+\overline{\delta}_1+\overline{\delta}_3, \sigma+\delta)$

i.e.

$$\left| \left\{ \frac{(\sum_{i=0}^{n-1} \gamma_i s^i)(\sum_{j=0}^{m} b_j s^j)}{(s+\beta)^{n+m}} + \frac{(\sum_{i=0}^{m} \theta_i s^i)(s^n + \sum_{k=0}^{n-1} a_k s^k)}{(s+\beta)^{n+m}} \right\} u \right|$$

$$< 0(\varepsilon^{1/(n-m+1)})$$

on $[\sigma + \overline{\delta}_1 + \overline{\delta}_3, \sigma + \delta]$.

Furthermore, defining $\overline{R}$ as

$$\overline{R} = \begin{bmatrix} R & 0 \\ a^T & 1 \end{bmatrix}$$

where $a^T = [\underbrace{0 \ldots 0}_{m} \quad \underbrace{a_0 \; a_1 \ldots a_{n-1}}_{n}]$

and R is the $(n+m) \times (n+m)$ resultant matrix

$$R = \left[\begin{array}{cccccccccc} b_0 & b_1 & \cdots & b_m & 0 & \cdots & & & & 0 \\ 0 & b_0 & b_1 & \cdots & & b_m & 0 & \cdots & & 0 \\ \vdots & & & & & & & & & \vdots \\ 0 & 0 & 0 & b_0 & \cdots & & & & & b_m \\ a_0 & a_1 & \cdots & & a_{n-1} & 1 & 0 & 0 & \cdots & 0 \\ 0 & a_0 & \cdots & & & a_{n-1} & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & & & & & \ddots & & \vdots \\ 0 & \cdots & 0 & a_0 & \cdots & & & & a_{n-1} & 1 \end{array}\right] \begin{array}{l} \left.\vphantom{\begin{array}{c} 0\\0\\0\\0 \end{array}}\right\} n \\ \left.\vphantom{\begin{array}{c} 0\\0\\0\\0 \end{array}}\right\} m \end{array}$$

we find that

$$\left(\sum_{i=0}^{n-1} \gamma_i s^i\right)\left(\sum_{j=0}^{m} b_j s^j\right) + \left(\sum_{i=0}^{m} \theta_i s^i\right)\left(s^n + \sum_{k=0}^{n-1} a_k s^k\right)$$

can be expressed as

$$\xi^T \overline{R} [1 \quad s \ \ldots \ s^{n+m}]^T \, .$$

Define β as $\beta = \overline{R}^T \xi$.

The coprimeness of $A(s)$ and $B(s)$ assures the nonsingularity of R and hence of $\overline{R}$. Thus, as $\| \xi \| = 1$ by hypothesis

$$\| \beta \| > 1/\| \overline{R}^{-1} \| = K_3 > 0$$

where K_3 depends on the a_i and b_i only. Hence (3.A.5) implies that

$$\int_{\sigma+\overline{\delta}_1+\overline{\delta}_3}^{\sigma+\delta} W(\tau)W^T(\tau)d\tau$$

cannot be uniformly positive definite, since ε is arbitrary. Defining $\delta' = \delta - \overline{\delta}_1 - \overline{\delta}_3$ and $\tau = t + \delta - \delta'$, we conclude the lower bound in (3.6) is also violated. Hence if (3.6) holds, (3.7) must hold.

The proof of theorem 3.2 in the opposite direction proceeds in much the same way. Here we shall simply outline the steps. Assume that (3.9) fails. Then for arbitrary $\varepsilon > 0$ there is a λ, $\|\lambda\| = 1$, with $\lambda^T W$ of $0(\varepsilon)$ on some $[\sigma, \sigma + \delta]$. Then

$$\left| \frac{(s+\beta)^{n+m}}{(s+\alpha)^n (s^n + \sum_{i=0}^{n-1} a_i s^i)} \lambda^T W \right| = \left| \sum_{i=0}^{n+m} \frac{(s+\beta)^{n+m-i}\lambda_i}{(s+\alpha)^n (s^n + \sum_{i=0}^{n-1} a_i s^i)} u \right|$$

is small over

$(\sigma+\overline{\delta}_1, \sigma+\delta)$ for some $\overline{\delta}_1$. [The stability of $(s^n + \sum_{i=0}^{n-1} a_i s^{n-1})$ is crucial here.] Now, using the coprimeness of

$\sum_{j=0}^{m} b_j s^j$ and $s^n + \sum_{i=0}^{n-1} a_i s^i$, define θ_i , $i=0,\ldots,n+m$,

by

$$\sum_{i=0}^{n+m} (s+\beta)^{n+m-i} \lambda_i = \left(\sum_{i=0}^{n-1} \theta_i s^i\right)\left(\sum_{j=0}^{m} b_j s^j\right)$$

$$+ \left(\sum_{i=0}^{m} \theta_{n+i} s^i\right)\left(s^n + \sum_{i=0}^{n-1} a_i s^i\right)$$

Then we find that

$$\left| \frac{\sum_{i=0}^{n-1} \theta_i s^i}{(s+\alpha)^n} y + \frac{\sum_{i=0}^{m} \theta_{n+i} s^i}{(s+\alpha)^n} u \right| < 0(\varepsilon)$$

on $[\sigma+\overline{\delta}_1 + \overline{\delta}_3, \sigma + \delta]$. Failure of (3.8) is then immediate.

APPENDIX 3.B : PROOF OF THEOREM 3.4

Suppose that the conclusion is not valid. Then for any $\varepsilon > 0$ and any $\delta > \delta_1(\varepsilon) + \delta_2(\varepsilon)$ (as defined in Lemma 3.1), we can find σ such that for a certain ξ with $\|\xi\| = 1$,

$$\int_{\sigma}^{\sigma+\delta} [\xi^T \bar{W}(t)]^2 dt < \varepsilon^4$$

while (3.17) holds. Following the same argument as for Theorem 3.2, we conclude that on $[\sigma, \sigma+\delta]$,

$$|\xi^T \bar{W}(t)| < 0(\varepsilon)$$

and so

$$\left| \sum_{i=0}^{\bar{n}+m} \frac{(s+\beta)^{\bar{n}+m-i}\xi_i}{(s+\beta)^{\bar{n}+m}} u \right| < 0(\varepsilon)$$

where $\bar{n} = n-p$.

Recognize that $\frac{(s+\beta)^m}{A(s)} u$ can be regarded as a linear functional of the state of the plant assuming zero initial condition. It follows by Lemma 3.1 that, with the λ_i satisfying conditions as specified in the Lemma,

$$\left| \sum_{i=0}^{n+m} \frac{(s+\beta)^{\bar{n}+m-i}\xi_i}{(s+\beta)^{\bar{n}+m}} \frac{(s+\beta)^m}{A(s)} u - \sum_{i=1}^{n-\bar{n}} \lambda_i e^{j\omega_i t} \right| < 0(\varepsilon)$$

on $(\sigma + \delta_1, \sigma + \delta - \delta_2)$. There is a monic polynomial of degree $n-\bar{n}$ with zeros $j\omega_i$, $i=1,\ldots,n-\bar{n}$. Call it $q(s)$. Then

$$\left| (s+\beta)^{\bar{n}} \frac{q(s)}{(s+\alpha)^{n}} \left[\sum_{i=0}^{\bar{n}+m} \frac{(s+\beta)^{\bar{n}+m-i}\xi_i}{(s+\beta)^{\bar{n}+m}} \frac{(s+\beta)^{m}}{A(s)} u - \sum_{i=1}^{n-\bar{n}} \lambda_i e^{j\omega_i t} \right] \right| < 0(\varepsilon)$$

i.e.

$$\left| \sum_{i=0}^{\bar{n}+m} \frac{(s+\beta)^{\bar{n}+m-i}\xi_i q(s)}{(s+\alpha)^{n} A(s)} u \right| < 0(\varepsilon)$$

Now find θ_i, $i=0,\ldots,n+m$, such that

$$\sum_{i=0}^{\bar{n}+m} (s+\beta)^{\bar{n}+m-i} \xi_i q(s) = \left(\sum_{i=0}^{n-1} \theta_i s^i\right) B(s) + \left(\sum_{i=0}^{m} \theta_{n+i} s^i\right) A(s)$$

(This is possible because A and B are coprime.)

Then we have

$$\left| \frac{\sum_{i=0}^{n-1} \theta_i s^i}{(s+\alpha)^n} \frac{B(s)}{A(s)} u + \frac{\sum_{i=0}^{m} \theta_{n+i} s^i}{(s+\alpha)^n} u \right| < 0(\varepsilon)$$

or $|[\theta_{n-1}\ \theta_{n-2}\ \cdots\ \theta_0\ \ \theta_{n+m}\ \cdots\ \theta_n] V(t)| < 0(\varepsilon)$ on $[\sigma+\delta_1, \sigma+\delta-\delta_2]$. This contradicts (3.17) and the theorem is established.

APPENDIX 3.C.

Proofs of the theorems in section 3.3 require the following lemmata.

Lemma 3.C.1

Consider an asymptotically stable SISO system with proper transfer function $T(s)$, arbitrary finite initial conditions and bounded input $v(t)$. For any $\varepsilon > 0$ define $\delta_1(\varepsilon)$ by

$$e^{-\lambda \delta_1(\varepsilon)} = \varepsilon$$

where λ is determined by the poles of $T(s)$. Suppose for some $\sigma > 0$ and $\delta > \delta_1(\varepsilon)$

$$\int_{\sigma}^{\sigma+\delta} |v(t)|dt < \varepsilon$$

Then with $w(t)$ defined as the output

$$\int_{\sigma+\delta_1(\varepsilon)}^{\sigma+\delta} |w(t)|dt < 0(\varepsilon)$$

Proof

Follows on the same lines as those in the proof of lemma 3.A.2.

$\nabla\nabla\nabla$

Note, that if $\{A,b,c,d\}$ is a minimal realization of $T(s)$, then λ is such that for some finite K

$$\| e^{At} \| \leq K e^{-\lambda t}$$

This applies to all situations where we talk of a λ "determined by the poles of an asymptotically stable transfer function".

Lemma 3.C.2

Let $v_1(t) \in \Omega_\Delta[0,\infty)$ and v_2, v_3 and v_4 be generated by (3.C.1 - 3.C.3), with arbitrary finite initial conditions

$$v_2(t) = \frac{B(s)}{A(s)} v_1(t) \tag{3.C.1}$$

$$v_3(t) = \sum_{i=1}^{\nu} \theta_i \frac{v_2(t)}{(s+\beta)^i} \tag{3.C.2}$$

$$v_4(t) = \frac{A(s)(s+\beta)^\nu}{A_1(s)} v_3(t) \tag{3.C.3}$$

where $\delta[A(s)] = n \geq \delta[B(s)]$, $\delta[A_1(s)] \geq n + \nu$, $A_1(s)$ is Hurwitz and θ_i are finite constants.

For any $\varepsilon > 0$ define $\delta'(\varepsilon)$ by

$$e^{-\lambda_2 \delta'(\varepsilon)} = \varepsilon$$

with λ_2 determined by the roots of $A_1(s)$. Suppose for some $\sigma > 0$ and $\delta > \delta'(\varepsilon)$

$$\int_\sigma^{\sigma+\delta} |v_3(t)|dt < \varepsilon \tag{3.C.4}$$

Then

$$\int_{\sigma+\delta'(\varepsilon)}^{\sigma+\delta} |v_4(t)|dt < 0(\varepsilon) \tag{3.C.5}$$

Remark:

The intuition behind this result is the following. Consider figure 3.C.1 which is a block diagram representation of (3.C.1 - 3.C.3). Let $v_3(t)$ be "small" in some sense over an interval $[\sigma,\sigma+\delta]$. As far as $v_4(t)$ is concerned it comprises two components : the initial condition component f_I , reflecting the history up to the instant σ and the forced component f_p , stemming from excitation beyond the time σ . While f_p is small, the same cannot be said about f_I as the latter depends on the past history of $v_3(t)$ and $v_3(t)$ maybe unbounded. The third operator block, however, exactly cancels the initial conditions coming from the blocks 1 and 2. Thus if δ' is the time required for the intial conditions due to block 3 to decay down to sufficiently small values then $v_4(t)$ would be small over an interval $[\sigma+\delta',\sigma+\delta]$.

Proof:

Rewriting (3.C.1) - (3.C.3) as

$$A(s)v_2(t) = B(s)v_1(t) \tag{3.C.6}$$

$$(s+\beta)^i \quad \phi_i(t) = v_2(t) \tag{3.C.7}$$

$$v_3(t) = \sum_{i=1}^{\nu} \theta_i \ \phi_i(t) \tag{3.C.8}$$

$$A_1(s)v_4(t) = A(s)(s+\beta)^{\nu} v_3(t) \tag{3.C.9}$$

it is easy to show that

$$A_1(s)v_4(t) = \sum_{i=1}^{\nu} \theta_i(s+\beta)^{\nu-i} B(s)v_1(t) \tag{3.C.10}$$

Let $\{F,G_1,H_1,J_1\}$ and $\{F,G_2,H_2,J_2\}$ be minimal realization of (3.C.9) and (3.C.10) respectively. Let χ_1,χ_2 denote the respective state vectors.Then as $A_1(s)$ is Hurwitz $\exists\ \lambda_1,\lambda_2 > 0$, such that

$$|| e^{F(t-\sigma)}|| \leq \lambda_1 e^{-\lambda_2(t-\sigma)}$$

Also

$$v_4(t) = H_1 e^{F(t-\sigma)}\chi_1(\sigma) + \int_\sigma^t H_1 e^{F(t-\tau)} G_1 v_3(\tau)d\tau + J_1 v_3(t)$$

$$\forall t \geq \sigma \qquad (3.C.11)$$

$$= H_2 e^{F(t-\sigma)}\chi_2(\sigma) + \int_\sigma^t H_2 e^{F(t-\tau)} G_2 v_1(\tau)d\tau + J_2 v_1(\tau)$$

$$\forall t \geq \sigma \qquad (3.C.12)$$

We shall now show the existence of a K_1 , independent of σ for which

$$|| H_1 e^{F(t-\sigma)}\chi_1(\sigma)|| \leq K_1 e^{-\lambda_2(t-\sigma)} \qquad (3.C.13)$$

The zero input response of (3.C.12) remains unaffected by the values of $v_1(t)$ beyond $t = \sigma$. Let $\bar{v}_4(t)$ be the value of $v_4(t)$ when $v_1(t) = 0$ $\forall t \geq \sigma$. Then

$$\bar{v}_4(t) = H_2 e^{F(t-\sigma)}\chi_2(\sigma)$$

$$= H_1 e^{F(t-\sigma)}\chi_1(\sigma) + \{\int_\sigma^t H_1 e^{F(t-\tau)} G_1 \bar{v}_3(\tau)d\tau +$$

$$J_1\bar{v}_3(t)\} \qquad (3.C.14)$$

where $\bar{v}_3(t)$ is the output of (3.C.8) whenever $v_1(t) = 0$, $\forall t \geq \sigma$. But $\bar{v}_3(t)$ lies at the zeros of (3.C.3) whence the term within parentheses in (3.C.14) is identically zero. Thus

$$H_1 e^{F(t-\sigma)} \chi_1(\sigma) = H_2 e^{F(t-\sigma)} \chi_2(\sigma)$$

Also, from (3.C.10) it is clear that $\chi_2(t)$ is bounded whence (3.C.13) follows. Hence with δ' such that $e^{-\lambda_2 \delta'} = \varepsilon$

$$\| H_1 e^{F(t-\sigma)} \chi_1(\sigma) \| \le 0(\varepsilon) \qquad \forall t \ge \sigma+\delta'$$

Thus

$$\int_{\sigma+\delta'}^{\sigma+\delta} |v_4(t)|dt \le \int_{\sigma+\delta'}^{\sigma+\delta} \| H_1 e^{F(t-\sigma)} \chi_1(\sigma) \| \, dt$$

$$+ \int_{\sigma+\delta'}^{\sigma+\delta} \int_{\sigma}^{t} \| H_1 e^{F_1(t-\tau)} G_1 v_3(\tau) \| \, d\tau \, dt$$

$$+ \int_{\sigma+\delta'}^{\sigma+\delta} \| J_1 v_3(t) \| \, dt$$

$$\le 0(\varepsilon) \quad \text{due to (3.C.4)} .$$

∇∇∇

The point of the next lemma is the following. Consider figure 3.C.2. Let

$$v_1(t) = \frac{\theta_1(s)}{(s+\alpha)^n} y(t) + \frac{\theta_2(s)}{(s+\alpha)^n} u(t)$$

be "small" over an interval $[\sigma, \sigma+\delta]$. Then

$$v_2(t) = \frac{A(s)(s+\alpha)^n}{A_1(s)} v_1(t)$$

$$= \frac{\theta_1(s)B(s)+\theta_2(s)A(s)}{A_1(s)} u(t)$$

is "small" over a smaller interval $[\sigma+\delta',\sigma+\delta]$ if

(i) $A_1(s)$ is Hurwitz and

(ii) the state $x_1(t)$ of any state variable realization of

$$\left\{\frac{\theta_1(s)\ B(s) + \theta_2(s)A(s)}{A_1(s)}\right\} \quad u(t)$$

is bounded by a constant M at $t = \sigma$. As before the last block serves to exactly cancel the initial condition effects generated in the preceding ones. The constant δ' is large enough to ensure that initial condition effects due to $\bar{x}_1(\sigma)$ decay by the time $\sigma+\delta'$. The proof of this lemma is similar to that of the previous one and is hence omitted.

Lemma 3.C.3.

Consider a proper time invariant system given by

$$A(s)y(t) = B(s)u(t) \tag{3.C.15}$$

with $A(s)$, $B(s)$ polynomials with degrees n and m respectively. Consider also the Hurwitz polynomial $A_1(s)$, $\delta[A_1(s)] \geq 2n$ and the differential equations

$$(s+\alpha)^n v_{11}(t) = \theta_1(s)y(t) \tag{3.C.16}$$

$$(s+\alpha)^n v_{12}(t) = \theta_2(s)u(t) \tag{3.C.17}$$

$$A_1(s)v_2(t) = A(s)(s+\alpha)^n[v_{11}(t)+v_{12}(t)] \tag{3.C.18}$$

for some polynomials $\theta_1(s)$, $\theta_2(s)$; $\delta[\theta_1(s)] \leq n-1$ and $\delta[\theta_2(s)] \leq m$. In the above systems (3.C.15 - 3.C.18), all initial conditions are arbitrary but finite.

From (3.C.15 - 3.C.18)

$$v_2(t) = \{\frac{\theta_1(s)B(s) + \theta_2(s)A(s)}{A_1(s)}\}u(t) \qquad (3.C.19)$$

Let $x_1(t)$ be the state vector in any minimal state variable realization of (3.C.19) and let it satisfy the boundary condition

$$\|x_1(\sigma)\| < M .$$

For any $\varepsilon > 0$ define $\delta'(\varepsilon)$ by

$$e^{-\lambda\delta'(\varepsilon)} = \varepsilon$$

with λ determined by the zeros of $A_1(s)$. Suppose for some $\sigma > 0$ and $\delta > \delta'(\varepsilon)$

$$\int_{\sigma}^{\sigma+\delta} |v_1(t)|dt < \varepsilon \qquad (3.C.20)$$

where

$$v_1(t) = v_{11}(t) + v_{12}(t)$$

Then

$$\int_{\sigma+\delta'(\varepsilon)}^{\sigma+\delta} |v_2(t)|dt < 0(\varepsilon) \qquad (3.C.21)$$

∇∇∇

We are now in a position to prove Theorems 3.6 - 3.8.

Proof of Theorem 3.6

Let (3.31) be violated. Then for arbitrary $\varepsilon > 0$, $\exists$ a unit vector $\theta^T = [\theta_1, \ldots, \theta_\nu]$ such that for some σ

$$\int_\sigma^{\sigma+\delta} |\theta^T V(t)_{n+m+1}|^2 dt < \varepsilon^2$$

By Schwarz's inequality

$$\int_\sigma^{\sigma+\delta} |\theta^T V(t)_{n+m+1}| dt \leq \delta^{\frac{1}{2}} \{\int_\sigma^{\sigma+\delta} |\theta^T V(t)_{n+m+1}|^2 dt\}^{\frac{1}{2}} \leq \delta^{\frac{1}{2}} \varepsilon \quad (3.C.22)$$

Then

$$\int_\sigma^{\sigma+\delta} \left| \frac{\sum_{i=0}^{n-1} \theta_i s \quad y(t)}{(s+\alpha)^n} + \frac{\sum_{j=0}^{m} \theta_{j+n} \, s^j u(t)}{(s+\alpha)^n} \right| dt < 0(\varepsilon)$$

Let the signal within the modulus signs bé $v_1(t)$. Then defining $v_2(t)$ as

$$v_2(t) = \frac{(s+\alpha)^n}{(s+\beta)^{2n}} A(s)$$

we have that

$$v_2(t) = \frac{(\sum_{i=0}^{n-1} \theta_i s^i) B(s) + (\sum_{j=0}^{m} \theta_{j+n} s^j) A(s)}{(s+\beta)^{2n}} u(t) \quad (3.C.23)$$

Now since $u(t) \in \Omega_\Delta[0,\infty)$ the state vector in any minimal realization of (3.C.23) is finite. Thus from Lemma 3.C.3 we have that $\exists$ a $\bar{\delta} < \delta$ such that

$$\int_{\sigma+\bar{\delta}}^{\sigma+\delta} \left| \frac{\left(\sum_{i=0}^{n-1} \theta_i s^i\right) B(s) + \left(\sum_{j=0}^{m} \theta_{j+n} s^j\right) A(s)}{(s+\beta)^{2n}} u(t) \right| dt < 0(\varepsilon)$$

Now, since $A(s)$ and $B(s)$ are coprime, by arguments similar to that used in the proof of Theorem 3.2, we have that there exists a $\gamma = [\gamma_1 \ldots \gamma_{n+m}]^T$, $\|\gamma\|$ bounded away from zero, for which

$$\int_{\sigma+\bar{\delta}}^{\sigma+\delta} \left| \frac{\sum_{i=0}^{n+m} \gamma_i (s+\beta)^i}{(s+\beta)^{2n}} u(t) \right| dt < 0(\varepsilon) \tag{3.C.24}$$

As $\beta > 0$, the term within the modulus sign in (3.C.24) is bounded and hence Lemma 3.A.1 is applicable. Thus as in theorem 3.2

$$\left| \sum_{i=0}^{n+m} \frac{\gamma_1 (s+\beta)^i}{(s+\beta)^{2n+m}} u(t) \right| < 0\left(\varepsilon^{1/(2(n-m+1))}\right) \tag{3.C.25}$$

whence (3.30) is violated. Thus the result follows.

∇∇∇

Proof of Theorem 3.7

Let (3.34) be violated. Then for arbitrary $\varepsilon > 0$, $\exists$ a unit vector $\theta^T = [\theta_1, \ldots, \theta_\nu]^T$ such that for some σ

$$\int_{\sigma}^{\sigma+\delta} |\theta^T Y_\nu(t)| dt < \delta^{\frac{1}{2}}\varepsilon$$

=>

$$\int_{\sigma}^{\sigma+\delta} \left| \left(\sum_{i=0}^{\nu} \frac{\theta_i}{(s+\beta)^i} Y_\nu(t) \right| dt \le \delta^{\frac{1}{2}}\varepsilon \qquad (3.C.26)$$

Let the quantity within the modulus signs in (3.C.26) be $v_3(t)$. Consider $v_4(t)$ defined by

$$v_4(t) = \frac{(s+\beta)^\nu A(s)}{(s+\alpha)^{\nu+n}} v_3(t)$$

Thus by Lemma 3.C.2

$$\int_{\sigma+\bar{\delta}}^{\sigma+\delta} \left| \left\{ \frac{\sum_{i=0}^{\nu} \theta_i (s+\beta)^i B(s)}{(s+\beta)^{\nu+n}} \right\} u \right| dt < 0(\varepsilon) \qquad (3.C.27)$$

for some $\bar{\delta}$ depending on α . Let $B(s) = B_s(s)B_o(s)B_+(s)$ where B_s is Hurwitz, $B_o(s)$ has zeros on the imaginary axis and $B_+(s)$ has zeros in the open right half plane.

Then by Lemma 3.C.1 $\exists$ a $\bar{\delta} < \delta$ such that

$$\int_{\sigma+\bar{\delta}}^{\sigma+\delta} \left| \frac{(s+\alpha)^{m-z-z_+}}{B_s(s)} \frac{\sum_{i=0}^{\nu} \theta_i (s+\beta)^i B(s)}{(s+\alpha)^{\nu+n}} u(t) \right| dt < 0(\varepsilon)$$

(3.C.28)

where $\delta[B_o(s)] = z$ and $\delta[B_+(s)] = z_+$.

Thus

$$\int_{\sigma+\bar{\delta}}^{\sigma+\delta} \left| \frac{\sum_{i=0}^{\nu} \theta_i (s+\beta)^i \, B_o(s) \, B_+(s)}{(s+\alpha)^{\nu+n-m+z+z_+}} u(t) \right| dt < 0(\varepsilon) \qquad (3.C.29)$$

The term within the modulus signs in (3.C.29) is bounded. Thus Lemma 3.A.1 implies

$$\left| \frac{\sum_{i=0}^{\nu} \theta_i (s+\beta)^i \, B_o(s) \, B_+(s)}{(s+\alpha)^{\nu+n-m+z+z_+}} u(t) \right| dt < 0(\varepsilon^{\frac{1}{2}})$$

whence by the repeated application of Lemma 3.2 we have

$$\left| \frac{\sum_{i=0}^{\nu} \theta_i (s+\beta)^i \, B_o(s)}{(s+\alpha)^{\nu+n-m+z+z_+}} u(t) \right| dt < 0(\varepsilon^{\frac{1}{2}})$$

Thus following on the lines of the proof of the previous theorem we have that the result follows.

∇∇∇

Proof of Theorem 3.8

As before the violation of the implied inequality means that $\exists$ a unit θ and a σ such that

$$\int_{\sigma}^{\sigma+\delta} \left| \frac{\sum_{i=0}^{n-1} \theta_i s^i}{(s+\alpha)^n} y(t) + \frac{\sum_{j=0}^{m} \theta_{j+n} \, s^j}{(s+\alpha)^n} u(t) \right| dt < 0\,(\varepsilon)$$

Consider the filtering of the quantity within the modulus signs through $(s+\alpha)^n A(s)/(s+\beta)^{2n}$. Let $\mathbf{x}(t)$ be any state vector in a minimal realization of

$$G(s) = \frac{(\sum_{i=0}^{n-1} \theta_i s^i)\ B(s) + (\sum_{j=0}^{m} \theta_{j+m}\ s^j)\ A(s)}{(s+\beta)^{2n}}$$

obeying $\| x(\sigma) \| < M$. Then by Lemma 3.C.3 there exists $\overline{\delta} < \delta$ such that

$$\int_{\sigma+\overline{\delta}}^{\sigma+\delta} \left| G(s)\ u(t) \right| dt < 0(\varepsilon)$$

whence by the coprimineṣs assumption $\exists$ a vector γ bounded away from zero in modulus such that

$$\int_{\sigma+\overline{\delta}}^{\sigma+\delta} \left| \gamma^T\ W(t) \right| dt \le 0(\varepsilon)$$

$\forall\ W(\cdot) \in W^{(\sigma,M)}$

Thus the result follows. $\nabla\nabla\nabla$

APPENDIX 3.D : PROOF OF THEOREM 3.9

To prove theorem 3.9 we need the following lemma, which follows the same rationale as that behind the lemmata 3.C.2 and 3.C.3 . The proof runs similarly to that of lemma 3.C.2 and is therefore omitted.

Lemma 3.D.1

Consider the n-dimensional, linear, time-invariant multi input - single output system

$$\begin{aligned} \dot{x}_1 &= Ax_1(t) + I_n \Gamma(t) \\ y_1(t) &= c^T x_1(t) \end{aligned} \tag{3.D.1}$$

$\Gamma(t)$ bounded and the SISO system

$$\begin{aligned} \dot{x}_2 &= Ax_2(t) + b\gamma(t) \\ y_2 &= c^T x_2 + d\gamma(t) \end{aligned} \tag{3.D.2}$$

$\gamma(t)$ bounded, and a signal $y_3(t)$ such that $|y_3(t)| \leq K, \forall t \in R_+$. Consider also $v(t)$ given by

$$v(t) = \frac{\det(sI-A)}{(s+\alpha)^{\nu+n}} \theta(s) \{y_1(t)+y_2(t)+y_3(t)\} \tag{3.D.3}$$

where $\alpha > 0$ and $\theta(s)$ is a polynomial with degree less than or equal to ν .

For any $\varepsilon > 0$ define $\delta'(\varepsilon)$ by

$$e^{-\lambda\delta'(\varepsilon)} = \varepsilon$$

λ determined by the zeros of $(s+\alpha)$.

Suppose for some $\sigma > 0$, $\delta > \delta'(\varepsilon)$, and for arbitrary finite initial conditions for the system within modulus signs,

$$\int_{\sigma}^{\sigma+\delta} \left|\frac{\theta(s)}{(s+\alpha)^{\nu}}\{y_1(t)+y_2(t)+y_3(t)\}\right|dt < \varepsilon \qquad (3.D.4)$$

Then

$$\int_{\sigma+\delta'(\varepsilon)}^{\sigma+\delta} |v(t)|dt < 0(\varepsilon) \qquad (3.D.5)$$

∇∇∇

<u>Proof of Theorem 3.9</u>

Let (3.38b) hold. Thus

$$\int_{\sigma}^{\sigma+\delta} \left|\frac{\sum_{i=0}^{\nu} \theta_i (s+\beta)^{\nu-i}}{(s+\beta)^{\nu}} u(t)\right| dt < 0(\varepsilon) \qquad (3.D.6)$$

Thus using $(s+\beta)^{\nu}/(s+\alpha)^{\nu}$ as $T(s)$ in Lemma 3.C.1 we have that for some $\bar{\delta} < \delta$

$$\int_{\sigma+\bar{\delta}}^{\sigma+\delta} \left|\frac{\sum_{i=0}^{\nu} \theta_i (s+\beta)^{\nu-i}}{(s+\alpha)^{\nu}} u(t)\right| dt < 0(\varepsilon) \qquad (3.D.7)$$

Let σ' be a point in $[\sigma+\bar{\delta},\sigma+\delta]$ for which

$$\| c^T(\sigma') \| \, \| b(\sigma') \| + |d(\sigma')| > m_1 \qquad (3.D.8)$$

By hypothesis ∃ such a σ' . Equation (3.37) can be re-written as

$$\dot{x}(t) = A(\sigma')x(t)+b(\sigma')r(t)+\Delta b(t)r(t)+\Delta A(t)x(t)$$

$$(3.D.9)$$

$$u(t) = c^T(\sigma')x(t)+d(\sigma')r(t)+\Delta c^T(t)x(t)+\Delta d(t)r(t)$$

where $\Delta A(t) = A(t)-A(\sigma')$ and Δb, Δc and Δd are similarly defined. Then $u(t)$ can be expressed as $y_1(t)+y_2(t)+y_3(t)$ where the $y_i(t)$ are defined as follows.

$$\begin{aligned} \dot{x}_1(t) &= A(\sigma')x(t) + I_n\{\Delta b(t)r(t) + \Delta A(t)x(t)\} \\ y_1(t) &= c^T(\sigma')x_1(t) \end{aligned} \tag{3.D.10}$$

$$\begin{aligned} \dot{x}_2(t) &= A(\sigma')x_2(t) + b(\sigma')r(t) \\ y_2(t) &= c^T(\sigma')x_2(t) + d(\sigma')r(t) \end{aligned} \tag{3.D.11}$$

and

$$y_3(t) = \Delta c^T(t)x(t) + \Delta d(t)r(t) .$$

Now, by associating $A(\sigma')$, $b(\sigma')$, $c(\sigma')$, $d(\sigma')$, $\Delta b(t)r(t)+\Delta A(t)x(t)$ and $r(t)$ with $A,b,c,d,\Gamma(t)$ and $\gamma(t)$ respectively, we find that the assumptions of Lemma 3.D.1 hold.

Thus $\exists$ a δ' such that

$$\int_{\sigma+\delta'}^{\sigma+\delta} \left| \frac{\sum_{i=0}^{\nu} \theta_i (s+\beta)^i}{(s+\alpha)^{\nu+n}} \xi(s)r(t) + f_1(t)+f_2(t) \right| dt < 0(\varepsilon)$$

where

$$\xi(s) = c^T(\sigma')\mathrm{Adj}(sI-A(\sigma'))b(\sigma') \ ,$$

$$f_1(t) = \left\{\frac{\sum_{i=0}^{\nu} \theta_i (s+\beta)^i}{(s+\alpha)^{\nu+n}}\right\} c^T(\sigma')\mathrm{Adj}(sI-A(\sigma')) \; I_n\{\Delta b(t)r(t)+\Delta A(t)x(t)\}$$

and

$$f_2(t) = \frac{\det(sI-A(\sigma'))\{\sum_{i=0}^{\nu} \theta_i (s+\beta)^i\}}{(s+\alpha)^{\nu+n}} \{\Delta c^T(t)x(t)+\Delta d(t)r(t)\}$$

Now if ξ_i are the coefficients of $\xi(s)$, then (3.D.8) ensures that $\|\xi\|$ is bounded away from zero where

$$\xi^T = [\xi_0 \ \dots \ \xi_{n-1}] \ .$$

Thus the result follows.

∇∇∇

APPENDIX 3.E : PROOF OF THEOREM 3.10

(a) We note first of all that from (3.60) and (3.61) the exponential stability of $\bar{k}$ implies the exponential stability of k and $\bar{e}$.

(b) Since the quantities $\phi, \theta, \bar{x}$ etc. are related indirectly, through stable, (strictly) proper transfer functions, to $r(t)$ it follows that they too must belong to the set $\Omega_{\Delta}[0,\infty)$.

Also, from (3.59, 3.60) we have that

$$\dot{\bar{k}} = \frac{Q_1 \bar{\phi}(\bar{\phi}^T \bar{k} g_p + \varepsilon(t))}{\phi^T Q \phi + \lambda_0}$$

which, due to the boundedness of $\phi^T Q \phi + \lambda_0$ and theorem 3.1, is exponentially asymptotically stable if there exist positive α_{14}, α_{13} and large enough positive $\bar{\Delta}$ such that :

$$\alpha_{13} I < \int_{\sigma}^{\sigma+\bar{\Delta}} \overline{\phi\phi}^T dt < \alpha_{14} I \tag{3.E.1}$$

for all $\sigma \varepsilon R_+$.

(c) We now show that (3.62) implies (3.E.1) . As before, the equivalence of the upperbounds follows as a consequence of the boundedness of the quantities in question. Consider then the violation of the lower bound in (3.E.1). Then, since $\bar{\phi} \varepsilon \; \Omega_{\Delta}[0,\infty)$, arguing as in the proof of theorem 3.2, there exists, for any arbitrary

$\varepsilon > 0$, a unit vector γ such that

$$|\gamma^T\bar{\phi}| < \varepsilon \quad \text{on} \quad [\sigma, \sigma+\bar{\Delta}]$$

Thus

$$\left|c^T(sI-A)^{-1}b\gamma^T\begin{bmatrix} I_{2n} & 0 \\ 0 & 0 \end{bmatrix}\theta\right| < \varepsilon \quad \text{on} \quad [\sigma+\nu, \sigma+\bar{\Delta}]$$

or

$$\left|\frac{1}{\beta_r}\gamma^T[\theta_u^T \quad \theta_y^T]^T\right| < \varepsilon \quad \text{on} \quad [\sigma+\nu, \sigma+\bar{\Delta}]$$

where ν is some positive number independent of σ. Thus as $\theta\varepsilon\Omega_\Delta[0,\infty)$ and as β_r has degree n^*, we have, by Lemma 3.A.1, that

$$|\gamma^T[\theta_u^T \quad \theta_y^T]^T| < 0(\varepsilon^{1/n^*}) \quad \text{on} \quad [\sigma+\nu, \sigma+\bar{\Delta}] .$$

Hence from (3.39), (3.49) and (3.53) and the stability of T_p and $\det(sI-A_0)$ we have that

$$\left|[\gamma_u^T + \frac{g_p\alpha_p(s)}{\beta_p(s)}\gamma_y^T](sI-A_o)^{-1}b_0u\right| < 0(\varepsilon^{1/n^*})$$

where $\gamma = [\gamma_u^T \; \gamma_y^T]$.

Now, α_p and β_p are coprime, β_p has degree n and $\gamma_u^T(sI-A_0)^{-1}b_0$ and $\gamma_y^T(sI-A_0)^{-1}b_0$ have numerator polynomials of degree at most $n-1$. Thus

$$(\gamma_u^T + g_p\frac{\alpha_p(s)}{\beta_p(s)}\gamma_y^T)(sI-A_0)^{-1}b_0 \not\equiv 0$$

as otherwise α_p/β_p can be represented by a rational function of degree less than n .

Thus, as in the proof of theorem 3.2, a combination of lemmas 3.A.1 and 3.A.2 leads us to infer the existence of a Δ' independent of σ and a vector $\eta \triangleq [\eta_0, \ldots, \eta_{2n-1}]^T$ such that

$$\left| \frac{\sum_{i=0}^{2n-1} \eta_i (s+\bar{\beta})^i}{(s+\bar{\beta})^{2n-1}} u \right| < 0(\varepsilon^{1/2n^*}) \quad \text{on} \quad [\sigma+\Delta', \sigma+\bar{\Delta}] \qquad (3.E.2)$$

where $\|\eta\|$ is uniformly bounded away from zero.

If we denote the quantity within the modulus sign in (3.E.2) by $w(s)u$, then (3.E.2), (3.39), (3.49) and (3.53) together with a trivial modification of lemma 3.A.2, imply that there exists a $\Delta'' > \Delta'$ such that

$$\| w(s) I [\theta_u^T , \theta_y^T]^T \| \leq 0(\varepsilon^{1/2n^*}) \quad \text{on} \quad [\sigma+\Delta'', \sigma+\bar{\Delta}] .$$

Moreover, we know from proposition 3.1 that $\| \hat{k}_u^T, \hat{k}_y^T \|$ is bounded. Rewriting (3.50) as

$$u = \hat{k}_u^T \theta_u + \hat{k}_y^T \theta_y + (1/g_p)\theta_r$$

a modification of Theorem 3.9 and the decaying of $\dot{\hat{k}}$ show that for a large enough Δ'''

$$|w(s)(u - \frac{\theta_r}{g_p})| = |w(s)(\hat{k}_u^T \theta_u + \hat{k}_y^T \theta_y)| \leq 0(\varepsilon^{1/2n^*}) \quad \text{on}$$

$[\sigma+\Delta''', \sigma+\bar{\Delta}]$ whence

$$|w(s)\theta_r| \leq 0(\varepsilon^{1/2n^*}) .$$

Thus, if $\theta_r \triangleq [\theta_r, \ldots \theta_r/(s+\beta)^{2n-1}]^T$ then following the reasoning employed in proving the previous theorems, if there exist positive α_{15} and $\bar{\bar{\Delta}}$ such that

$$\alpha_{15} I < \int_{\sigma}^{\sigma+\bar{\bar{\Delta}}} \theta_r \theta_r^T dt \qquad (3.E.3)$$

for all $\sigma \varepsilon R_+$, then the lower bound of (3.E.1) must hold. Now as θ_r is related to r by a stable transfer function having no more than n_r imaginary axis zeros, a direct application of theorem 3.7 shows that the lower bound of (3.62) implies (3.E.3) and hence the lower bound of (3.E.1) . Thus (3.62) must ensure the exponential stability of the scheme in question.

∇∇∇

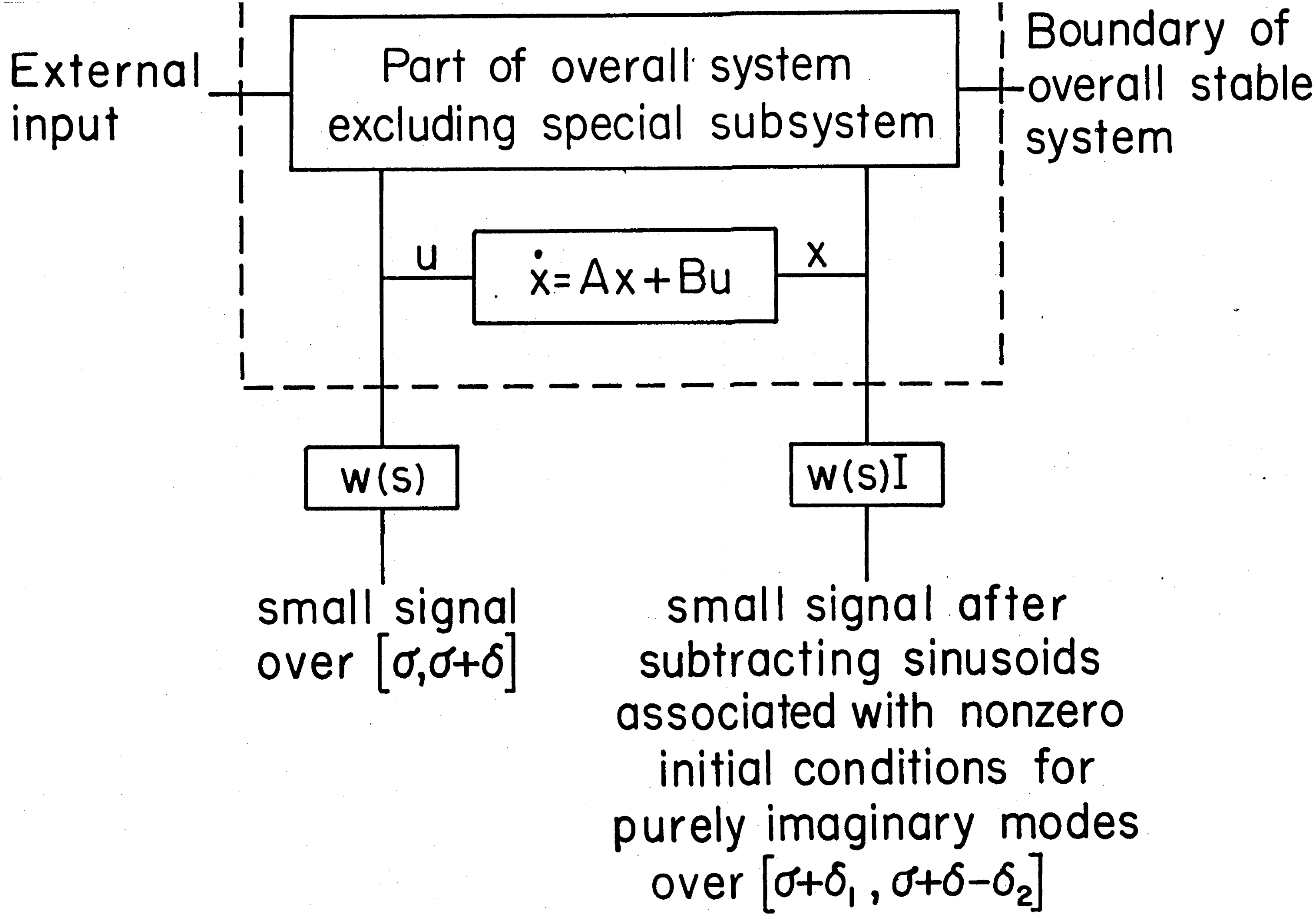


Illustration of Lemma 3.1

Figure 3.1

RJG 10-2-83

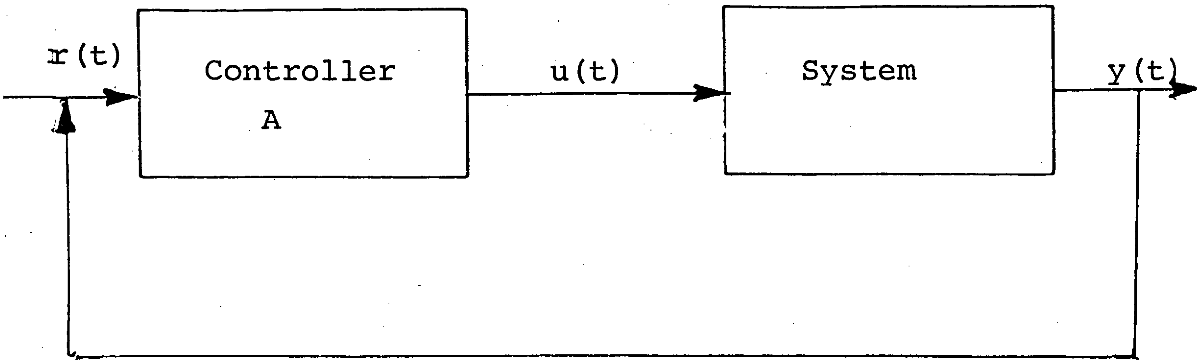


Figure 3.2 System with controller.

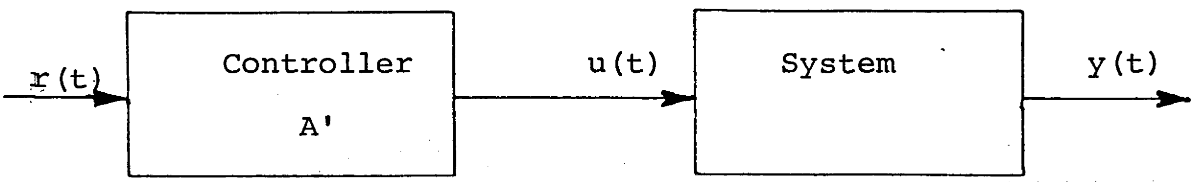


Figure 3.3 An equivalent representation of fig. 3.2 with block A' possibly nonminimal.

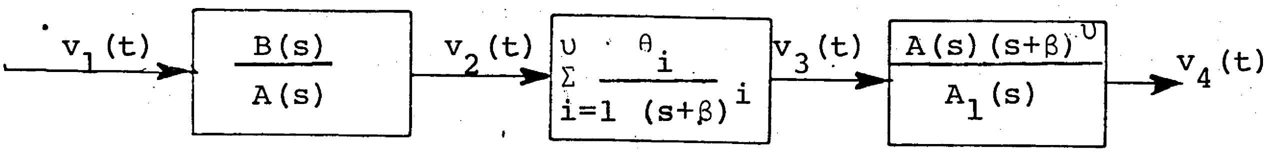


Figure 3.C.1. Representation of equations (3.C.1 - 3.C.3)

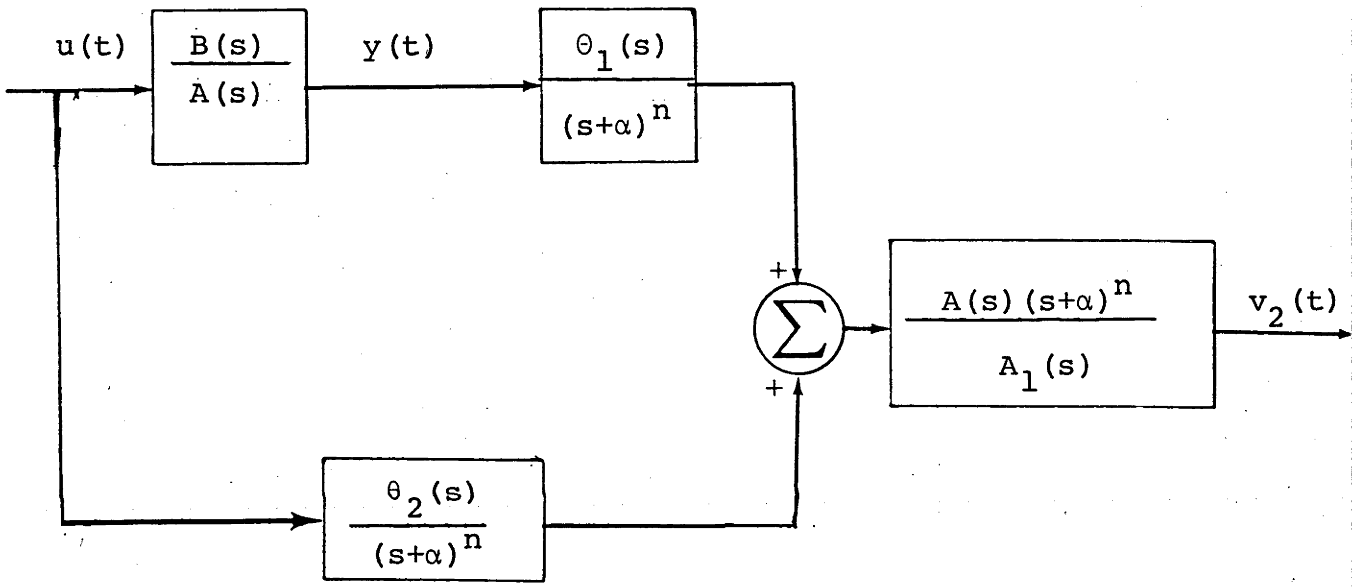


Figure 3.C.2 Representations of equations (3.C.15-3.C.19)

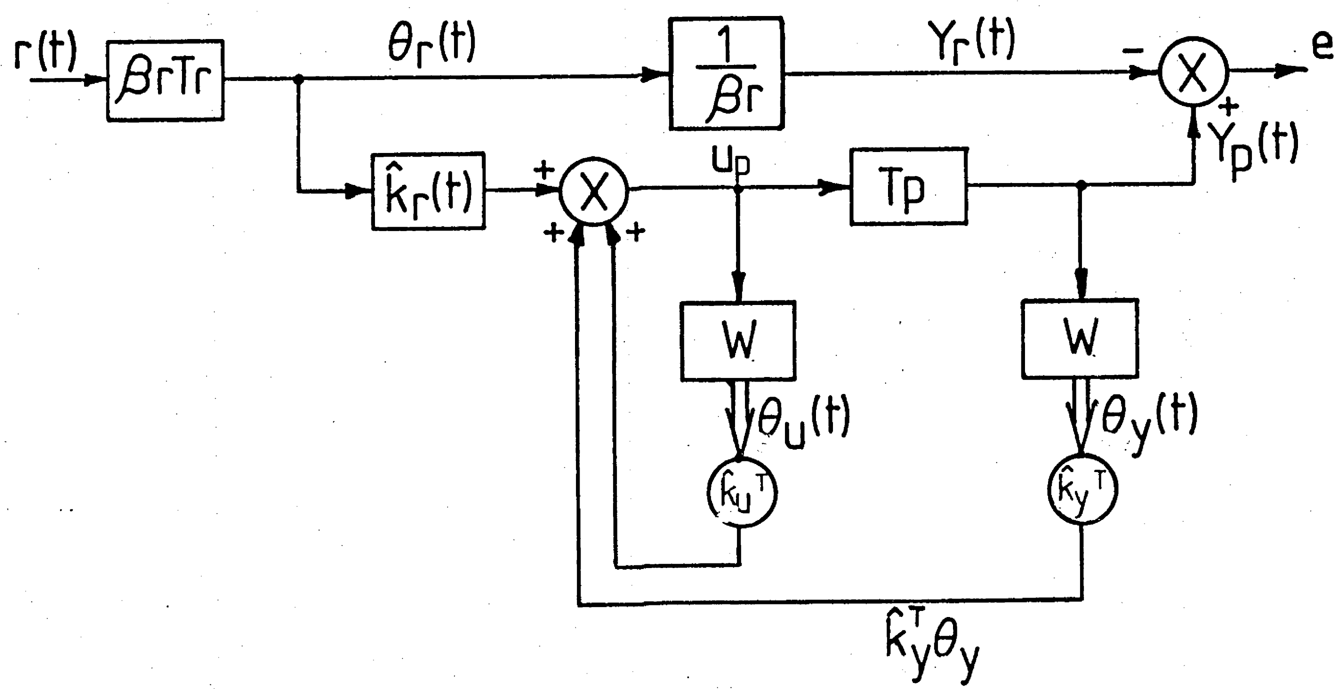


Figure 3.4 A block diagram representation of Morse's algorithm.

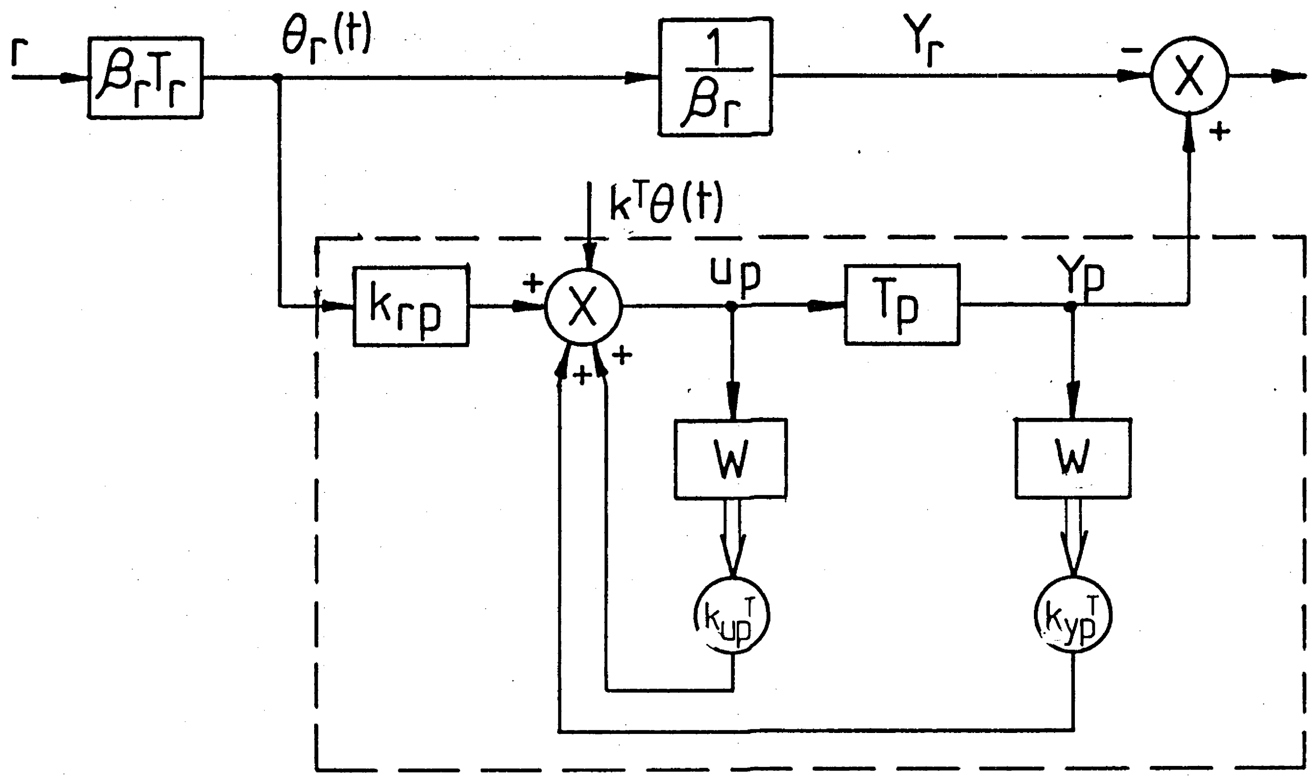


Figure 3.5 An equivalent representation of Morse's algorithm.

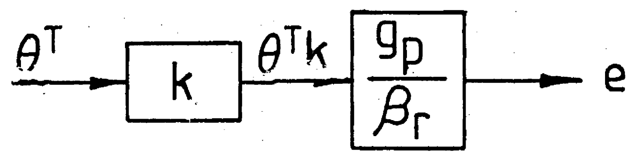


Figure 3.6 An error model for Morse's algorithm.

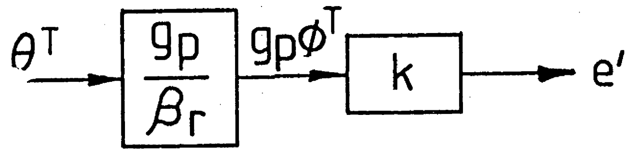


Figure 3.7 Error model not requiring strictly positive real condition for convergence.

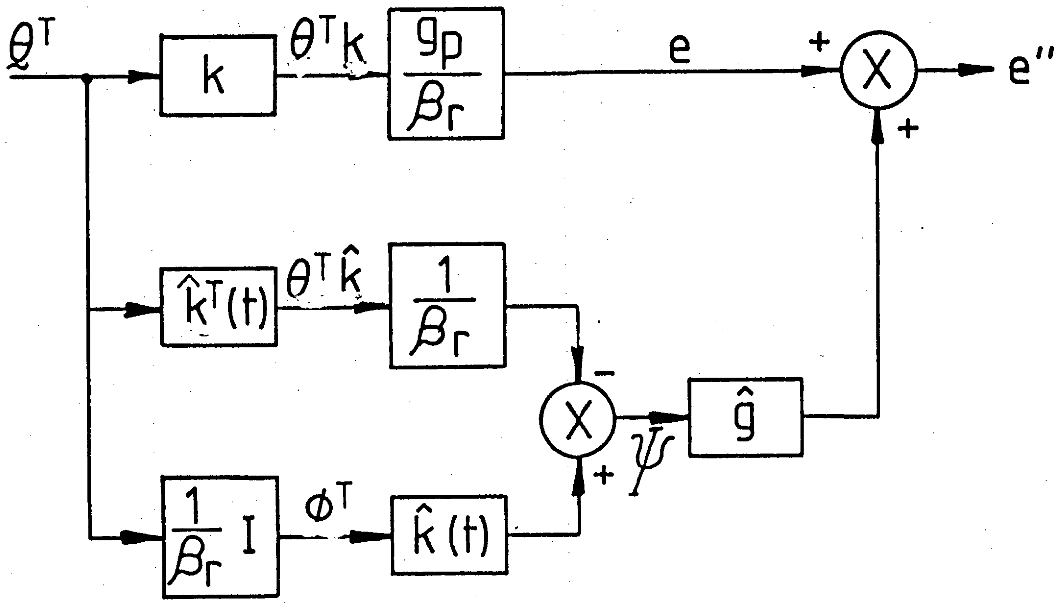


Figure 3.8 Augmented error model.

References for Chapter 3

1. B.D.O. Anderson,"Adaptive Systems, lack of persistence of excitation and bursting phenomena". Submitted to Automatica

2. B.D.O. Anderson and R.M. Johnstone, "When will adaptive systems really adapt? The robustness issue." Proc. 2nd Conference on Control Engg.IE Australia. 1982, pp 59-66.

3. W. Hahn, Stability of motion, Berlin Springer Verlag, 1967.

4. A.P. Morgan and K.S. Narendra, "On the uniform asymptotic stability of certain linear nonautonomous differential equqtions", SIAM J Contr. vol 15, Jan 1977, pp 5-24.

5. A.P. Morgan and K.S. Narendra, "On the stability of non-autonomous differential equations $\dot{x} = (A+B(t))x$ with skew-symmetric matrix B(t)",SIAM J Cont and Opt, vol 15, Jan 1977, pp 163-176.

6. G. Kreisselmeier, "Adaptive observers with arbitrary exponential rates of convergence", IEEE Trans Auto Cont, AC-21, Jan 1977, pp2-8.

7. G. Kreisselmeier, "Algebraic separation in realizing a linear state feedback control law by means of adaptive observers". IEEE Trans Auto Cont, AC-25, April 1980, pp 238- 243.

8. M.M. Sondhi and D. Mitra,"New results on the performance of a well known class of adaptive filters", Proc. IEEE, vol 64, Nov 76, pp 1583- 1597.

9. B.D.O. Anderson, "An approach to multivariable system identification", Automatica, vol 13, pp 401-408, 1977.

10. B.D.O. Anderson, "Exponential stability of linear equations arising in adaptive identification", IEEE Trans Auto Cont, AC-22, Feb 1977, pp 83-88.

11. J.S.C. Yuan and W.M. Wonham,"Probing signals for model reference identification",IEEE Trans Auto Cont, AC-22, Aug 1977, pp 530-538.

12. B.D.O. Anderson and C.R. Johnson,Jr.,"Exponential convergence of adaptive identification and control algorithms", Automatica, vol 18, Jan 1982, pp 1-13.

13. R.M. Johnstone and B.D.O. Anderson,"Exponential convergence of recursive least squares with exponential forgetting factor adaptive control",Systems and Control Letters, vol 2, August 1982, pp 69-76.

14. S.Boyd and S. Sastry,"On parameter convergence in adaptive control", Systems and Control Letters, pp311-319 Dec 1983.

15. S. Boyd and S. Sastry,"Necessary and sufficient conditions for parameter convergence in adaptive control",Tech.Report University of California, Berkeley, UCB/ERL M84/25, 1984.

16. A.S. Morse,"Global stability of parameter adaptive control systems", IEEE Trans Auto Contr, AC-25, June 1980, pp 433-439.

17. B.D.O. Anderson and R.M. Johnstone,"Adaptive systems and time varying plants",Int J. Cont, 1983, vol 37, No. 2, 367-377.

18. K.S. Narendra, Y.H. Lin and L.S. Valvani, "Stable adaptive controller design, part II: Proof of stability", IEEE Trans Auto Contr, AC-25, June 1980, pp 440-448.

19. S. Dasgupta, B.D.O. Anderson and A.C. Tsoi,"Input conditions for continuous time adaptive system problems", Proc. 22nd CDC, San Antonio, Texas, Dec 1983. Submitted to IEEE Trans Aut Contr.

20. B.D.O. Anderson, S. Dasgupta and A.C. Tsoi,"On the convergence of a model reference adaptive control algorithm with unknown high frequency gain",Submitted for publication.

21. R.R. Bitmead, Convergence properties of discrete time stochastic adaptive estimation algorithms, Ph.D thesis University of Newcastle, 1979.

22. D.S. Mitrinovic, Analytic Inequalities, Springer Verlag, 1970.

23. G. Kreisselmeier, "An approach to stable indirect adaptive control", submitted for publication.

24. H. Elliot, R. Cristi and M. Das,"Global stability of a direct adaptive pole placement algorithm", Technical Report, University of Massachusetts, Amherst, #UMASS-ECE-NO81-1,November 1982.

25. K.S. Narendra,I.H. Khalifa and A.M. Annaswamy,"Error models for hybrid adaptive control" Tech. Report No. 8207 Center for Systems Science, Yale University, 1982.

26. N.N. Krasovskii, Stability of motion, Stanford University press,1963.

27. G. Kreisselmeier, Personal communication.

§4. EQUATION ERROR IDENTIFICATION

This chapter considers a class of equation error identification algorithms for systems whose unknown parameters have direct physical relevance. In the last two decades a large number of algorithms for identifying linear, time-invariant systems of known finite order have appeared in the literature [1-7]. The standard approach in these and many others algorithms is to presuppose a complete lack of knowledge about the unknown system (aside from the degree and relative degree) and to ignore all additional information available to the modeller. The algorithms thus estimate the numerator and denominator coefficients of the transfer function, having first assumed all of them to be unknown. In practice, however, a great deal of partial knowledge is often available, which if exploited should give rise to parametrisations involving fewer unknowns and better identification schemes.

Formulated here are identification algorithms which exploit a form of partial knowledge commonly encountered in practice. Of particular interest are situations where uncertainities in a system are restricted to specific parameter values (usually with physical significance), eg. the value of a moment of inertia, a frictional coefficient or a capacitor, with the remainder of the system known a priori. As argued in chapter 2, for such systems, the tranfer function can be viewed as a ratio of two polynomials with the polynomial coefficients multilinear in the unknown parameters.

In section 4.1 are presented three identification algorithms which exploit the multilinearity mentioned above. Two of these involve two step procedures. The first step obtains an unconstrained estimate of the parameter vector by ignoring the inherent nonlinearity while the second step uses this latter knowledge to constrain the parameters on to the desired mainfold. The third algorithm essentially combines these two steps into one by using penalty functions. Two of these algorithms require that the bounds on the magnitude of the parameters be known. Given that these parameters are directly related to physical element values, such a requirement is easy to satisfy. Section 4.1.3 gives an interpretation of two of the algorithms. Section 4.2 presents persistence of excitation conditions on the system input which guarantee the uniform asymptotic stability (u.a.s.) of these algorithms. The analysis used draws heavily upon the tools developed in chapter 3. Section 4.3 considers in turn the convergence properties of the three algorithms while section 4.4 discusses certain ideas related to identifiability and implementability of one of the algorithms. Section 4.5 presents simulation results. Most of the work contained in this chapter appears in [8] and [9].

4.1. PARAMETER ADJUSTMENT LAWS

4.1.1 The System and Some Notations

In this chapter, we consider the problem of identifying a stable, single input, single output, time invariant linear system with proper transfer function

$$W(s) = \frac{P(s,k)}{Q(s,k)} \tag{4.1}$$

Here, k is an N-vector of the unknown parameters $k_1,\ldots,k_N$ which are to be identified and $P(s,k)$ and $Q(s,k)$ are polynomials in s. Specifically, we study the case where P and Q are multilinear in $k_1,\ldots,k_N$. Thus, with the set S defined as

$$S \triangleq \{1,\ldots,N\} \tag{4.2}$$

then

$$P(s,k) \triangleq p_0(s) + \sum_{r\subset S} [\prod_{i\varepsilon r} k_i]\, p_r(s) \tag{4.3}$$

$$Q(s,k) \triangleq q_0(s) + \sum_{r\subset S} [\prod_{i\varepsilon r} k_i]\, q_r(s) \tag{4.4}$$

For each subset r of S, $p_r(.)$ and $q_r(.)$ (as well as $p_0(.)$ and $q_0(.)$) are known polynomials. In this section, adaptive algorithms for identifying k from the measurements of the input signal u and the output signal y and knowledge of the coefficient polynomials $p_r(.)$ and $q_r(.)$ is presented. Each algorithm produces at time t an estimate $\hat{k}(t) \triangleq [\hat{k}_1(t), \ldots ,k_N(t)]^T$ of the unknown vector k. Define the parameter error $x(t)$ by

$$x(t) \triangleq \hat{k}(t) - k \tag{4.5}$$

and let K and $\hat{K}(t)$ be vectors in $\mathbb{R}^{(2^N-1)}$ with entries $\prod_{i \varepsilon r} k_i$ and $\prod_{i \varepsilon r} \hat{k}_i(t)$ respectively for each $r \subset S$. The ordering of entries must be consistent and the singelton subsets $\{1\}, \{2\}, \ldots, \{N\}$ are placed first. We shall abuse notation by writing, for example, K_{123} for $K_{\{1,2,3\}}$, being the $r = \{1,2,3\}$ entry of K. Thus, for a three parameter system,

$$k = [k_1, k_2, k_3]^T$$

$$\hat{k}(t) = [\hat{k}_1(t), \hat{k}_2(t), \hat{k}_3(t)]^T$$

$$x(t) = [\hat{k}_1(t)-k_1, \hat{k}_2(t)-k_2, \hat{k}_3(t)-k_3]^T$$

$$K = [k_1, k_2, k_3, k_1k_2, k_2k_3, k_3k_1, k_1k_2k_3]^T$$

$$\hat{K}(t) = [\hat{k}_1, \hat{k}_2, \hat{k}_3, \hat{k}_1\hat{k}_2, \hat{k}_2\hat{k}_3, \hat{k}_3\hat{k}_1, \hat{k}_1\hat{k}_2\hat{k}_3]^T$$

Further, define $D(x,k) \quad \varepsilon \mathbb{R}^{(2^N-1)}$ by

$$D(x,k) \triangleq \hat{K}(t) - K \tag{4.6}$$

and let $d(x,k)$ be the r-entry of $D(x,k)$ so that

$$d_r(x,k) \triangleq \prod_{i \varepsilon r} (k_i + x_i) - \prod_{i \varepsilon r} k_i \tag{4.7}$$

Let the matrix $\Lambda \, \varepsilon \mathbb{R}^{(2^N-1)\times(2^N-1)}$ be defined by

$$\Lambda = \text{diag}\, \{\lambda_r | r \subset S\} \tag{4.8}$$

where $\lambda_1, \ldots, \lambda_N > 0$ and $\lambda_{12}, \ldots, \lambda_{12\ldots N} \geqslant 0$. The ordering of Λ must be consistent with that of $K, \hat{K}$ and D.

One approach to identifying the system described by equation 4.1 is to rewrite it as

$$W(s) = \frac{p_0(s) + \sum_{r \subset S} K_r p_0(s)}{q_0(s) + \sum_{r \subset S} K_r q_r(s)} \quad (4.9)$$

$$= \frac{\sum_{i=0}^{p} a_i s^i}{\sum_{i=0}^{n} b_i s^i} \quad (4.10)$$

where the a_i and b_i coefficients are affine functions of K_r, $r \subset S$. These coefficients and hence K_r can then be identified by standard techniques [eg.1-7]. However, such an approach ignores several forms of à priori knowledge. First, several of the a_i and b_i , or linear combinations there of, are known. Second the coefficients K_r obey nonlinear relationships typified by

$$K_r = \prod_{i \varepsilon r} k_i \quad (4.11)$$

Further, since the unknown parameters k_i, $i \varepsilon S$ often represent physical element values or coefficients in many cases there will be à priori bounds on their values or knowledge of their signs which can be usefully exploited in their identification. This à priori information is more difficult to use in the identification of the a_i and b_i coefficients in equation(4.10).

In this section, three identification schemes for the system described by equation (4.1) are presented. The first two, referred to as the least squares two step algorithm and the generalized two step algorithm involve a two step procedure outlined as follows.

Step 1 : Ignoring the à priori information expressed in the nonlinear relationships between the parameters (equation 4.11), an unconstrained estimate, referred to as K_u, of K is produced

Step 2 : A constrained estimate $\hat{K}$ which is in some sense close to K_u is produced so that it obeys the nonlinear relations of equation (4.11).

Note that these two steps are performed simultaneously.

For example in a two parameter system with transfer function

$$\frac{p_0(s)+k_1p_1(s)+k_2p_2(s)+k_1k_2p_{12}(s)}{q_0(s)+k_1q_1(s)+k_2q_2(s)+k_1k_2q_{12}(s)}$$

the first step generates $K_u = [K_{u1},K_{u2},K_{u12}]^T$ where K_{u12} need not equal $K_{u1}K_{u2}$. The second step generates a $[\hat{k}_1,\hat{k}_2]^T$ such that $[\hat{k}_1,\hat{k}_2,\hat{k}_1\hat{k}_2]^T$ is in some sense the closest to K_u . It is not true, however, that the first step does not utilize any <u>à priori</u> information whatsoever. The only information that it is ignoring is the nonlinear relationship between K_{u12} and K_{u1} and K_{u2} , for example. Thus, as compared to parameter update laws which estimate a_i's and b_i's in (4.10), having assumed them all to be unknown, the number of estimated parameters in the first step is still smaller.

The third algorithm, referred to as the single step algorithm, combines these two steps into one by using a penalty function.

The remainder of this section contains detailed descriptions of these algorithms and their implementation. The last subsection is an heuristic justification of the approach taken in the design of these algorithms. These algorithms are designed on the premise of ideal system behaviour; i.e. the system satisfies certain idealizing assumptions of no noise, no time variation and no modelling errors. In face of modest departures from ideality, we shall argue that these algorithms still behave in a more or less acceptable manner. What is more, simulation results in section 4.5 demonstrate, that in the two step algorithms outlined above the second step leads to substantial improvement over the first step, whenever the idealizing assumptions are violated.

4.1.2 Least Squares Two Step Algorithm

Define for each $r \subset S$,

$$v_r(s) = p_r(s)u(s) - q_r(s)y(s) \tag{4.12}$$

and

$$h_r(s) = \frac{v_r(s)}{(s+\gamma)^n} \tag{4.13}$$

where $\gamma > 0$ and n is the highest degree among the q_r (i.e. the order of the system in (4.1)). We have thus introduced state variable filters in the fashion of [1], to avoid explicit differentiation of the measurements. In order to simplify presentation, the notation will be abused by writing, for example, $h_i(t)$ to denote the inverse Laplace transform of $h_i(s)$. Let for $t > 0$, $V(t)$ and $H(t)$ be vectors with elements $v_r(t)$ and $h_r(t)$, respectively, for all $r \subset S$. (The ordering of entries runs in the same manner as the ordering of entries of K.) Then the input-output relation in (4.1) can be rewritten through (4.13) as

$$h_0(t) + K^T H(t) \equiv 0 \tag{4.14}$$

Let us now consider the vector $r_0(t)$ and the matrix $R(t)$ given by

$$r_0(t) = \int_0^t e^{-\alpha(t-\tau)} h_0(\tau) H(\tau) d\tau \tag{4.15}$$

and

$$R(t) = \int_0^t e^{-\alpha(t-\tau)} H(\tau) H^T(\tau) d\tau \tag{4.16}$$

where $\alpha > 0$. Under certain persistence of excitation conditions and certain restrictions on the polynomials

$p_r(s)$ and $q_r(s)$ it will be shown in the next section that for some $t \geq t_0$ $R(t)$ is nonsingular for all $t > t_0$. Hence, for the remainder of this section, the following interim assumption is made. It can be removed after section 4.2.

Interim Assumption I1

There exists $t_0 > 0$ such that $R(t)$ is nonsingular for all $t \geq t_0$. ∇ ∇ ∇

To state the least squares two step algorithm define the function $L : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}_+$ by

$$L(x,k) = \frac{1}{2} D(x,k)^T \Lambda D(x,k) \tag{4.17}$$

where D and Λ are defined in equations (4.6) and (4.8) respectively.

Then the proposed parameter update law is

$$\dot{\hat{K}}(t) = - \left[\frac{\partial L(x(t),k)}{\partial x}\right]^T \tag{4.18}$$

$$= - \frac{\partial \hat{K}^T(t)}{\partial \hat{k}(t)} \Lambda [\hat{K}(t) - K] \tag{4.19}$$

Equation (4.18) can be implemented using the following result.

Proposition 4.1

Consider the parameter update law in equation (4.18) and suppose that interim assumption I1 holds. Then (4.18) can be implemented by the following equation

$$\dot{\hat{k}}(t) = - \frac{\partial \hat{K}^T(t)}{\partial \hat{k}} \Lambda [R(t)^{-1} r_0(t) + \hat{K}(t)] \tag{4.20}$$

$$\dot{R}(t) = -\alpha R(t) + H(t) H(t)^T , R(0) = 0 \tag{4.21}$$

for $t \varepsilon [0,t_0]$

$$\frac{d}{dt}[R(t)^{-1}] = \alpha\, R(t)^{-1} - R(t)^{-1}H(t)H(t)^{T}R(t)^{-1} \tag{4.22}$$

for $t \geqslant t_0$

$$\dot{r}_0(t) = -\alpha\, r_0(t) + h_0(t)\, H(t) \tag{4.23}$$

where $H(t)$ is the vector of $h_r(t)$ defined in equation (4.13) and Λ is defined in equation (4.8) and $\hat{K}(t)$ is the vector with entries $\prod_{i \varepsilon r} \hat{k}_i(t)$.

∇ ∇ ∇

Remarks

(4.1) Equations (4.21), (4.22) and (4.23) are on-line implementations for the variables $R(t)$ $(R(t)^{-1})$ and $r_0(t)$ defined in equations (4.15) and (4.16).

(4.2) Starting with equation (4.14) and post-multiplying by $e^{-\alpha T}\, H(t)^{T}$ yields that

$$e^{-\alpha t}(h_0(t)\; H(t)^{T} + K^{T}\; H(t)\;\; H(t)^{T}) = 0, \forall t \varepsilon R_{+}$$

By integrating this equation, it follows that

$$r_0(t) + R(t)\; K = 0 \tag{4.24}$$

i.e.

$$K = -R(t)^{-1}\; r(t) \tag{4.25}$$

Thus equations (4.21), (4.22) and (4.23) form K. This is step 1 of the algorithm referred to above. Equation (4.20) can be viewed as a steepest descent minimization of L which is the second step of the algorithm.

(4.3) The use of quantities $R(t)$ and $r(t)$ in identification is not new. For example, in [6] one can find an identification algorithm using an error formed as $r(t)+R(t)\hat{K}(t)$ which can be verified to equal $R(t)\ [-K(t)+\hat{K}(t)]$.

(4.4) Precision demands that (4.14) be rewritten as

$$h_0(t) + K^T H(t) + \varepsilon_1(t) \equiv 0$$

where $\varepsilon_1(t)$ consists of linear combinations of exponentially decaying signals, arising from initial conditions in the system and the state variable filters. In conjunction with the boundedness of $H(t)$, this in turn requires (4.24) and (4.25) to be rewritten as

$$r(t) + R(t)K + \varepsilon_2(t) \equiv 0$$

and

$$-R^{-1}(t)r(t) = K + R^{-1}(t)\varepsilon_2(t)$$

respectively, where $\varepsilon_2(t)$ is exponentially decaying as well. Since $R^{-1}(t)\varepsilon_2(t)$ is exponentially decaying, we will henceforth ignore it in our convergence analysis.

(4.5) Even if we neglect the initial conditions, (4.24) and (4.25) hold only under ideal circumstances, i.e. when there are no noise and unmodelled modes. In face of departures from these assumptions $R^{-1}(t)r(t) \neq -K$. As we shall argue in the next subsection the second step is designed to improve upon the estimate generated through $-R^{-1}(t)r(t)$, a fact which is borne out by the simulation results of section 4.5.

4.1.3 An Interpretation of the Least Squares Two Step Algorithm

It could be argued that equation (4.20) in the least squares two step algorithm is unnecessary since the vector K can be formed by multiplying $-R^{-1}(t)$ with $r_0(t)$. The unknown parameters k are then the first N components of K. However, this ignores round-off errors, effects of initial conditions and spurious signals such as noise, all of which combine to make $-R(t)^{-1}\ r_0(t)$ (as calculated) different from K. In fact, $-R(t)^{-1}\ r_0(t)$ (as calculated) will probably not obey the nonlinear relations implied by equation (4.11) i.e. $-R^{-1}(t)\ r_0(t)$ may not belong to the correct model set. Equation (4.20) can be viewed as an attempt to use these nonlinear relations to filter out the imperfections and to constrain the eventual estimate to lie in the correct model set.

To examine this problem in more detail by way of an anology, consider the problem of estimating a parameter vector Φ given a sequence of scalar measurements y_i and vectors x_i and model

$$y_i = x_i^T \Phi + n_i \tag{4.26}$$

where n_i is a noise sequence. A common estimate for Φ is the least squares estimate $\overline{\Phi}_u$ which minimizes $\sum_i (y_i - x_i^T \Phi)^2$ i.e.

$$\overline{\Phi}_u = R^{-1} a \tag{4.27}$$

where $R \triangleq \sum_i x_i x_i^T$ and $a \triangleq \sum_i y_i x_i$. But now suppose that Φ is constrained so that for some $\theta \varepsilon \mathbb{R}^2$

$$\Phi = [\theta_1, \theta_2, \theta_1\theta_2]^T \tag{4.28}$$

We now seek an estimate $\hat{\Phi} = [\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_1\hat{\theta}_2]^T$ minimizing $\sum(y_i - x_i^T\Phi)^2$. If the measurements y_i were noise free, one could use the same $\hat{\Phi}$ as before. But in general this will not be possible, and the minimizing θ will satisfy

$$\begin{bmatrix} 1 & 0 & \hat{\theta}_2 \\ 0 & 1 & \hat{\theta}_1 \end{bmatrix} \left(a - R \begin{bmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \\ \hat{\theta}_1\hat{\theta}_2 \end{bmatrix}\right) = 0 \tag{4.29}$$

which is obtained by differentiating $\sum_i(y_i - x_i^T\Phi)^2$ with respect to θ . This equation could be very awkward to solve, and alternative approaches to parameter estimation with obvious intuitive appeal might prove more practical. One such possible approach involves the following two step procedure.

(i) Find the value of Φ , call it $\bar{\Phi}_u$, which minimizes $\sum_i(y_i - x_i^T\Phi)^2$.

(ii) Find the value of Φ , call it $\hat{\Phi}$, which minimizes $||\Phi - \Phi_u||_2$ and such that $\exists\theta$ such that $\Phi = (\theta_1, \theta_2, \theta_1\theta_2)^T$.

The solution to step (i) is $\bar{\Phi}_u = R^{-1}a$ and the gradient solution to step (ii) is

$$\begin{bmatrix} 1 & 0 & \hat{\theta}_2 \\ 0 & 1 & \hat{\theta}_1 \end{bmatrix} \left(R^{-1}a - \begin{bmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \\ \hat{\theta}_1\hat{\theta}_2 \end{bmatrix}\right) = 0 \tag{4.30}$$

Now let us relate these ideas to our identification algorithm. If there is some noise or other modelling error present, the quantity $-R^{-1}r_0$ has to be interpreted as a least squares estimate, based on measurements up to time t , and with exponential discounting of old data.

As such one cannot expect that $-R^{-1}r_0$ will have the correct constrained structure which K has. Accordingly, our identification algorithm attempts to implement the second step of the two-step procedure described above - it introduces a gradient descent algorithm to locate the value of K (satisfying the constraints) which is a best least squares match of $-R^{-1}r_0$ i.e. a gradient descent algorithm for finding the $\hat{k}$ which minimizes $L(x,k) = [\hat{K} + R^{-1}r_0]^T\Lambda[\hat{K} + R^{-1}r_0]$.

Note that the gradient descent algorithm provides a practical procedure for solving the equivalent equation to (4.30). It may provide some opportunity to track time-varying parameters and offer the possibility of averaging out the fluctuations in $-R^{-1}(t)r_0(t)$, which stem from noise. In fact simulation results to be presented in section 4.5 illustrate that this is indeed the case and that the second step leads to a significant improvement in the accuracy of the parameter estimates.

4.1.4 General Two Step Algorithm

The algorithm presented in this subsection is a generalization of the least squares two step algorithm. The term $-R(t)^{-1}r_0(t)$ is replaced by $K_u(t) \varepsilon \mathbb{R}^{(2^N-1)}$. It is assumed that $K_u(t)$ is generated by a differential equation and that, in the absence of noise or modelling errors,

$$\lim_{t\to\infty} K_u(t) = K \tag{4.31}$$

For example, $K_u(.)$ might be the solution of

$$\dot{K}_u(t) = -\beta H(t) [h_0(t) + H(t)^T K_u(t)] \tag{4.32}$$

for some $\beta>0$. The quantity within the brackets can be verified to be $[K_u(t)-K]^T H(t)$ and as such is obtainable in an ideal (eg noiseless) case from measurements, even when K is unknown. The implementation of this law is then precisely equation(4.20)with $R(t)^{-1}r_0(t)$ replaced by $K_u(t)$ i.e.

$$\dot{\hat{k}}(t) = -\frac{\partial \hat{K}^T(t)}{\partial \hat{k}} \Lambda [K(t) - K_u(t)] \tag{4.33}$$

4.1.5 Single Step Algorithm

The third algorithm studied in this chapter does not produce an estimate $\hat{k}(t)$ of k. Instead, it produces an estimate $K_s(t)$ of K which tends in the limit to obey the nonlinear relations of equation (4.11). To achieve this, define the function $L_s : \mathbb{R}^{(2^N-1)} \times \mathbb{R}^N \to \mathbb{R}_+$ by

$$L_s(K_s,k) = \frac{1}{2}\,[\lambda(K_s-K)^T(K_s-K) + \sum_{r\subset S}\lambda_r(K_{sr}-\prod_{i\varepsilon r} K_{si})^2] \tag{4.34}$$

where K_{sr} is the r-th entry of K_s, K_{si} is the {i}-th entry and λ and λ_r are positive constants. Recall that the first N entries of K are k and that the rth entry is $K_r = \prod_{i\varepsilon r} k_i$. For example, if $N=2$,

$$L_s(K_s,k) = \frac{1}{2}\,[\lambda(K_{s1}-k_1)^2 + \lambda(K_{s2}-k_2)^2 + \lambda(K_{s12}-k_1k_2)^2 + \lambda_{12}(K_{s12}-K_{s1}K_{s2})^2] \tag{4.35}$$

Thus the first term in equation (4.34) is the square of the Euclidean distance between K_s and K and the second term is the measure of the amount by which K_s fails to obey the nonlinear relations implied by equation (4.11).

The single step algorithm is obtained by a steepest descent type minimization of L_s. That is, for some $\beta>0$,

$$\dot{K}_s(t) = -\beta\, R(t)\left[\frac{\partial L_s(K_s(t),k)}{\partial K_s}\right]^T \tag{4.36}$$

where $R(t)$ is defined in equation (4.16). Note that equation (4.36) is a differential equation with state space $\mathbb{R}^{(2^N-1)}$.

Define, for $K_s \varepsilon \mathbb{R}^{(2^N-1)}$ and $r\subset S$, the vector $X_r(K_s)$ in $\mathbb{R}^{(2^N-1)}$ to be

$$X_r(K_s) \triangleq [\frac{\partial}{\partial K_s} (K_{sr} - \prod_{i \in r} K_{si})]^T \tag{4.37}$$

For example, for N=2 and r = {1,2}

$$X_r(K_s) = [-K_{s2},-K_{s1},1]^T$$

Then the single step algorithm of equation (4.36) can be implemented as

$$\dot{K}_s(t) = -\beta \; [\lambda R(t) K_s(t) + \lambda r_0(t) + R(t) \sum_{r \subset S} \lambda_r$$

$$(K_{sr} - \prod_{i \in r} K_{si}) \; X_r(K_s(t))] \tag{4.38}$$

This follows by differentiation of equation (4.34) and observing that

$$r_0(t) = -R(t) \; K \tag{4.39}$$

4.2 PERSISTENCE OF EXCITATION

Section 4.1 showed that the implementation of the least squares two step parameter update law is conditional on the non-singularity of the matrix $R(t)$ defined as

$$R(t) = \int_0^t \bar{e}^{\alpha[t-\tau]} H(\tau)H^T(\tau)d\tau$$

With this in mind we now state p.e. conditions - first on $H(t)$ and subsequently on the input $u(t)$ - which guarantee the nonsingularity of $R(t)$. Moreover, it will be shown in subsequent sections that the conditions on $u(t)$ which guarantee the nonsingularity of $R(t)$ are also conditions which are necessary for the uniform asymptotic convergence of the other two algorithms.

4.2.1 A PE Condition on H(t)

Theorem 4.1:

Consider a vector $H(t)\varepsilon\Omega\ [0,\infty)$ and a positive scalar α. Define the function $g : \mathbb{R}^1 \rightarrow R^1$ by

$$g(\tau) = \begin{cases} \tau & \tau > 0 \\ 0 & \tau \leqslant 0 \end{cases}$$

Suppose there exist some σ_0, $\alpha_1 > 0$ and some t_1 such that $\forall\sigma\geqslant\sigma_0$ and $T<t_1$

$$\int_{g(\sigma-T)}^{\sigma} H(\tau)H^T(\tau)d\tau > \alpha_1 I \qquad (4.40)$$

Then there also exists an $\alpha_2 > 0$, such that $\forall\sigma\geqslant\sigma_0$

$$R(\sigma) = \int_0^{\sigma} \bar{e}^{\alpha(\sigma-\tau)} H(\tau)H^T(\tau)d\tau > \alpha_2 I \qquad (4.41)$$

Conversely, if (4.41) holds for all $\sigma\geqslant\sigma_0$, then there exist α_3 and t_2 such that $\exists\ \sigma\geqslant\sigma_0$ and $\bar{T}\geqslant t_2$

$$\int_{g(\sigma-\bar{T})}^{\sigma} H(\tau)H^T(\tau)d\tau > \alpha_3 I \qquad (4.42)$$

Remark

(4.3) Equations (4.40) and (4.42) are usually involved in a definition of persistence of excitation. The point of the theorem is that the finite interval in the p.e. definition is equivalent to an arbitrarily long interval with exponential forgetting applied to measurements in that interval. From another point of view, looking at measurements over a finite interval is equivalent to looking at measurements over an arbitrarily long interval with an infinite discounting factor on all but a finite subinterval, where there is no discounting. So the theorem is concerned with a form of equivalence of different discounting rates in defining p.e.

Proof

The violation of (4.41) implies that for arbitrary $\varepsilon>0$ there exist a $\sigma > \sigma_0$ and a unit vector ξ such that

$$\left|\int_0^t e^{-\alpha(\sigma-\tau)} \{\xi^T H(\tau)\}^2 d\tau\right| \leq \varepsilon \qquad \forall t\varepsilon[0,\sigma]$$

Thus as $H\varepsilon\Omega_\Delta[0,\infty)$, by Lemma 3.A.1 and arguments similar to those in the proof of theorem 3.2 we have that

$$\left|e^{-\alpha(\sigma-\tau)}\{\xi^T H(t)\}^2\right| \leq 0(\varepsilon^{\frac{1}{2}}) \qquad \forall t\varepsilon[0,\sigma]$$

Whence

$$\left|\xi^T H(t)\right|^2 \leq e^{\alpha(\sigma-t)}\, 0(\varepsilon^{\frac{1}{2}}) \qquad \forall t\varepsilon[0,\sigma]$$

Thus for any finite $T<t_1$

$$\left|\xi^T H(t)\right|^2 \leq 0(\varepsilon^{\frac{1}{2}}) \qquad \forall t\varepsilon[g(\sigma-T),\sigma]$$

Thus the violation of (4.41) implies the same for (4.40).

Hence (4.40) implies (4.41).

(ii) Note first of all that $H \epsilon \Omega_{\Delta}[0,\infty)$ implies that there exists a K such that

$$||H(t)|| < K \qquad \text{on} \quad [0,\infty)$$

and that finite one sided limits for $H(t)$ and $\dot{H}(t)$ exist at all points in $[0,\infty)$. As before the violation of (4.42) implies that for arbitrary $\varepsilon > 0$, there exist a $\sigma > \sigma_0, \overline{T} > t_2$ and a unit vector ξ such that

$$|\xi^T H| < \varepsilon \qquad \text{on} \quad [g(\sigma-\overline{T}),\sigma]$$

Thus $$|\xi^T R(\sigma)\xi| = \int_0^{g(\sigma-\overline{T})} e^{-\alpha(\sigma-g(\sigma-\overline{T}))} e^{-\alpha(g(\sigma-\overline{T})-\tau)} \{\xi^T H(\tau)\}^2 d\tau$$

$$+ \int_{g(\sigma-\overline{T})}^{\sigma} e^{-\alpha(\sigma-\tau)} \{\xi^T H(\tau)\}^2 d\tau$$

$$\leqslant \begin{cases} 0\ (\varepsilon^2) & \sigma - \overline{T} \leqslant 0 \\ e^{-(\sigma-g(\sigma-\overline{T}))} \dfrac{K^2}{\alpha} + 0(\varepsilon^2) & \sigma - \overline{T} > 0 \end{cases}$$

By definition, $g(\sigma-\overline{T}) = \sigma - \overline{T}$ when $\sigma - \overline{T} > 0$ whence $\sigma - g(\sigma-\overline{T}) = \overline{T} \geqslant t_2$ when $\sigma - \overline{T} > 0$

Thus $|\xi^T R(\sigma)\xi| \leqslant e^{-\alpha t_2} \dfrac{K^2}{\alpha} + 0(\varepsilon^2)$ when $\sigma - \overline{T} > 0$

So selecting t_2 to force $e^{-\alpha t_2} \leqslant \varepsilon^2$ ensures

$$|\xi^T R(\sigma)\xi| \leqslant 0(\varepsilon^2)$$

Thus (4.41) is violated as well. In other words (4.41) implies (4.42).

Remark

(4.4) Conditions (4.40) and (4.42) in a sense indicate that the linear independence of the $h_r(t)$ and hence of the $v_r(t)$, $r \neq 0$ is necessary for the algorithm in (4.20 - 4.23) to be implementable. In some cases the linear dependence of the $v_r(t)$ may lead to a lack of identifiability of the k_i. For example if there exist nonzero constants c_1, c_2, c_3 such that

$$c_1 v_1(t) + c_2 v_2(t) + c_3 v_3(t) \equiv 0$$

one can readily verify by substituting into

$$v_0(t) + k_1 v_1(t) + k_2 v_2(t) + k_1 k_2 v_3(t)$$

that k_1 and k_2 cannot be distinguished from $k_1 + c_1 d$ and $k_2 + c_2 d$ respectively, where

$$d = \frac{c_3 - k_1 c_2 - k_2 c_1}{c_1 c_2}$$

There are instances, however, when this may not be the case. For example if any v_r is identically zero, then the v_r are linearly dependent, though the system may well be identifiable. Thus if

$$v_0(t) + k_1 v_1(t) + k_2 v_2(t) \equiv 0$$

then k_1, k_2 remain identifiable as long as v_1 and v_2 are linearly independent. Thus we have a situation where the unknown parameters are uniquely identifiable although the algorithm cannot be implemented. The question of possible remedies to this is discussed later in Section 4.4.

4.2.2 A PE Condition on u(t)

The theorem below relates the p.e. condition on $H(t)$ to one on $u(t)$, subject to the following assumption on

$p_1, \ldots, p_{12\ldots N}$ and $q_1, \ldots, q_{12\ldots N}$

<u>Assumption 4.1</u>

If $\{\theta_r\}$ are a collection of scalars such that

$$\sum_{r \subset S} \theta_r p_r(s) \equiv \sum_{r \subset S} \theta_r q_r(s) \equiv 0 \qquad (4.43)$$

Then $\theta_r = 0 \ \forall r \subset S$.

<u>Remarks on Assumption 4.1</u>.

(4.5) Should (4.43) hold for some nonzero θ_r, then the v_r would be linearly dependent <u>irrespective of the input signal</u>, and so $R(t)$ would be singular for all input signals and times t. Again as argued in remark (4.4) this does not always mean a lack of identifiability in the system, e.g. for the example cited there, when $q_{12}(s) \equiv p_{12}(s) \equiv 0$.

(4.6) Let n denote the maximum degree of any of the polynomials $p_r(s)$, $q_r(s)$. Rewrite (4.43) as

$$[\theta_1 \;\; \theta_2 \;\; \theta_{12\ldots N}] \begin{bmatrix} p_1^0 & p_1^1 & \cdots & p_1^n & q_1^0 & q_1^1 & \cdots & q_1^n \\ p_2^0 & p_2^1 & \cdots & p_2^n & q_2^0 & q_2^1 & \cdots & q_2^n \\ \vdots & & & & \vdots & \vdots & & \vdots \\ p_{12\ldots N}^0 & p_{12\ldots N}^1 & & p_{12\ldots N}^n & q_{12\ldots N}^0 & q_{12\ldots N}^1 & & q_{12\ldots N}^n \end{bmatrix} = 0$$

where p_j^i denotes the coefficient of s^i in $p_j(s)$. Then if $2^N > 2(n+1)$ or $2^{N-1} > n+1$, it is guaranteed that there exist nontrivial θ_i. So the number of parameters which can be identified by the algorithm presented to this point can never exceed a bound imposed by the complexity of the system.

Theorem 4.2:

Consider an asymptotically stable nth order system having a proper transfer function of the form

$$\frac{y(s)}{u(s)} = \frac{P(s,\ k)}{Q(s,\ k)}$$

with $k \varepsilon R^N$ $P(s,k) = p_0(s) + \sum_{r \subset S} (\prod_{i \varepsilon r} k_i) p_r(s)$ and $Q(s,k) = q_0(s) + \sum_{r \subset S} (\prod_{i \varepsilon r} k_i) q_r(s)$. Assume that $P(s,k)$ and $Q(s,k)$ are such that they are coprime, that the transfer function $P(s,k)/Q(s,k)$ is never identically independent of any of the k_i, in the sense that for no i do we have $\frac{\partial}{\partial k_i} [\frac{P}{Q}] \equiv 0$ for all s and all $k_1, k_2, \ldots, k_N$, that

$$\delta[q] = n > \delta[q_r] \quad \forall r \subset S$$

$u(t) \varepsilon \Omega_\Delta[0,\infty)$ and that Assumption 4.1 holds. Define m as the maximum degree among the polynomials

$$p_r \bar{q}_r - q_r p_{\bar{r}} \quad , \quad \forall r, \bar{r} \subset S, r \neq \bar{r}$$

and

$$p_r q_0 - q_r p_0 \quad , \quad \forall r \subset S$$

Define also

$$H^T(t) = [\frac{v_1(t)}{(s+\gamma)^n}, \frac{v_2(t)}{(s+\gamma)^n}, \ldots, \frac{v_N(t)}{(s+\gamma)^n}, \frac{v_{12}(t)}{(s+\gamma)^n}, \ldots, \frac{v_{12\ldots N}(t)}{(s+\gamma)^n}]$$

and

$$U^T(t) = [u(t), \frac{u(t)}{(s+\bar{\gamma})}, \ldots, \frac{u(t)}{(s+\bar{\gamma})^m}]$$

Suppose $\exists \alpha_4, \delta' > 0$ such that $\forall \sigma \varepsilon R_+$

$$\alpha_4 I \leqslant \int_\sigma^{\sigma+\delta'} U(t) U^T(t) dt \qquad (4.44)$$

Then $\exists \alpha_5 > 0$ and $\delta > \delta'$ such that $\forall \sigma \varepsilon R_+$

$$\alpha_5 I \leqslant \int_{\sigma}^{\sigma+\delta} H(t) H^T(t) dt \tag{4.45}$$

Proof

In the sequel, notation will be abused in the following way:

$$c(t) = \frac{a(s)}{b(s)} d(t)$$

will refer to the solution of the differential equation

$$b(p)c(t) = a(p)d(t)$$

$p \triangleq d/dt$, with arbitrary finite initial conditions. In the definitions of $H(t)$ and $U(t)$, as noted earlier all initial conditions are zero.

Suppose (4.45) is violated.

By Lemma 3.A.3 we have that $u(t) \varepsilon \Omega_\Delta[0,\infty)$ implies that both $H(t)$ and $U(t) \varepsilon \Omega_\Delta[0,\infty)$. Thus using lemma 3.A.1 and arguing as in the proof of theorem 3.2, we have that for an arbitrary $\varepsilon > 0$, there exists a unit vector $\theta^T = [\theta_1, \theta_2, \ldots \theta_{12\ldots N}]$ such that

$$|\theta^T H(t)| \leqslant \varepsilon \qquad \text{on } [\sigma, \sigma+\delta]$$

Thus $\exists$ a $\delta_1' < \delta$ such that

$$\left| \sum_{r \subset S} \theta_r \frac{\left(q_r \frac{P(s,k)}{Q(s,k)} - p_r\right) u}{(s+\gamma)^n} \right| \leqslant 0(\varepsilon) \quad \text{on } [\sigma+\delta_1, \sigma+\delta]$$

Thus by arguing as in theorem 3.2 $\exists$ a δ_2 such that

$$\left| \frac{(s+\gamma)^n Q(s,k)}{(s+\bar{\gamma})^{2n}} \sum_{r \subset S} \theta_r \frac{(q_r P(s,k) - p_r Q(s,k))}{(s+\gamma)^n Q(s,k)} u \right| \leqslant 0(\varepsilon)$$
$$\text{on } [\sigma+\delta_2, \sigma+\delta]$$

whence

$$\left| \sum_{r \subset S} \theta_r \frac{(p_r Q(s,k) - P(s,k) q_r)}{(s+\bar{\gamma})^{2n}} u \right| \leqslant 0(\varepsilon) \qquad (4.46)$$

Now $\sum_{r \subset S} \theta_r (p_r Q - q_r P) \not\equiv 0$

as otherwise either

$$\frac{y}{u} = \frac{P}{Q} \equiv \{ \sum_{r \subset S} \theta_r p_r(s) \} / \{ \sum_{r \subset S} \theta_r q_r(s) \}$$

or $\sum_{r \subset S} \theta_r p_r(s) \equiv \sum_{r \subset S} \theta_r q_r(s) \equiv 0$

neither of which can hold by hypotheses (the first falls down by the coprimeness and $\delta[q_0] > \delta[q_i]$ assumptions and the second by Assumption 4.1.) Now, (4.46) can be rewritten as

$$\left| \sum_{r \subset S} \theta_r \{ \frac{p_r (q_0 + \sum_{r \in S} (\prod_{i \in r} k_i) q_r) - q_r (p_0 + \sum_{r \in S} (\prod_{i \in r} k_i) p_r)}{(s+\bar{\gamma})^{2n}} \} u \right| \leqslant 0(\varepsilon) \qquad (4.47)$$

Thus there exists a nonzero vector $\bar{\theta}^T = [\theta_0, \ldots, \bar{\theta}_m]$ such that (4.47) is equivalent to

$$\left| \sum_{i=0}^{m} \frac{\theta_i (s+\bar{\gamma})^i}{(s+\bar{\gamma})^{2n}} u(t) \right| \leqslant 0(\varepsilon) \quad \text{on } [\sigma+\delta_2, \sigma+\delta]$$

where m has been defined in the hypothesis. Thus by lemma 3.A.1 and arguments similar to those in theorem 3.2

$$\left| (s+\bar{\gamma})^{2n-m} \sum_{i=0}^{m} \frac{\theta_i (s+\bar{\gamma})^i}{(s+\bar{\gamma})^{2n}} u \right| \leqslant 0(\varepsilon^{1/(2n-m+1)})$$

$$\text{on } [\sigma+\delta_2, \sigma+\delta]$$

whence

$$\left| (s+\bar{\gamma})^{2n-m} \sum_{i=0}^{m} \frac{\theta_i (s+\bar{\gamma})^i}{(s+\bar{\gamma})^{2n}} \; u \right| \leqslant O(\varepsilon^{1/(2n-m+1)})$$

on $[\sigma+\delta_2, \sigma+\delta]$

$$\int_{\sigma+\delta_2}^{\sigma+\delta} U(t)U^T(t)dt > \alpha_4 I$$

is violated. Thus (4.44) implies (4.45)

Remarks

(4.7) A single sinusoid carries with it two pieces of information - namely, its magnitude and phase - and a linear combination of $\frac{N}{2}$ sinusoids (1/2 a sinusoid is tantamount to a dc signal) should in general identify an N-parameter system. But for a system of the form

$$v_0(t) + K^T V(t) \equiv 0 \tag{4.48}$$

the application of an input $\sum_{i=1}^{N/2} u_i \sin\omega_i t$ would result in the generation of N equations of the form

$$a_{0i} + \sum_{r \subset S} a_{ri} \prod_{j \varepsilon r} k_j = 0 \tag{4.49}$$

and

$$b_{0i} + \sum_{r \subset S} b_{ri} \prod_{j \varepsilon r} k_j = 0 \tag{4.50}$$

where $v_r = \sum_{i=1}^{N/2} (a_{ri} \sin\omega_i t + b_{ri} \cos\omega_i t)$. In general, the nonlinearity in (4.49) and (4.50) means that they may not have a unique real point of intersection, and thus a linear combination of $N/2$ frequencies may not suffice to identify the parameter vector K. On the other hand a linear combination of $(2^N-1)/2$ sinusoids will generate altogether (2^N-1) equations in the k_i, from which the elimination of the (2^N-1-N) nonlinear terms is possible. This will leave N linearly independent linear equations. Thus an input with $(2^N-1)/2$ different frequency components should suffice to uniquely identify (4.48). However, in actuality one may require more than this number as the transfer function may become independent of some of the unknowns at all of these frequencies. For example, in a 3-parameter system all the ω_i may be the zeroes of the polynomials $p_1, p_{12}, p_{123}, q_1, q_{12}$ and q_{123} simultaneously, making the transfer function independent of k_1 and the latter simply unidentifiable for this particular choice of restriction of input frequencies. Theorem 4.2, quantifies, we believe, the above argument, whenever, the linear dependence of the v_r coincides with the loss of identifiability of (4.48). As far as linear independence is concerned, if all the ω_i are simultaneously the zeroes of both p_r and q_r for any $r \subset S$, then the linear dependence of the $\{v_r\}$ and hence the lack of implementability of the least squares two step algorithm is immediate. Thus $u(t)$ must be sufficiently rich in frequencies to preclude such a possibility.

(4.8) It is worth noting that $m \leq 2n - \ell$, where ℓ is the relative order of the system. Thus the required number of frequencies in the input signal will in general be smaller than that required for the algorithms in [1-7] (see Corollary 3.1).

(4.9) A reversal of the arguments in theorem 4.2 shows that if there exist α_6 and δ'' independent of σ such that

$$\alpha_6 I \leq \int_{\sigma}^{\sigma+\delta''} H(t)H^T(t)dt \qquad (4.51)$$

for all $\sigma \varepsilon R_+$, then there also exist α_7 and a suitably large $\delta > \delta''$, independent of σ such that

$$\alpha_7 I < \int_{\sigma}^{\sigma+\delta} U(t) U^T(t) dt \qquad (4.52)$$

Furthermore, when Q is not Hurwitz analysis on lines similar to that presented in Theorem 3.7 reveals that while (4.44) still implies (4.45), in going from (4.51) to (4.52) one must replace U by $\bar{U}$, where

$$\bar{U}^T = [u, \frac{u}{s+\bar{\gamma}}, \ldots, \frac{u}{(s+\bar{\gamma})^{m-\nu}}]$$

where ν is the number of imaginary axis zeros of Q . Thus when the unknown system has $j\omega$-axis poles, the precise locations of which may not be known, some relaxation on the complexity of u may be permitted.

4.3 Convergence Analysis

In this section the three parameter adjustment algorithms presented in section 4.1 are analysed for convergence.

4.3.1 Convergence of Least Squares Two Step Algorithm

The error model of the least squares two step algorithm is given by

$$\dot{x}(t) = \dot{\hat{k}}(t) = -\left[\frac{\partial \hat{K}(t)}{\partial \hat{k}(t)}\right]^T (\hat{K}(t)-K) \quad (4.53)$$

$$= -\frac{\partial L(x(t),k)}{\partial \hat{k}(t)} \quad (4.54)$$

with

$$L(x(t),k) \triangleq \frac{1}{2}(\hat{K}(t)-K)^T \Lambda (\hat{K}(t)-K) = \frac{1}{2}D^T(x,k)\Lambda D(x,k) \quad (4.55)$$

Recall that k is the N-dimensional parameter vector, $\hat{k}(t)$ its estimate, K and $\hat{K}$ are the 2^N-1 dimensional vectors containing the multilinear combinations of the elements of k and $\hat{k}$ respectively, $x(t) \triangleq \hat{k}(t)-k$, $D \triangleq \hat{K}(t)-K$ and Λ is a $(2^{N-1}\times 2^{N-1})$ dimensional diagonal matrix with diagonal elements λ_r. The first N diagonal elements of Λ are positive while the rest are non-negative.

In this subsection we are interested in the behaviour of (4.53). Several results are derived.

First, it is shown to be locally exponentially stable about the true parameter values. With a restriction on the magnitude of the true parameter values, global uniform asymptotic stability is also established.

Next, it is shown that regardless of the true parameter values the algorithm is uniformly convergent whenever the initial guesses for the entries of $\hat{k}$ are either zero or have the correct signs. Though these stability results on (4.53) are apparently independent of any p.e. conditions, without p.e. (4.53) cannot be implemented so the p.e. requirement is after all present.

Finally, a modification of (4.53) involving time varying Λ is proposed and shown to be globally uniformly asymptotically convergent.

As the understanding of the structure of (4.53) is crucial to the understanding of the proofs we note that (4.53) can be written more explicitly as

$$\dot{x}_i(t) = - \sum_{\substack{r \subset S \\ i \varepsilon r}} \lambda_r \left(\prod_{\substack{j \varepsilon r \\ j \neq i}} \hat{k}_j(t) \right) \left(\prod_{j \varepsilon r} \hat{k}_j - \prod_{j \varepsilon r} k_j \right) \qquad \forall i \varepsilon S \tag{4.56}$$

For example in the three parameter case this becomes

$$\dot{\hat{k}} = - \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} x - \lambda_{12} d_{12} \begin{bmatrix} \hat{k}_2 \\ \hat{k}_1 \\ 0 \end{bmatrix} - \lambda_{23} d_{23} \begin{bmatrix} 0 \\ \hat{k}_3 \\ \hat{k}_2 \end{bmatrix} - \lambda_{13} d_{13} \begin{bmatrix} \hat{k}_3 \\ 0 \\ \hat{k}_1 \end{bmatrix}$$

where $d_r \triangleq \left(\prod_{j \varepsilon r} \hat{k}_j(t) \right) - \left(\prod_{j \varepsilon r} k_j(t) \right)$.

The arguments in this section all refer to the behaviour of the algorithm in the idealized case of no noise, time-invariance of parameters, no nonlinearity, etc. Later, we shall comment on what can happen given departures from the ideal, using our conclusions about behaviour in the ideal case.

4.3.1.1 Local and Semiglobal Stability

The theorem stated below establishes the local and, what we call, "semiglobal" uniform asymptotic convergence of the algorithm in (4.53). The seimglobal result states that global uniform asymptotic convergence can be claimed if the unknown parameters or some of the elements of the gain matrix Λ fall in a certain region of space (of nontrivial size). In the sequel we shall assume that

$$\Lambda = \begin{bmatrix} \Lambda_1 & | & 0 \\ \hline 0 & | & \Lambda_2 \end{bmatrix} \qquad (4.57)$$

where

$$\Lambda_1 = \text{diag}\ \{\lambda_1, \ldots \lambda_N\}$$

and

$$\Lambda_2 = \text{diag}\ \{\lambda_{12}, \ldots \lambda_{123\ldots N}\}$$

where the elements of Λ_1 are positive and those of Λ_2 non-negative.

Theorem 4.3:

The parameter adjustment law

$$\dot{x} = - \frac{\partial L(x,k)}{\partial x}$$

is u.a.s. if any of the following hold

(i) The initial $x(0)$ lies in a ball around the origin of arbitrarily large radius R, and k lies in a ball around the origin of radius $r(R,\Lambda)$.

(ii) The initial $x(0)$ lies in a ball of arbitrarily large radius R and the elements of Λ_2, see (4.57), lie in a ball around the origin of radius

$r'(R,k)$.

(iii) The initial $x(0)$ lies in a ball around the origin of radius $r''(k,\Lambda)$.

Proof

See Appendix 4.A

∇ ∇ ∇

Remarks

(4.10) Stated simply the result claims the following: The algorithm will converge uniformly asymptotically to zero, regardless of how large $x(0)$ is, provided that the actual parameter magnitude is "small enough" or if the elements of Λ_2 are "small enough". It is also locally u.a.s. (condition (iii)) irrespective of k and Λ, though the extent of convergence will depend on k and Λ.

However, the theorem fails to specify the exact extent of $\| k \|$ and $\| \Lambda_2 \|$ for which, global (with respect to x) uniform asymptotic convergence can be claimed, i.e. the formula for the radii r and r' in (i) and (ii). By contrast, for the two parameter case, this region is exactly known. In fact theorem 4.4 given below shows that this region is defined by

$$k^T \Lambda_1 k < \frac{8\lambda_1\lambda_2}{\lambda_{12}}$$

and r and r' are seen to be independent of R. Whether this dependence extends to $N > 2$ parameters is unknown.

(4.11) Of the above (ii) is particularly interesting as it shows that the algorithm is exponentially stable if

$\Lambda_2 = 0$. This is because $\Lambda_2 = 0$ implies

$$\dot{x}_i = -\lambda_i x_i$$

The question may well arise as to why Λ_2 need be made non-zero at all if global stability can be claimed even otherwise. By making Λ_2 zero the multilinearities are no longer explicitly accounted for in our algorithm. (This notion is also consistent with our discussion of the rationale of the algorithm is section 4.1.3) Thus, intuitively, useful information is being discarded, with, one would conjecture, an accompanying degradation of performance.

Theorem 4.4:

Consider (4.53) when N , the number of unknown parameters, is two. Then (4.53) is u.a.s if

$$k^T \Lambda_1 k \leqslant \frac{8\ \lambda_1 \lambda_2}{\lambda_{12}}$$

with Λ_1 defined in (4.57).

Proof

For N=2 , (4.53) becomes

$$\dot{\hat{k}} = -\ [I\ \ E(x,k)]\ \Lambda \begin{bmatrix} I \\ (\tfrac{1}{2}x+k)^T E \end{bmatrix} x \qquad (4.58)$$

where

$$E = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Taking L(x,k) (see (4.55) as a Lyapunov function from (4.53) one can see that (4.53) is u.a.s. if the right side of (4.58) is zero only when $x=0$. A sufficient condition for this is

$$\det \; [I \quad E(x,k)] \; \Lambda \begin{bmatrix} I \\ (\tfrac{1}{2}x+k)^T E \end{bmatrix} > 0$$

$$\Leftarrow \det \Lambda_1 (1 + \lambda_{12} (\tfrac{1}{2}x+k)^T E \Lambda_1^{-1} E(x+k)) > 0$$

$$\Leftarrow 1 + \lambda_{12} (\tfrac{1}{2}x+k)^T \frac{\Lambda_1}{\lambda_1 \lambda_2} (x+k) > 0 \qquad (4.58a.)$$

By minimizing the left hand side of (4.58a.) with respect to x one can see that it is greater than

$$1 - \frac{\lambda_{12}}{8\lambda_1\lambda_2} k^T \Lambda_1 k.$$

Thus (4.53) is u.a.s. if

$$1 - \frac{\lambda_{12}}{8\lambda_1\lambda_2} k^T \Lambda_1 k > 0$$

whence the result follows. ∇ ∇ ∇

4.3.1.2. Assured uniform asymptotic convergence

In this subsection we demonstrate that initializing the parameter estimates with correct signs guarantees uniform asymptotic convergence. Indeed, in many real situations such sign knowledge will in fact be available as unknown parameter may be a moment of inertia, a frictional coefficient and so on. If, on the other hand, such knowledge is not available, we show that $\hat{k}(0) = 0$ will suffice. Our result is proved in two steps. We show first of all that all trajectories starting from the closed orthant of the $\hat{k}$ space, which contains $\hat{k} = k$, remain in that orthant. We then show that this orthant is free from false equilibria, i.e. $\dot{x}$ in (4.53) is zero iff $x=0$ $(\hat{k} = k)$.

Lemma 4.1

In $\hat{k}$ space, let O denote the closed orthant

$$O = \{(\hat{k}_1, \ldots, \hat{k}_N) \,\big|\, k_i \hat{k}_i > 0, \forall i\} \tag{4.59}$$

Then all trajectories starting from any point in O (including the origin) remain in O for all time.

Lemma 4.2

None of the equilibria of the parameter adjustment law (4.53), except $\hat{k} = k$, lie in the orthant O.

Both these lemmata are proven in appendix 4.B.

Theorem 4.5:

Suppose the initial estimate of the parameter vector $\hat{k}$ be such that

$$k_i \hat{k}_i(0) > 0 \quad \forall i \varepsilon S \tag{4.60}$$

Then (4.53) is u.a.s.

Proof

If the initial estimates satisfy (4.60) we have, by lemma 4.1 that the parameter estimates remain in O. Taking L to be the Lyapunov function we have from (4.54) that

$$\dot{L} = -\left[\frac{\partial L(x,k)}{\partial x}\right]^T \left[\frac{\partial L(x,k)}{\partial x}\right]$$

which is negative everywhere except at points at which

$$\frac{\partial L(x,k)}{\partial x} = 0$$

i.e. points of equilibrium of (4.53). Thus $\dot{L}$ is negative definite in a region free from false equilibria.

Lemma 4.2 thus ensures that $\dot{L}$ is negative definite in O. Thus whenever (4.60) satisfied the trajectories remain in a region where $\dot{L}$ is negative definite. The result is then immediate. ∇ ∇ ∇

Remarks

(4.12) In proving theorem 4.3 it can be shown that the local rates of convergence of (4.53) are exponential, in that if $x(t_0)$ lies in a ball of radius $r''(k,\Lambda)$ around $x=0$, then $x(t)$ decays exponentially fast for all $t > t_0$. Furthermore, starting from $\hat{k}(0) = 0$ (4.53) is u.a.s. and $\hat{k}(t)$ so generated is thus uniformly bounded so that one can always construct an exponentially decaying function which overbounds $x(t)$. This shows that (4.53) is exponentially stable whenever $\hat{k}(0)$ is in the orthant O.

(4.13) It is a well known fact that adaptive algorithms with exponential rates of convergence are substantially immune to noise and a variety of modelling deficiencies. Thus our algorithm should, as long as the trajectories remain in Orthant O, be robust. If on the other hand, the presence of noise, of reasonable magnitudes, forces the trajectories to leave O, one may well ask if convergence can still be expected. Observe that if $\hat{k}_j = 0$

$$\dot{x}_j = -\Big[\sum_{\substack{r \subset S \\ j \varepsilon r}} \lambda_r \Big(\prod_{\substack{\alpha \varepsilon r \\ \alpha \neq j}} \hat{k}_\alpha k_\alpha\Big)\Big] x_j$$

and if $\hat{k} = 0$

$$\dot{x} = -\Lambda_1 x$$

It is easy to see from these that in the vicinity of $\hat{k} = 0$

and/or the hyperplanes bounding O , the trajectories still point towards the interior of O . Thus for sufficiently small excursions from O , they can be expected to reenter O and eventually to converge to $x=0$.

(4.14) Of interest also is to see what happens if once $\hat{K}$ has converged to the correct value K , the latter undergoes a step change to $\bar{K}$. Thus

$$R(t)K = -r_0(t) \qquad \forall t \leq t_0 \tag{4.61}$$

and

$$\hat{K}(t_0) = K \tag{4.62}$$

Now, let K become $\bar{K}$ $\forall t>t_0$. Thus the new system is given by

$$h_0(t) + \bar{K}^T H(t) \equiv 0 \tag{4.63}$$

By definition

$$\begin{aligned} R(t) &= \int_0^t e^{-\alpha(t-\tau)} H(\tau)H^T(\tau)d\tau \\ &= \int_0^{t_0} e^{-\alpha(t-\tau)} H(\tau)H^T(\tau)d\tau + \int_{t_0}^t e^{-(t-\tau)} H(\tau)H^T(\tau)d\tau \\ &= e^{-\alpha(t-t_0)} R(t_0) + \int_{t_0}^t e^{-\alpha(t-\tau)} H(\tau)H^T(\tau)d\tau \end{aligned} \tag{4.64}$$

Similarly,

$$r_0(t) = e^{-\alpha(t-t_0)} r_0(t_0) + \int_{t_0}^t e^{-\alpha(t-\tau)} h_0(\tau)H(\tau)d\tau \tag{4.64a.}$$

Equations (4.61 - 4.64a.) imply that $\forall t>t_0$

$$\begin{aligned} R(t)\bar{K} = {} & e^{-\alpha(t-t_0)} R(t_0)(\bar{K}-K) + e^{-\alpha(t-t_0)} R(t_0)K \\ & + \int_{t_0}^t e^{-\alpha(t-\tau)} H(\tau)H^T(\tau)Kd\tau \end{aligned}$$

$$= e^{-\alpha(t-t_0)} R(t_0)(\bar{K}-K) - e^{-\alpha(t-t_0)} r_0(t_0) - \int_{t_0}^{t} e^{-\alpha(t-\tau)} h_0(\tau) H(\tau) d\tau$$

$$= e^{-\alpha(t-t_0)} R(t_0)(\bar{K}-K) - r_0(t)$$

Thus

$$-R^{-1}(t) r_0(t) = \bar{K} - e^{-\alpha(t-t_0)} R^{-1}(t) R(t_0)(\bar{K}-K)$$

Thus $-R^{-1}(t)r_0(t)$ approaches $\bar{K}$ exponentially. In light of theorem 4.7 in section 4.3.2, this implies that $\hat{k}$ converges uniformly asymptotically to $\bar{k}$, $\bar{k}$ obviously defined, as long as $\bar{k}$ has the same sign as k . Since the parameters have physical significance signs of the elements of k are unlikely to change.

4.3.1.3 A modified least squares two step algorithm

In this subsection we propose a modification of the algorithm in (4.53) by requiring Λ to be time variant. Recall from the proof of theorem 4.5 that with L (see 4.55) as a Lyapunov function $\dot{L}$ is negative semidefinite equalling zero at the stationary points of (4.53). Unfortunately (4.53) in general has stationary points apart from $x=0$ and thus global u.a.s. of (4.53) cannot be claimed. In remark (4.16) we show, however, that all stationary points apart from $x=0$ are unstable with respect to changes in Λ_2 and thus in principle one can find perturbations in Λ_2 which will induce the parameter estimates to drift away from these "false" points of equilibrium. One way of achieving this is to continuously alter the elements of Λ_2 in a manner indicated by Theorem 4.6.

Theorem 4.6:

With L and Λ_2 defined as in (4.55) and (4.57) respectively, the parameter update law

$$\dot{x} = -\frac{\partial L[x(t),\Lambda_2(t),k]}{\partial x} \tag{4.65}$$

and

$$\dot{\Lambda}_2(t) = \text{diag}\ \{f_r(\Lambda_2(t),t)\}_{\substack{r \subset S \\ |r|>1}} \tag{4.66}$$

is u.a.s. if f_r are continuous, $\lambda_r(0)>0, \forall\lambda_r(0)$ a diagonal element of $\Lambda_2(0)$ and $-f_r(\Lambda_2,t) > \phi_r(\Lambda_2)$ with ϕ_r obeying

$$\begin{aligned} &\text{(i)} \quad \phi_r = 0 \quad \text{iff}\lambda_r = 0 \\ &\text{(ii)} \quad \phi_r > 0 \quad \forall\lambda_r > 0 \end{aligned} \tag{4.67}$$

Note $|r|$ = number of elements of r

Proof

See Appendix 4.C.

Remarks

(4.15) A possible $f_r(\Lambda_2,t)$ is $-a_r\lambda_r^{\ell_r}$, $a_r > 0$. More generally, the rth element of f must be negative for all positive λ_r. Thus in the limit λ_r can be expected to converge to zero though this convergence can be made arbitrarily slow

(4.16) A heuristic interpretation of the above theorem is that by changing Λ_2 we are continuously driving the false equilibria away from the trajectory. To see why this is so observe that for a fixed Λ_2

$$\dot{x}(x_e,\Lambda_2) = -\Lambda_1 x_e - \sum_{\substack{r\subset S \\ |r|>1}} \left[\lambda_r d_r \frac{\partial d_r}{\partial x}\right]\Bigg|_{x = x_e}$$

Suppose x_e is an equilibrium point and $x_e \neq 0$. Then as $\dot{x} = 0$, the second term on the right hand side of the above equation must be nonzero. Suppose we alter Λ_2 by $\Delta\Lambda_2$, such that λ_r is charged by $-\lambda_r a$, $a>0$. Then

$$\dot{x}(x_e,\Lambda_2 + \Delta\Lambda_2) = a \sum_{\substack{r\subset S \\ |r|>1}} \left[\lambda_r d_r \frac{\partial d_r}{\partial x}\right]\Bigg|_{x = x_e}$$

Thus by changing Λ_2 we cause the trajectory to move from the false stationary point.

(4.17) It is evident, nonetheless, that in the limit we could drive Λ_2 to zero. An attractive modification of the algorithm would then be the following:

(i) leave Λ_2 constant until $\dot{x}$ slows down;

(ii) then change Λ_2 according to $\Delta\Lambda_2 = -a\Lambda_2$, until the convergence rate picks up again.

Then if the equilibrium point being approached is $x = 0$ changing Λ_2 will not alter matters. False equilibrium points on the other hand will be driven away.

4.3.2 Convergence of the generalized two step algorithm

This subsection analyses the convergence properties of the generalized two step algorithm. Exponential convergence is shown to be conditional on a p.e. condition and the knowledge of the magnitude of the parameters. The latter information we reiterate is in practice available as magnitude bounds on physical parameters are usually known.

Examined here is the parameter update law

$$\dot{\hat{k}}(t) = - \frac{\partial \hat{K}^T(t)}{\partial \hat{k}(t)} \Lambda \, [\hat{K}(t) - K_u(t)] \qquad (4.68)$$

where $K_u(t)$ is generated from a differential equation eg (4.32) and has the property that

$$\lim_{t\to\infty} K_u(t) = K$$

In this subsection we put forward conditions on $K_u(t)$ which will force (4.68) to retain the convergence characteristics of (4.53). In particular we shall require $K_u(t)$ to approach K exponentially fast i.e.

$$\| K_u(t) - K\| < \nu_1 \| K_u(0) - K\| e^{-\nu_2 t} \qquad (4.69)$$

where ν_1 and $\nu_2 > 0$.

Theorem 4.7:

The parameter estimate update law (4.68) is u.a.s. if:

(i) the parameter estimate $\hat{k}(t)$ remains in a region where $\|\partial L/\partial x\|$, with L defined in (4.55), is positive definite i.e. it equals zero if $x = 0$, and

(ii) The adaptive law generating $K_u(t)$ is such that (4.69) is satisfied (in the absence of noise, structural modelling error, etc).

Proof

See Appendix 4.D. ∇ ∇ ∇

Remarks

(4.18) One way of satisfying condition (i) is to force $\hat{k}(t)$ to remain in $\mathcal{O}$. See also remark (4.19).

(4.19) It is evident from the foregoing that (4.68) will be globally u.a.s. whenever $\Lambda_2 = 0$ and $K_u(t)$

satisfies the condition of the above theorem. For nonzero Λ_2 , however, $K_u(t)$ needs to be so generated as to force each of its elements to either have the same sign as the corresponding element in K or else to equal zero. Then by arguing as in lemma 4.1 it can be shown that $\hat{k}(t)$ remains in O as long as $\hat{k}(0)$ is in O , whence subject to the satisfaction of (ii) in theorem 4.7, uniform asymptotic convergence of (4.68) is immediate. We describe below such a scheme which can be implemented whenever the signs of the k_i and the bounds on them are known a priori, viz if

$$m_i \leqslant k_i \leqslant M_i \quad \forall i \varepsilon S .$$

Consider the scheme

$$\dot{K}_u(t) = -\beta\ H(t)[K_u^T(t)H(t) + h_0(t)] \quad (4.70)$$

As the system is described by

$$K^T H(t) + h_0(t) \equiv 0$$

(4.70) is equivalent to

$$\dot{K}_u(t) = -\beta\ H(t)H^T(t)(K_u(t) - K(t)) \quad (4.71)$$

Thus by theorem 3.1 (4.69) is satisfied as long as $\exists$ constants $\alpha_1, \alpha_2, \delta > 0$ such that $\forall \sigma \varepsilon R_+$

$$\alpha_1 I \leqslant \int_{\sigma}^{\sigma+\delta} HH^T dt \leqslant \alpha_2 I \quad (4.72)$$

For $u(t) \varepsilon \Omega_\Delta[0,\infty)$, theorem 4.2 gives the conditions on $u(t)$ for which (4.72) holds.

Conditions (i) of theorem 4.7 is satisfied if $\hat{k}(t) \varepsilon O$ for all $t \varepsilon R_+$. This can be achieved by adding a penalty function in the fashion of [12] to (4.68) viz

$$\dot{\hat{k}}(t) = -\frac{\partial \hat{K}^T(t)}{\partial \hat{k}(t)} \Lambda [\hat{K}(t) - K_u(t)] - \Gamma_1\Psi_1(\hat{k}(t)) ,$$

$$\hat{k}(0)\varepsilon O$$

where Γ is a diagonal matrix and the ith element of Ψ is

$$\Psi_u(\hat{k}(t)) \triangleq \begin{cases} \hat{k}_i(t) - m_i & \text{when } \hat{k}_i(t)\leqslant m_i \\ 0 & \text{when } m_i\leqslant \hat{k}_i(t)\leqslant M_i \\ \hat{k}_i(t) - M_i & \text{when } \hat{k}_i(t)\geqslant M_i \end{cases}.$$

In the above m_i is assumed positive, a fact that can be achieved by translating the parameters k_i . The weighting matrix Γ_1 is chosen so that for every $i\varepsilon S$ and $t\varepsilon R_+$

$$0\leqslant\hat{k}_i(t) \leqslant \tilde{M}_i$$

when $\tilde{M}_i$ are arbitrarily chosen constants, greated than M_i. Analysis of equations of this form will be given in Chapter 5. Another possibility is to add a penalty function of the above form to (4.70) so that all the elements of $K_u(t)$ are always positive. This in turn ensures that $\hat{k}_i(t)\varepsilon O\forall t\varepsilon R_+$, a fact proved by a simple extension of the proof of lemma 4.1. The choice of the weighting matrix here, however, requires the knowledge of magnitude bounds on $H(t)$. If these are not available a normalization of the form in [12], given below may need to be introduced

$$\dot{K}_u(t) = -\beta \frac{H(t)[K_u^T(t)H(t)+h_0(t)]}{\bar{\beta} + H^T(t)H(t)} - \Gamma_2 \Psi_2(K_u(t))$$

where $\forall r\subset S$

$$\Psi_{2r}(K_u(t)) = \begin{cases} K_{ur} - \prod_{i\varepsilon r} m_i & \text{when } K_{ur} \leqslant \prod_{i\varepsilon r} m_i \\ 0 & \text{elsewhere} \end{cases}$$

The choice of Γ_2 and a comprehensive analysis can be found in Chapter 6.

4.3.3 Convergence of the single step law

Consider the differential equation (4.36). It can be rewritten as

$$\dot{K}_s = -\beta\, R(t)\, \left[\lambda(K_s(t)-K) + \sum_{r\subset S} \lambda_r (K_{sr}(t) - \prod_{i\varepsilon r} K_{si}(t)) \frac{\partial(K_{sr}(t) - \prod_{i\varepsilon r} K_{si}(t))}{\partial K_s(t)}\right] \tag{4.73}$$

with $\lambda_1\lambda_2 > 0$.

Then the following theorem describes conditions for its uniform asymptotic stability.

Theorem 4.8:

The adaptive law in (4.73) is u.a.s. in the large if

(i) $K_{si}(t)k_i \geqslant 0 \qquad \forall i\varepsilon S \qquad (4.74)$

and

(ii) an $\exists\alpha_1 > 0$ such that for some t_0 and all $t>t_0 \quad R(t) > \alpha_1 I \qquad (4.75)$

Proof

Consider the Lyapunov function

$$L_s = \frac{1}{2}\, [\lambda(K_s-K)^T(K_s-K) + \sum_{\substack{r\subset S\\ |r|>1}} \lambda_r (K_{sr} - \prod_{i\varepsilon r} K_{si})^2] \tag{4.76}$$

where $|r|$ denotes the number of elements in the set r . Clearly, L_s is positive definite with respect to K_s-K. Now (4.73) can be rewritten as

$$\dot{K}_s(t) = -\beta\, R(t) \left[\frac{\partial L_s(K_s(t),k)}{\partial K_s(t)}\right]$$

Then

$$-\dot{L}_s(t) = \beta \left[\frac{\partial L_s(K_s(t),k)}{\partial K_s(t)}\right]^T R(t) \left[\frac{\partial L_s(K_s(t),k)}{\partial K_s(t)}\right]$$

$$\geq \beta\alpha_1 \left[\frac{\partial L_s(K_s(t),k)}{\partial K_s(t)}\right]^T \left[\frac{\partial L_s(K_s(t),k)}{\partial K_s(t)}\right] \qquad (4.76a.)$$

Now $[\partial L_s(K_s(t),k)/\partial K_s(t)]$ has no explicit dependence on t. Thus (4.73) is u.a.s. if

$$\frac{\partial L_s(K_s(t),k)}{\partial K_s(t)} \equiv 0 \qquad (4.77)$$

$$\Longleftrightarrow \quad K_s(t) \equiv K \qquad (4.78)$$

Now, for all r not a sigleton the rth element of $\frac{\partial L_s(K_s(t),k)}{\partial K_s(t)}$ is given by

$$\lambda(K_{sr} - \prod_{i\varepsilon r} k_i) + \lambda_r(K_{sr} - \prod_{i\varepsilon r} K_{si})$$

Thus (4.77) $\Rightarrow$

$$K_{sr} = \frac{\lambda \prod_{i\varepsilon r} k_i + \lambda_r \prod_{i\varepsilon r} k_{si}}{\lambda+\lambda_r} \qquad \forall r\subset S, |r|>1$$

(4.79)

Now the ith element of $\partial L_s(K_s(t),k)/\partial K_s(t)$, $\forall i\varepsilon S$, is

$$\lambda(K_{si}-k_i) + \sum_{\substack{r\subset S\\ i\varepsilon r}} \lambda_r (\prod_{\substack{j\varepsilon r\\ j\neq i}} K_{si})(\prod_{j\varepsilon r} K_{si} - K_{sr})$$

Thus by (4.79), (4.77) $\Rightarrow$

$$\lambda(K_{si}-k_i) + \sum_{\substack{r\subset S\\ i\varepsilon r}} \lambda_r (\prod_{\substack{j\varepsilon r\\ j\neq i}} K_{si})(\prod_{j\varepsilon r} K_{si} - \frac{\lambda \prod_{i\varepsilon r} k_i + \lambda_r \prod_{i\varepsilon r} K_{si}}{\lambda+\lambda_r}) = 0 \quad \forall i\varepsilon S$$

$$\Rightarrow \lambda(K_{si}-k_i) + \sum_{\substack{r\subset S \\ i\varepsilon r}} \frac{\lambda r \lambda}{\lambda+\lambda_r} (\prod_{\substack{j\varepsilon r \\ j\neq i}} K_{si})(\prod_{j\varepsilon r} K_{si} - \prod_{j\varepsilon r} k_i) = 0 \qquad \forall i\varepsilon S \tag{4.80}$$

If we substitute $\hat{k}_i$ by K_{si} $\forall i\subset S$ in the proof of lemma 4.2, given in appendix 4.B, then we find that (4.80) holds iff

$$K_{si} = k_i \qquad \forall i\varepsilon S \tag{4.81}$$

as (4.74) holds and λ_1 and $\lambda_r>0$. Substituting (4.81) in (4.79) we find

$$K_{sr} = \prod_{i\varepsilon r} k_i \qquad \forall r\subset S$$

Thus (4.77) $\Leftrightarrow$ (4.78), whence the result follows.

∇ ∇ ∇

Remark

(4.20) Theorems 4.1 and 4.2 state conditions under which (4.75) is satisifed. To ensure that the first N elements of K_s always have the same signs as the corresponding elements of k we introduce a linear translation in the k_i's . Let us illustrate with a two parameter example.

Consider figure 4.0. Suppose the true parameters k_1 and k_2 are known to lie in the region

$$m_1 \leqslant k_1 \leqslant M_1$$

$$m_2 \leqslant k_2 \leqslant M_2$$

depicted by region A . Suppose $m_1 \leqslant m_2$ and the extents of the region in the X and Y directions are Δ_1 and Δ_2, respectively. Now suppose, the true parameters are both translated by a positive number Δ , so that their translated values $\tilde{k}_i$ lie in region B . Obviously the extents of region B in the x and y directions are unchanged. One obvious constraint on Δ is that $\bar{m}_1$ and $\bar{m}_2$ are both positive.

Suppose

$$L_s(t) = \lambda[(K_{s1}(t) - \tilde{k}_1)^2 + (K_{s2}(t) - \tilde{k}_2)^2 + (K_{s12}(t) - \tilde{k}_1\tilde{k}_2)^2$$

$$+ \lambda_{12}(K_{s12}(t) - K_{s1}(t)K_{s2}(t))^2$$

Now, (4.76a.) $\Rightarrow$

$$L_s(t) \leqslant L_s(0) \qquad \forall t \in R_+$$

Select $K_s(0)$ so that it belongs to region B and $K_{s12}(0) = K_{s1}(0)K_{s2}(0)$.

Then

$$L_s(0) \leqslant \lambda[\Delta_1^2 + \Delta_2^2 + (\bar{M}_1\bar{M}_2 - \bar{m}_1\bar{m}_2)^2]$$

$$= \lambda[\Delta_1^2 + \Delta_2^2 + (m_1\Delta_2 + m_2\Delta_1 + \Delta_1\Delta_2)^2]$$

$$= \lambda[\Delta_1^2 + \Delta_2^2 + [(\Delta_1 + \Delta_2)\bar{m}_1 + (m_2 - m_1)\Delta_1 + \Delta_1\Delta_2]^2] \tag{4.82}$$

Suppose $K_{s1}(t_1) = 0$ for some $t_1 > 0$. Then

$$L_s(t_1) \geqslant \lambda\tilde{k}_1^2 + \lambda(K_{s12} - \tilde{k}_1\tilde{k}_2)^2 + \lambda_{12}K_{s12}^2$$

$$\geqslant \lambda\tilde{k}_1^2 + \left[\frac{\lambda\lambda_{12}^2}{(\lambda+\lambda_{12})^2} + \frac{\lambda_{12}\lambda^2}{(\lambda+\lambda_{12})^2}\right]\tilde{k}_1\tilde{k}_2^2$$

$$= \lambda\tilde{k}_1^2 + \frac{\lambda\lambda_{12}}{\lambda+\lambda_{12}}\tilde{k}_1^2\tilde{k}_2^2 \tag{4.83}$$

Now, the highest power of Δ in equation (4.83) is four while that in (4.82) is two. Thus for suitable large Δ

$$L_s(t_1) > L_s(0)$$

Thus if we translate the parameters by a suitable extent then

$$K_{s1}(t) > 0 \qquad \forall t \varepsilon R_+$$

This procedure can be easily extended to $N>2$.

We remark that one could also conceivably use a penalty function term, of the form used in the gradient descent algorithm, to restrict the K_{si} to be positive. The stability analysis, however, becomes too awkward to handle.

4.4 SOME NOTIONS RELATED TO P.E. AND PARAMETER IDENTIFIABILITY

As noted earlier, the linear independence of the $v_r(t)$, $r \subset \{1, \ldots N\}$ is necessary for

$$R(t) = \int_0^t \bar{e}^{\alpha(t-\tau)} H(\tau) H^T(\tau) d\tau \tag{4.84}$$

to be nonsingular. If $R(t)$ is singular the least squares algorithm defined via (4.20 - 4.23) is not implementable. Recall from section 4.2 that the singularity of (4.84) is itself a consequence of the linear dependence of the $v_r(t)$ as

$$H(t) \triangleq \left[\frac{v_1(t)}{(s+\gamma)^n}, \frac{v_2(t)}{(s+\gamma)^n}, \quad --- , \frac{v_{12-N}(t)}{(s+\gamma)^n} \right]^T$$

However, the linear dependence of the $v_r(t)$'s need not always make the system

$$v_0(t) + K^T V(t) \equiv 0 \tag{4.85}$$

unidentifiable. In this section we consider modifications to the least squares two step algorithm to cope with situations where $R(t)$ is singular, yet the unknown parameters k_i are uniquely identifiable.

One form of lack of persistence of excitation arises when some of the $v_r(t) \equiv 0$. Below, we consider this situation in a greater depth.

Consider the error signal

$$e(t) = v_0(t) + \hat{K}^T(t) \; V(t) \tag{4.86}$$

where $\hat{K}(t)$ is a constrained estimate of K , defined in section 4.1. By subtracting (4.85) from (4.86) we have that

$$e(t) = D^T(t)\ V(t) \qquad (4.87)$$

where

$$D(t) \triangleq [d_1(t), \ldots, d_N(t), d_{12}(t), \ldots, d_{123\ldots N}(t)]^T$$

and

$$d_r(t) \triangleq \prod_{i \varepsilon r} \hat{k}_i(t) - \prod_{i \varepsilon r} k_i \qquad \forall r \subset S.$$

Now clearly if for any r , $v_r(t) \equiv 0$, then the corresponding d_r is not reflected in $e(t)$. If too many of these d_r are not reflected in $e(t)$ then error signals like $e(t)$ will not identify the parameters k_i uniquely. For example for a two parameter system

$$v_0(t) + k_1 v_1(t) + k_2 v_2(t) + k_1 k_2 v_{12}(t) \equiv 0 ,$$

suppose $v_2(t)$ and $v_{12}(t)$ are both identically zero. Then only d_1 is reflected in $e(t)$ and k_2 cannot be identified. The following definitions are meant to capture this situation.

Definition 4.1

For the system defined by (4.85), the error term $d_r(.)$, $r \subset S$, is observable if $v_r(.)$ is not identically zero with time.

Definition 4.2

The system in (4.85) is fundamentally identifiable under the following conditions: all the observable d_r equal zero if and only if $\hat{k} = k$.

Recall that the step II of the original least squares twp step algorithm

$$\dot{x}(t) = -\frac{\partial \hat{K}^T(t)}{\partial \hat{k}(t)} \Lambda D(t) \tag{4.88}$$

with

$$x(t) \triangleq [\hat{k}(t)-k]$$

Then an obvious modification

$$\dot{x} = -\frac{\partial \hat{K}^T(t)}{\partial \hat{k}(t)} \Lambda \bar{D} \tag{4.89}$$

where $\bar{D}$ is the same as D with the unobservable d_r's set equal to zero. The implementation of (4.89) is straight forward. With $H_{obs}(t)$, the vector of all $h_r(t)$ which are not identically zero, we define

$$R_{obs}(t) = \int_0^t \bar{e}^{\alpha(t-\tau)} H_{obs}(\tau) H^T_{obs}(\tau) d\tau$$

and

$$r_{obs}(t) = \int_0^t \bar{e}^{\alpha(t-\tau)} h_0(\tau) H_{obs}(\tau) d\tau \ .$$

Here $h_r(t) = \frac{1}{(s+\gamma)^n} v_r(t)$

Then $D(t)$ is obtained by $\hat{K}_{obs} + R^{-1}_{obs}(t) r_{obs}(t)$, where $\hat{K}_{obs}$ is obviously defined.

As far as the uniform asymptotic convergence of (4.88) is concerned an obvious pre-requisite is that the system be fundamentally identifiable as only then will $\bar{D}(t)$ have enough elements to ensure identifiability. But that by itself is not enough. One must also ensure that the closed orthant O of the $\hat{k}$ space, which contains the true parameter value, has no points other than $\hat{k}=k$ for which $\dot{x}=0$ in (4.89). Condition (ii) of theorem 4.9, below, guarantees this situation and is explained at a greater length in remark (4.21)

Theorem 4.4:

For the system defined in (4.89) let the following conditions hold:

(i) the gain element λ_r is positive whenever d_r is observable

(ii) for every $i \varepsilon S$ there exists at least one observable d_{r_i} such that $i \varepsilon r_i \subset S$ and for all other $j \varepsilon r_i$, d_j is observable; (note that r_i may equal i whenever d_i is observable).

Then for a fundamentally identifiable system (4.85), (4.89) is u.a.s. whenever

$$\hat{k}_i(0)k_i > 0 , \qquad \forall i \varepsilon S \tag{4.90}$$

Proof

See Appendix 4.E. ∇ ∇ ∇

Remarks

(4.21) Condition (ii) in the statement of theorem 4.9 implies that at least one among $d_1 \ldots d_N$ is observable. Should this be violated then $\hat{k} = 0$ will be a stationary point of (4.89) making $\hat{k}(0) = 0$ insufficient to guarantee convergence. For example when $N=3$, let $d_1, \ldots d_3$ be unobservable, but let d_{12}, d_{123}, d_{23} and d_{13} be observable. Then it can be checked that the system in question is fundamentally identifiable. However, the parameter law (4.89) becomes

$$\dot{\hat{k}}_1 = -\lambda_{12}\hat{k}_2 d_{12} - \lambda_{13}\hat{k}_3 d_{13} - \lambda_{123}\hat{k}_2\hat{k}_3 d_{123}$$

$$\dot{\hat{k}}_2 = -\lambda_{12}\hat{k}_1 d_{12} - \lambda_{23}\hat{k}_3 d_{23} - \lambda_{123}\hat{k}_1\hat{k}_3 d_{123}$$

$$\dot{\hat{k}}_3 = -\lambda_{13}\hat{k}_1 d_{13} - \lambda_{23}\hat{k}_2 d_{23} - \lambda_{123}\hat{k}_2\hat{k}_1 d_{123}$$

Clearly, when $\hat{k}_1 = \hat{k}_2 = \hat{k}_3 = 0$, $\dot{\hat{k}} = 0$. Thus the origin of the $\hat{k}$ space is a stationary point.

More generally, even if some of the d_i , $i\varepsilon S$, are observable, the nonsatisfaction of (ii) would cause some of the hyperplanes bounding O to be stationary with respect to (4.89). For example if in a 3-parameter system only d_3, d_{23} and d_{123} are observable, the trajectories will never leave the plane $\hat{k}_1 = 0$. Thus, unless $k_1 = 0$, $\hat{k}(0) = 0$ will not lead to convergence. On the other hand if (4.90) can be replaced by

$$\hat{k}_i(0)k_i > 0 \qquad \forall i\varepsilon S \tag{4.91}$$

which is possible if signs of the k_i are known <u>a priori</u> , then as long as for every $i\varepsilon S$ there exists at least one observable d_{r_i} with $i\varepsilon r_i \subset S$, (4.89) would be uniformly convergent. To understand this observe that if any $\hat{k}_i = 0$, the parameter estimate update law becomes

$$\dot{x}_i = - \sum_{\substack{r\subset S\\ i\varepsilon r}} \lambda_r \Big(\prod_{\substack{\alpha\varepsilon r\\ \alpha\neq i}} \hat{k}_\alpha k_\alpha \Big) x_i$$

where $\lambda_r = 0$ for all unobservable d_r ; whence it is easy to see that in the interior of the orthant O , all trajectories point away from $\hat{k} = 0$ and the boundaries $\hat{k}_i = 0$. Thus the satisfaction of (4.91) guarantees that all trajectories remain confined to a region of negative definite $\dot{L}$ as long as the system is fundamentally identifiable.

(4.22) Lack of p.e. could arise as a consequence of deficient inputs i.e. those not satisfying (4.44), or due to a fundamental property of the system, namely the violation of assumption 4.1. As far as the modification in (4.89) is concerned, it works for both cases as long as the nonzero $v_r(t)$ are linearly independent. When some of the nonzero $v_r(t)$ are dependent due to the violation of assumption 4.1, it is sometimes possible to find an affine transformation in k_i which makes all nonzero transformed $v_r(t)$ linearly independent, whence one can apply (4.89); e.g. for N=2, v_1 and v_{12} linearly dependent and

$$v_2 + \alpha v_{12}(t) \equiv 0 \qquad (4.92)$$

let $\hat{k}_1 = (k_1 - \alpha)a^{-1}$ for some nonzero a. Then observe that

$$v_0 + (a\hat{k}_1 + \alpha)v_1 + k_2 v_2 + (a\hat{k}_1 + \alpha)k_2 v_{12} \equiv 0$$

whence

$$(v_0 + \alpha v_1) + a\hat{k}_1 v_1 + a\hat{k}_1 k_2 v_{12} \equiv 0 \ .$$

Thus theorem 4.9 is applicable whenever $k_1 \neq \alpha$. If however $k_1 = \alpha$ and (4.92) holds, it is easily checked that the system is fundamentally unidentifiable, in the sense that it is impossible to identify k_2. If linear dependence such as (4.92) arises due to insufficiently frequency-rich inputs, then the required affine transformation cannot be determined a priori, as the value of α would in this instance depend on the unknown k_i. Moreover, in some situations a convenient affine transformation cannot be obtained e.g. when N=2, and v_1 and v_2 are linearly dependent. Whether an appropriate modification to (4.89) exists in these two cases, remains an open question.

4.5 SIMULATION RESULTS

In this section we present simulation results for the two step least squares adaptive identifier. The primary goal is to demonstrate that inclusion of the second step does indeed improve performance in the face of a whole range of deviations from ideality. The system considered has transfer function

$$T_p(s) = \frac{(s+2)(s+3)}{(s+1)(s+4)(s+5)} \tag{4.93}$$

parameterized as

$$T_p(s) = \frac{s^2+k_1 s+k_2}{(s^3+6s^2-s-10) + 2k_1 s^2-k_2 s^2+k_1 k_2 (s+1)} \tag{4.94}$$

with $k_1 = 5$ and $k_2 = 6$.

By applying theorem 4.2 it is clear that p.e. is guaranteed whenever the input has two distinct frequencies. This requirement is satisfied in all the simulations. All the λ_r's are unity, unless otherwise specified.

Throughout this section $-R^{-1}(t)\ r_0(t)$ will be denoted by $K_u(t)$ and the second step will be treated as

$$\dot{\hat{k}} = - \frac{\partial K^T(t)}{\partial k(t)} \Lambda \ [\hat{K}(t)-K_u(t)] \tag{4.95}$$

Ideally, of course $K_u(t) = K$.

4.5.1 Simusoidal drift in the parameters

In this subsection we introduce simusoidal drift in the parameters. Thus $k_1 = 5 + .5 \sin .01t$ and $k_2 = 6 + .5 \sin .01t$. The time t_0 in (4.22) is 5. Figs 4.1 and 4.2, show the tracking abilities of $\hat{k}_1$, K_{u1} and $\hat{K}_2$, K_{u2} respectively. The curve marked 1, in each plot represents k_i while the curve 2 represents $\hat{k}_i$ or

K_{ui}, as the case maybe. The first plot in each figure is an exploded view of k_i, $\hat{k}_i$ vst and the second is k_i, K_u vst. It is evident that the $\hat{k}_2$ tracks k_2 much better than K_{u2}, while the performance of $\hat{k}_1$ and K_{u1} are comparable. The third plot in each figure show $k_i, \hat{k}_i$ vst on the parameter axis scale 0-6. It is evident that tracking is almost perfect. Thus the second step has indeed improved matters.

4.5.2 Step change in the parameters

In this subsection we consider step changes in the parameters after the identifier has estimated the old parameter values. The values of k_1 and k_2 are

$$k_1 = \begin{cases} 5 & t<50 \\ 6 & t\geq 50 \end{cases}$$

$$k_2 = \begin{cases} 6 & t<50 \\ 7 & t\geq 50 \end{cases}$$

Figures (4.3) and (4.4) show the values of K_{u1} and $\hat{k}_1$ respectively; Fig (4.5) has both plots superimposed on one another. It is clear that convergence to the second value is smoother for $\hat{k}_1$ than it is for K_{u1}. This is because the second step also acts as a low pass filter. Similar results were obtained for $\hat{k}_2$ and K_{u2}. Thus here again the second step leads to improvement.

4.5.3 Identification when system has high frequency unmodelled modes:

Here we consider the case where the system has unmodelled modes at $-100 \pm 14i$. Thus it has a transfer function

$$T_{pu}(s) \quad \frac{10196(s+2)(s+3)}{(s+1)(s+4)(s+5)(s^2+200s+10196)}$$

although the identifier is designed for (4.93). The first and second plots on figure (4.6) are $\hat{k}_1$ vst and K_{u1} vst, respectively. Similarly, figure (4.7) gives $\hat{k}_2$ K_{u2} vst. The respective steady state values are

$$\hat{k}_1 = 4.956 \qquad K_{u1} = 4.5029$$

$$\hat{k}_2 = 5.2582 \qquad K_{u2} = 4.8594$$

Clearly, both the $\hat{k}_i$ track the values of 5 and 6 closer than do K_{u1} and K_{u2}. Also, the $\hat{k}_i$ are far less bumpy than their K_{ui} counterparts.

In the next two subsections we dispense with the first step, but run (4.33) after introducing deliberate errors in K_u.

4.5.4 Sinusoidal disturbance in the unconstrained estimate

The two plots in fig (4.8) give K_{u2} and $\hat{k}_2$ vst when the unconstrained estimate K_u has a sinusoidal disturbance. Thus $K_u(t)$ is

$$K_u^T(t) = [5 + .05 \sin .01t,\ 6 + .05 \sin .01t,\ 30 + 05 \sin .01t]$$

It is clear that the effect of the disturbances on $\hat{k}_2$ is much less marked than on K_{u2}. Similar results were obtained for $\hat{k}_1$ and K_{u1}.

4.5.5 Biased noise in the unconstrained parameter estimate

In this section we introduce noise in the unconstrained estimate K_u. Both the bias and the standard diviation will be varied as will be the parameter λ_{12}. The nominal parameter values will be 5 and 6.

Figure (4.9) gives $\hat{k}_1$ and K_{u1} vst when K_u is $[5,6,30]^T + \eta(t)$ where

$$E(\eta_i(t)) = .5 \qquad i=1,2,3$$

$$\sigma(\eta_i(t)) = 1 \qquad i=1,2,3$$

Also $\lambda_{12} = 1$.

Figure (4.10) gives the corresponding plots for $\hat{k}_2$ and K_{u2}. The following features are of interest.

(i) There is a significant smoothing effect due to the second step.

(ii) The approximate bias for $\hat{k}_1$ and $\hat{k}_2$ are -.0249 -.0047 respectively. Thus there is a significant reduction in bias as well.

Figure (4.11) gives $\hat{k}_1$ and K_{u1} vst when

$$E(\eta_i(t)) = .1 \qquad i=1,2,3$$

and $\sigma(\eta_i(t)) = .1 \qquad i=1,2,3$.

There is again a perceptible filtering effect. The resulting biases on $\hat{k}_1$ and $\hat{k}_2$ are -.0001 and .0083 respectively.

Figures 4.12 and 4.13 represent cases when

$E[K_u] = [4.5, 6.5, 29.5]^T$. But, in the former, noise of standard deviation .1 is present in each element of K_u while, the latter has no noise.

The following observations can be made

(i) The steady state values of $\hat{k}_1$ and $\hat{k}_2$ in fig 4.12 are 4.5222 and 6.5059, respectively. Though the bias in $\hat{k}_2$ is apparently bigger than .5, the distance $(\hat{k}_1-5)^2 + (\hat{k}_2-6)^2$ is smaller than $(4.5-4)^2 + (6.5-5)^2$.

(ii) The steady state values of $\hat{k}_1$ and $\hat{k}_2$ in fig (4.13) are 4.9812 and 5.9059.

(iii) Apparently there are more than one combination of $(\hat{k}_1, \hat{k}_2, \hat{k}_1\hat{k}_2)$ which are equally close to (4.5,6.5,29.5). Two such are given by $(\hat{k}_1, \hat{k}_2)$ = (4.5222, 6.5059) and (4.9812, 5.9059). It appears that in such situations external factors such as absence or presence of noise determine which parameter set is selected.

Thus the second step has two effects on noise in $-R^{-1}(t)\ r_0(t)$:

(i) it has a low pass filtering effect and

(ii) it results in significant reduction in bias.

The effect of λ_{12} is interesting. Figure (4.14) gives $\hat{k}_1$ vst when $E\{K_u\} = [4.4, 5.5, 29.9]^T$ and $\sigma(\eta_i(t)) = 1\ \forall\ i$, for the cases where $\lambda_{12} = 6$ and 0. The respective biases were (.0073, .0016) and (-.6023, -.4041). Thus though the low pass filtering effect is retained, with $\lambda_{12} = 0$, bias reduction is almost absent and convergence is substantially slower.

Finally, figs 4.15 and 4.16 deal with the case when the bias in K_u is [-.6, -.5, .1] and the σ_i are .1. The values of λ_{12} were 1 and 3 respectively. The resultant biases in $[\hat{k}_1, \hat{k}_2]$ were [.0003, .009] and [.0035, .0065] respectively. Bearing in mind that we have reached the limits of accuracy for the routine employed, it is nonetheless clear, that no significant improvement in bias is obtained beyond λ=1.

In conclusion one notes that the second step of the two step least squares algorithm improves upon the estimate $-R^{-1}(t)\ r_0(t)$ on several counts. It thus leads to a better response in tracking unmodelled modes, tracks step changes more smoothly, has a better ability to track sinusoidal drift in the unknown parameters, leads to the reduction of noise in terms of both bias and standard deviation and reduces sinusoidal disturbances in $-R^{-1}(t)\ r_0(t)$. However, although its response speed is high, it is nonetheless constrained to be slower than the estimate generated in the first step, i.e. $-R^{-1}(t)\ r_0(t)$, for obvious reasons.

APPENDIX 4.A : PROOF OF THEOREM 4.3.

Subject to the condition that all diagonal elements of Λ_1 are strictly positive and those of Λ_2 are nonnegative we see immediately that the Lyapunov function $L(x,k,\Lambda_2)$ defined in (4.55) is positive definite in x .

Also

$$\dot{L}(x,k,\Lambda_2) = -\left[\frac{\partial L(x,k,\Lambda_2)}{\partial x}\right]^T \left[\frac{\partial L(x,k,\Lambda_2)}{\partial x}\right] \quad (4.A.1)$$

is negative semidefinite, having zeros at the stationary points of (4.53) , and is continuous with respect to x, k and Λ_2 . To prove (i) and (ii) we need simply prove then that $\dot{L}(x,0,\Lambda_2)$ and $\dot{L}(x,k,0)$ are both negative definite in x . Then by continuity there exists a k-ball and a Λ_2-ball such that for any k and Λ_2 lying, respectively, in them, the algorithm is u.a.s. Now, for $k=0$ from (4.56) we have that

$$\dot{x} = - \sum_{\substack{r \subset S \\ i\varepsilon r}} \lambda_r [\prod_{\substack{\alpha\varepsilon r \\ \alpha\neq i}} x_\alpha][\prod_{\beta\varepsilon r} x_\beta]$$

$$= - [\sum_{\substack{r \subset S \\ i\varepsilon r}} \lambda_r (\prod_{\substack{\alpha\varepsilon r \\ \alpha\neq i}} x_\alpha)^2] x_i$$

which has no stationary point apart from $x=0$. This proves (i).

For $\Lambda_2 = 0$ (4.53) becomes

$$\dot{x} = -\Lambda_1 x$$

which too has only one stationary point $x = 0$. This proves (ii).

To prove (iii) observe that

$$\dot{x} = -\left[\frac{\partial \hat{K}^T(t)}{\partial \hat{k}(t)}\right] \Lambda D(x,k)$$

Now $D(0,k) = 0$. Thus (see [10, p 71, eqn (10)])

$$D(x,k) = \left[\int_0^1 \frac{\partial D}{\partial \hat{k}}(\nu x, k)\, d\nu\right] x$$

Thus

$$\dot{x} = -A(x,k)x \tag{4.A.2}$$

where

$$A(x,k) = \left[\frac{\partial \hat{K}^T}{\partial \hat{k}}\right] \Lambda \left[\int_0^1 \frac{\partial D}{\partial \hat{k}} (\nu x, k)\, d\nu\right] \tag{4.A.3}$$

Now $\dot{L} = x^T A^T(x,k) A(x,k) x$.

Thus if $A(0,k)$ is nonsingular for all k, then by continuity of A with respect to both its arguments, $\dot{L}$ is negative definite in the vicinity of $x = 0$, whence (4.A.2) is locally u.a.s. But by (4.A.3)

$$A(0,k) = \left.\frac{\partial\hat{K}^T}{\partial\hat{k}}\right|_{x=0} \Lambda \left.\frac{\partial\hat{K}^T}{\partial\hat{k}}\right|_{x=0}$$

$$= [I_N \ \bar{A}^T] \ \Lambda \begin{bmatrix} I_N \\ -- \\ \bar{A} \end{bmatrix}$$

for some matrix $\bar{A}$. Now the first N diagonal elements of Λ are positive and the rest are non-negative. Thus $A(0,k)$ is clearly positive definite.

APPENDIX 4.B : PROOF OF LEMMAS 4.1 and 4.2

Proof of Lemma 4.1

Recall

$$\dot{x}_i = - \sum_{\substack{r \subset S \\ i \varepsilon r}} \lambda_r [\prod_{\substack{\alpha \varepsilon r \\ \alpha \neq i}} \hat{k}_\alpha] [\prod_{\beta \varepsilon r} \hat{k}_\beta - \prod_{\beta \varepsilon r} k_\beta] \qquad (4.B.1)$$

Now, consider a point on one of the hyperplanes, e.g. $\hat{k}_j = 0$, bounding the orthant O defined in (4.59). Then

$$\dot{x}_j = - [\sum_{\substack{r \subset S \\ j \varepsilon r}} \lambda_r (\prod_{\substack{\alpha \varepsilon r \\ \alpha \neq j}} \hat{k}_\alpha k_\alpha)] x_j \qquad (4.B.2a)$$

where the quantity in square brackets is by hypothesis positive, being underbounded by λ_j . Thus $\dot{x}_j$ always points towards the interior of O . If $\hat{k} = 0$ then

$$\dot{x}_j = - \lambda_j x_j \quad \forall j \varepsilon S \; . \qquad (4.B.2b)$$

This completes the proof.

Proof of Lemma 4.2

We shall show that $\dot{x} = 0$ iff all $\lambda_r d_r = 0$. Thus since $\lambda_1, \ldots \lambda_N$ are positive this can happen iff $x = 0$. The "if" part of the proof is trivial. For the "only if" part consider the following cases.

Case I

Suppose $k_i \hat{k}_i > 0$ for all i . Consider

$\eta^T = \phi^T \text{diag}\{\hat{k}_i\}_{i=1}^N$, where $\phi_i = \ell n \frac{\hat{k}_i}{k_i}$. Then (4.B.1) implies (see also (4.56)) that

$$\eta^T\dot{x} = - \sum_{r \subset S} \lambda_r (\prod_{\alpha \varepsilon r} \hat{k}_\alpha) d_r \phi^T e_r$$

where e_r is a vector whose ith element is 1 if $i\varepsilon r$ and is zero otherwise. Now if $\dot{x} = 0$ then $\eta^T\dot{x} = 0$. We will show, then that unless all $\lambda_r d_r = 0$, $\eta^T\dot{x}$ and hence $\dot{x}$ does not equal zero.

$$(\prod_{\alpha \varepsilon r} \hat{k}_\alpha) d_r \phi^T e_r = (\prod_{\alpha \varepsilon r} \hat{k}_\alpha) d_r \sum_{i \varepsilon r} \ell n \frac{\hat{k}_i}{k_i}$$

$$= (\prod_{\alpha \varepsilon r} \hat{k}_\alpha) d_r \{\ell n((\prod_{i \varepsilon r} \hat{k}_i)(\prod_{j \varepsilon r} \hat{k}_j)) -$$

$$\ell n((\prod_{i \varepsilon r} \hat{k}_i)(\prod_{j \varepsilon r} k_j))\} \qquad (4.B.3)$$

$$= a_r .$$

Now

$$(\prod_{\alpha \varepsilon r} \hat{k}_\alpha) d_r \gtrless 0$$

<=>

$$(\prod_{\alpha \varepsilon r} \hat{k}_\alpha)[\prod_{j \varepsilon r} \hat{k}_j - \prod_{j \varepsilon r} k_j] \gtrless 0$$

<=> (using the case I assumption)

$$\frac{(\prod_{\alpha \varepsilon r} \hat{k}_\alpha)(\prod_{j \varepsilon r} \hat{k}_j)}{(\prod_{\alpha \varepsilon r} \hat{k}_\alpha)(\prod_{j \varepsilon r} k_j)} - 1 \gtrless 0$$

<=>

$$\ell n[(\prod_{\alpha \varepsilon r} \hat{k}_\alpha)(\prod_{j \varepsilon r} \hat{k}_j)] - \ell n[(\prod_{\alpha \varepsilon r} \hat{k}_\alpha)(\prod_{j \varepsilon r} k_j)] \gtrless 0$$

Thus a_r is always positive unless $d_r = 0$. Thus $\eta^T \dot{x} = 0$ iff $\lambda_r d_r = 0 \ \forall \ r \subset S$.

Case II

Suppose $k_i \hat{k}_i \geq 0$, with the equality holding for one or more $i \varepsilon S$. Suppose further that $k_j \hat{k}_j = 0$ and $k_j \neq \hat{k}_j$ for at least one j . If $\hat{k}_j = 0$, the calculation of $\dot{x}_j$ in the proof of Lemma 4.1 gives $\dot{x}_j \neq 0$ and so $\dot{x} \neq 0$. If $k_j = 0$, a very similar calculation again yields $\dot{x}_j \neq 0$. Finally suppose that whenever $k_j \hat{k}_j = 0$, $k_j = \hat{k}_j = 0$. If there are ℓ such k_j , then without loss of generality assume these to be $k_{N-\ell+1}, \ldots k_N$. Then since $\dot{x}_{N-\ell+1}$ to $\dot{x}_N$ will be zero one need consider only the parameters $k_1, \ldots k_{N-\ell}$, which have not been identified. Then using the analysis of case I on the reduced system one can show that $\dot{x} = 0$ implies

$\lambda_r d_r = 0 \ \forall \ r \subset \{S - \{N-\ell+1, \ldots, N\}\}$. Thus as $d_{N-\ell+1} \ldots d_N = 0$, this implies that $\lambda_r d_r = 0$, $\forall \ r \subset S$.

∇∇∇

APPENDIX 4.C : PROOF OF THEOREM 4.6

Notice that the conditions on $f_r(\cdot,\cdot)$ and non-negative $\lambda_r(0)$ ensure that $\lambda_r(t) \geq 0$ $\forall t$. Consider the Lyapunov function

$$\bar{L}(x(t),k,\lambda(t)) = L(x(t),k,\lambda(t)) + \frac{1}{2}\lambda^T(t)\lambda(t)$$

with L as in (4.55) and λ a vector having the diagonal elements of Λ_2 as its elements. Thus, for the region of space considered $\bar{L}$ is positive definite with respect to $[x^T(t),\lambda^T(t)]$ and is also decrescent and radially unbounded. Now, using (4.A.1) one can see that

$$-\dot{\bar{L}} = [\Lambda_1 x + \sum_{\substack{r \subset S\\|r|>1}} \lambda_r d_r \frac{\partial d_r}{\partial x}]^T [\Lambda_1 x + \sum_{\substack{r \subset S\\|r|>1}} \lambda_r d_r \frac{\partial d_r}{\partial x}]$$

$$- \sum_{\substack{r \subset S\\|r|>1}} f_r(\lambda) d_r^2 - \sum_{\substack{r \subset S\\|r|>1}} f_r(\lambda)\lambda_r$$

$$\geq [\Lambda_1 x + \sum_{\substack{r \subset S\\|r|>1}} \lambda_r d_r \frac{\partial d_r}{\partial x}]^T [\Lambda_1 x + \sum_{\substack{r \subset S\\|r|>1}} \lambda_r d_r \frac{\partial d_r}{\partial x}]$$

$$+ \sum_{\substack{r \subset S\\|r|>1}} \phi_r(\lambda) d_r^2 + \sum_{\substack{r \subset S\\|r|>1}} \phi_r(\lambda)\lambda_r = P(x,\lambda) .$$

$P(x,\lambda)$ is independent of t and depends on x and λ. Clearly $P(0,0) = 0$. If $\lambda_r > 0$ for any r then $P(x,\lambda) > 0$ by (4.67). If $\lambda = 0$, $P(x,\lambda) = x^T \Lambda_1^2 x$. Thus $P(x,\lambda)$ is positive definite in $[x^T(t), \lambda^T(t)]^T$ and hence the combined system of equations (4.65) and (4.66) is u.a.s.

APPENDIX 4.D: PROOF OF THEOREM 4.7

We shall prove theorem 4.7 by showing that all trajectories $[x^T(t), K_u^T(t) - K^T]^T$ converge uniformly asymptotically to zero.

Observe first of all that, denoting the elements of K_u by K_{ur}, $\forall r \subset S$, one can rewrite (4.68) as

$$\dot{x}_i = -\sum_{\substack{r \subset S \\ i \varepsilon r}} \lambda_r [\prod_{\substack{\alpha \varepsilon r \\ \alpha \neq i}} \hat{k}_\alpha][(\prod_{\beta \varepsilon r} \hat{k}_\beta) - K_{ur}]$$

Thus if $\hat{k}_i$ becomes large

$$\dot{x}_i \simeq -\{\sum_{\substack{r \subset S \\ i \varepsilon r}} \lambda_r [\prod_{\substack{\alpha \varepsilon r \\ \alpha \neq i}} \hat{k}_\alpha^2]\}\hat{k}_i$$

whence if $\hat{k}_i$ becomes large the adaptive law acts to reduce its magnitude. Thus it is clear that $\hat{K}$ remains bounded. Furthermore, by assumption (ii) K_u is also bounded.

Now, the exponential asymptotic stability of the adaptive law which generates $K_u(t)$ ensures the existence of a Lyapunov function L_1 such that

$$c_1 \| K_u(t) - K \|^2 \leq L_1 \leq c_2 \| K_u(t) - K \|^2 \qquad (4.D.1)$$

and

$$\dot{L}_1 \leq -c_3 \| K_u(t) - K \|^2 \qquad (4.D.2)$$

where c_1, c_2 and $c_3 > 0$ (see [11, p 86]).

Moreover as $\hat{k}(t)$ is bounded so is $\frac{\partial \hat{K}(t)}{\partial \hat{k}(t)}$; viz

$$\left\| \frac{\partial \hat{K}^T(t)}{\partial \hat{k}(t)} \Lambda \right\| < M$$

Then consider the Lyapunov function $\frac{1}{2}L + \beta L_1$ with $\beta = 1 + M^2/2c_3$, and L defined as in (4.55)

$$\frac{1}{2}\dot{L} + \beta\dot{L}_1 = -\left\| \frac{\partial L}{\partial x} \right\|^2 + \left[\frac{\partial L}{\partial x} \frac{\partial \hat{K}^T(t)}{\partial \hat{k}(t)} \Lambda\ (K_u(t) - K)\right]$$

$$+ \beta\dot{L}_1$$

$$\leq -\left\| \frac{\partial L}{\partial x} \right\|^2 + M \left\| \frac{\partial L}{\partial x} \right\| \ \|K_u(t) - K\| - \beta c_3 \| K_u(t) - K\|^2$$

$$= -\tfrac{1}{2}\left\| \frac{\partial L}{\partial x} \right\|^2 - c_3\|K_u(t) - K\|^2 - \left(\frac{1}{\sqrt{2}} \left\|\frac{\partial L}{\partial x}\right\| \right.$$

$$\left. - \sqrt{c_3(\beta-1)}\ \|K_u(t) - K\|\right)^2$$

$$\leq -\tfrac{1}{2}\left\| \frac{\partial L}{\partial x} \right\|^2 - c_3 \| K_u(t) - K\|^2$$

which is negative definite in $[x^T, K_u^T(t) - K^T]^T$ by condition (i) of the theorem. This completes the proof.

∇∇∇

APPENDIX 4.E : PROOF OF THEOREM 4.9

We shall show first of all that the result of Lemma 4.1 applies for every fundamentally identifiable system. To do so we consider in turn the following two cases.

Case I   Let $\hat{k}_i = 0$ and d_i be observable. Clearly by (4.B.2a), the trajectory points in towards O.

Case II   Let $\hat{k}_i = 0$ and d_i be unobservable. Then by (ii) and (4.B.2a) we have that

$$\dot{x}_i = -\{\lambda_{r_i} (\prod_{\substack{\alpha \varepsilon r_i \\ \alpha \neq i}} \hat{k}_\alpha k_\alpha) + a(\hat{k})\} x_i \qquad (4.E.1)$$

where $a(\hat{k}) \geq 0$, $\lambda_{r_i} > 0$ and all $\hat{k}_\alpha$ in the first term, by Case I, must eventually become nonzero with the same sign as k_α. Moreover, there must be at least one r_i for which $k_\alpha \neq 0$ for every $\alpha \varepsilon r_i$ and $\alpha \neq i$, as otherwise k_i will not be uniquely identifiable. Thus the right hand side of (4.E.1) has the opposite sign to x_i and the trajectory points in towards O.

Thus lemma 4.1 applies. Furthermore, considerations similar to those in lemma 4.2, together with (i), reveal that $\dot{x} = 0$ if and only if all observable d_r equal zero. Thus fundamental identifiability of (4.14) and arguments similar to those presented in proving theorem 4.5 show that $\dot{L}$ is negative definite everywhere in O. This completes the proof.

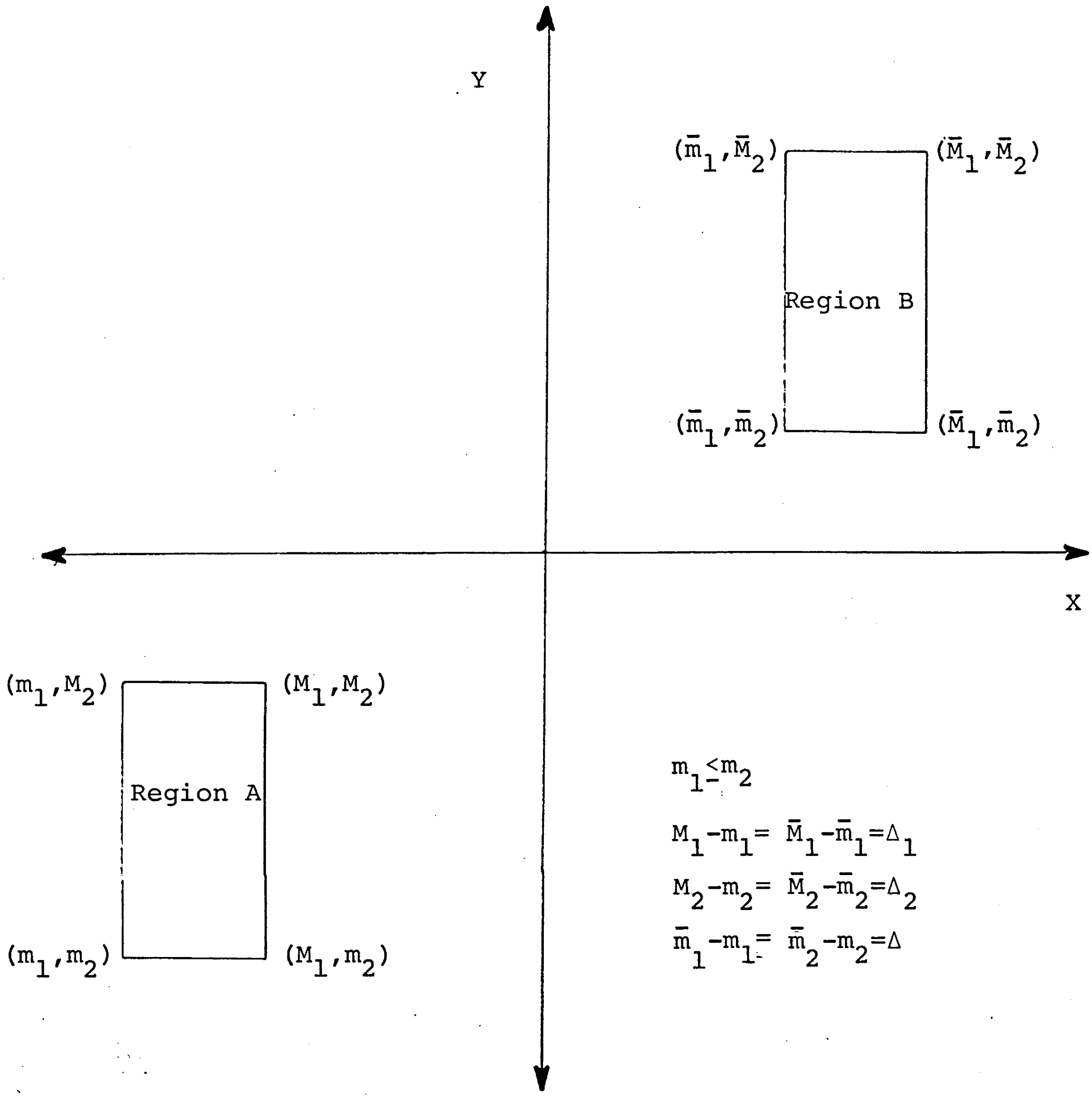


Figure 4.0 Illustration of the parameter translation scheme.

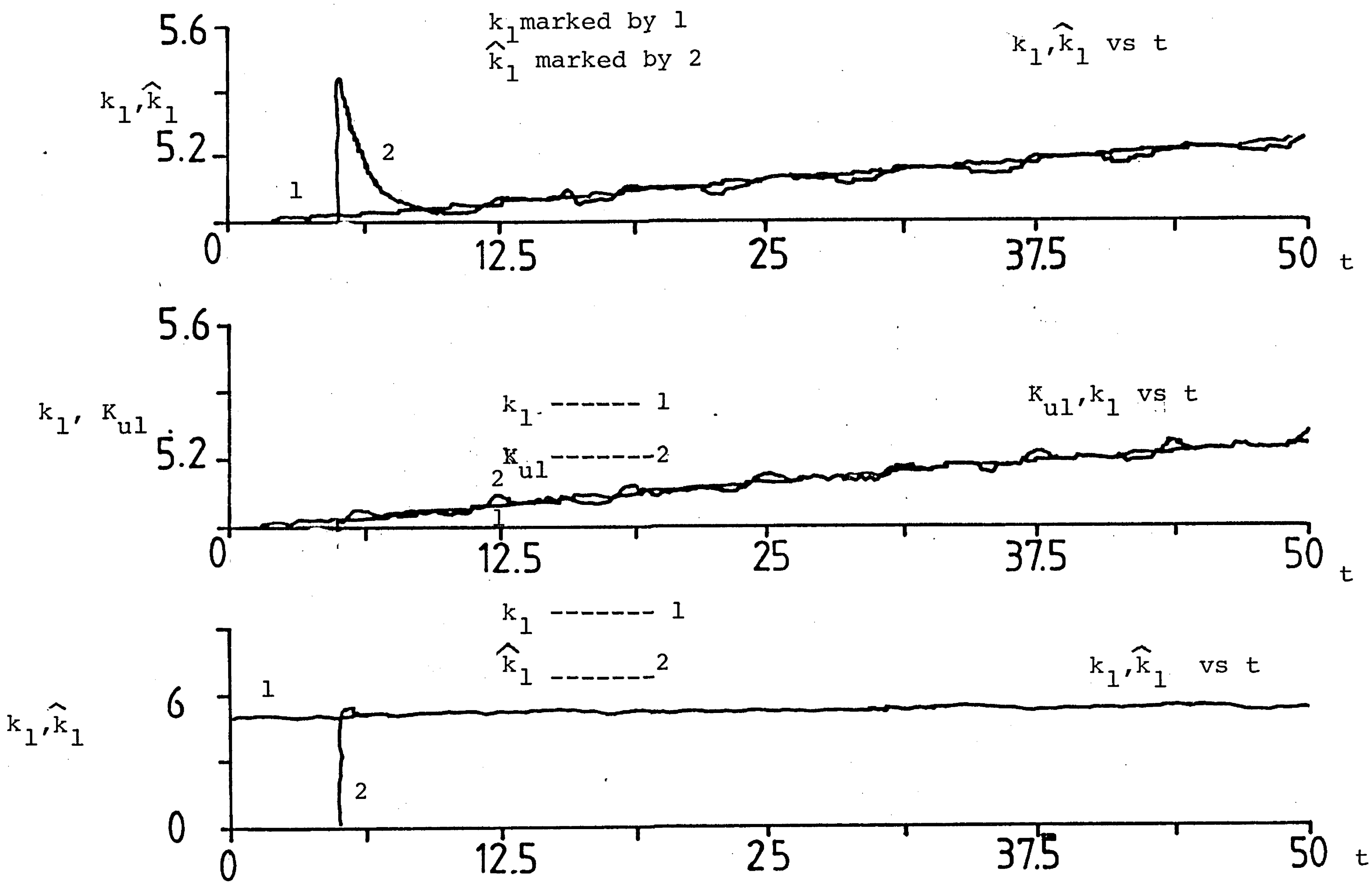


Figure 4.1 Sinusoidal drift in the k_i

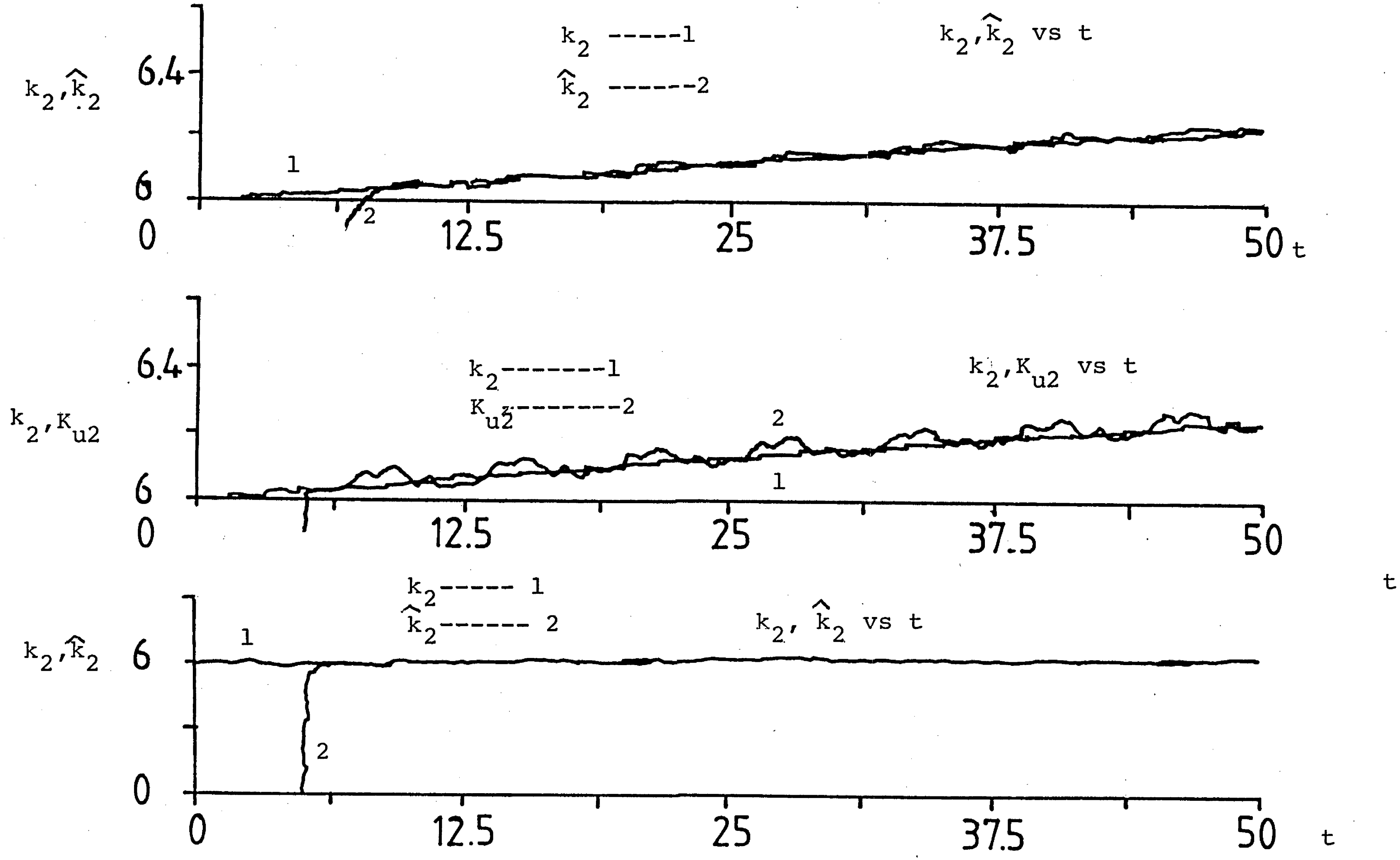


Figure 4.2 Sinusoidal drift in the k_i.

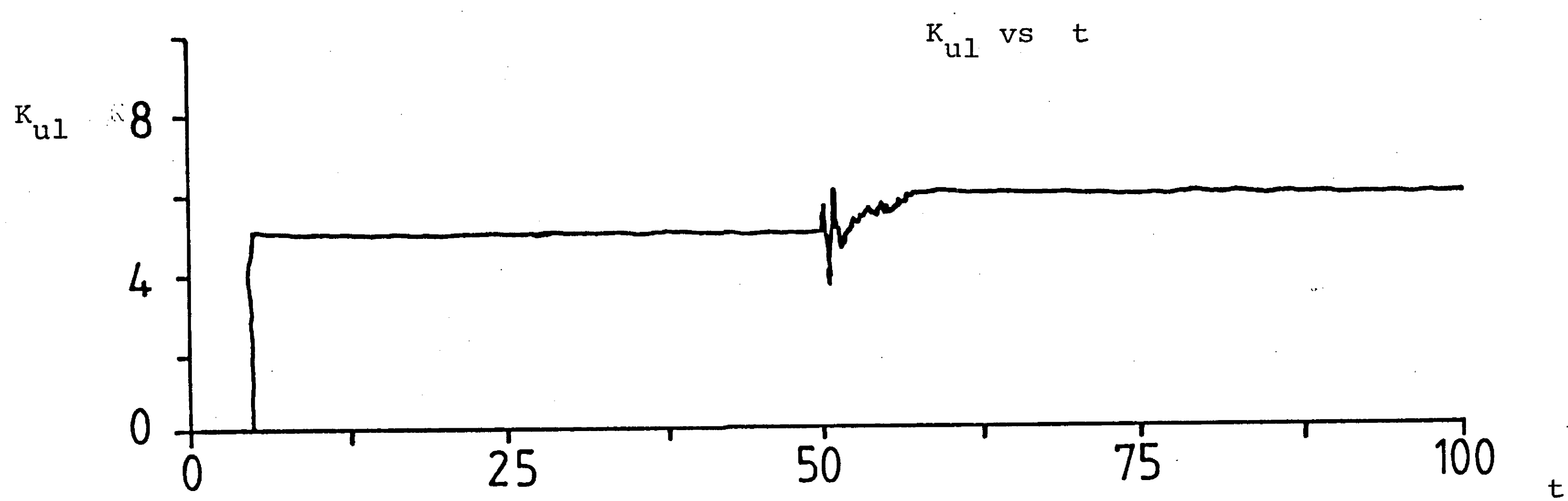


Figure 4.3 Step change in the k_i of 1 each.

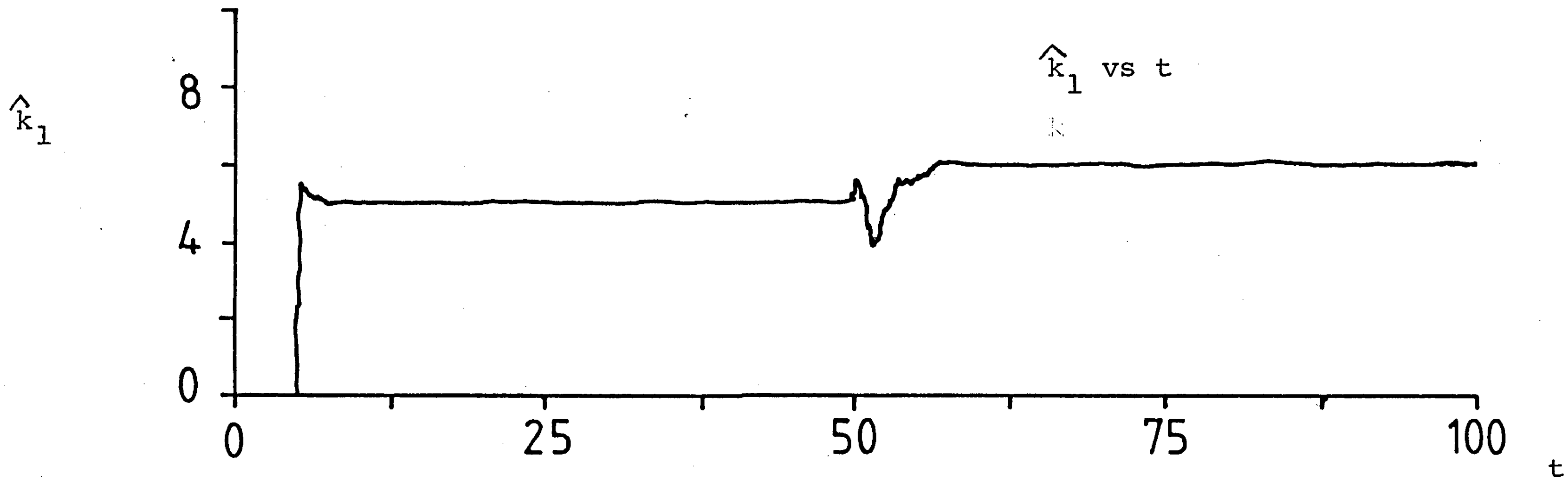


Figure 4.4 Step change in each k_i of 1.

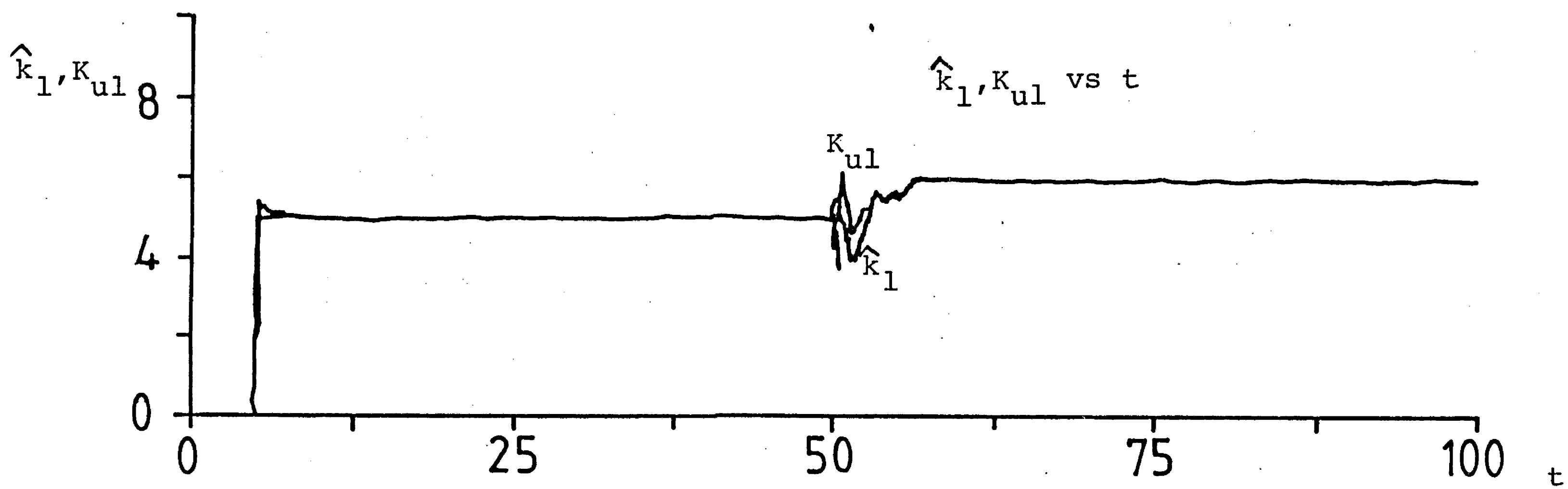


Figure 4.5 Step change in the k_i of one each.

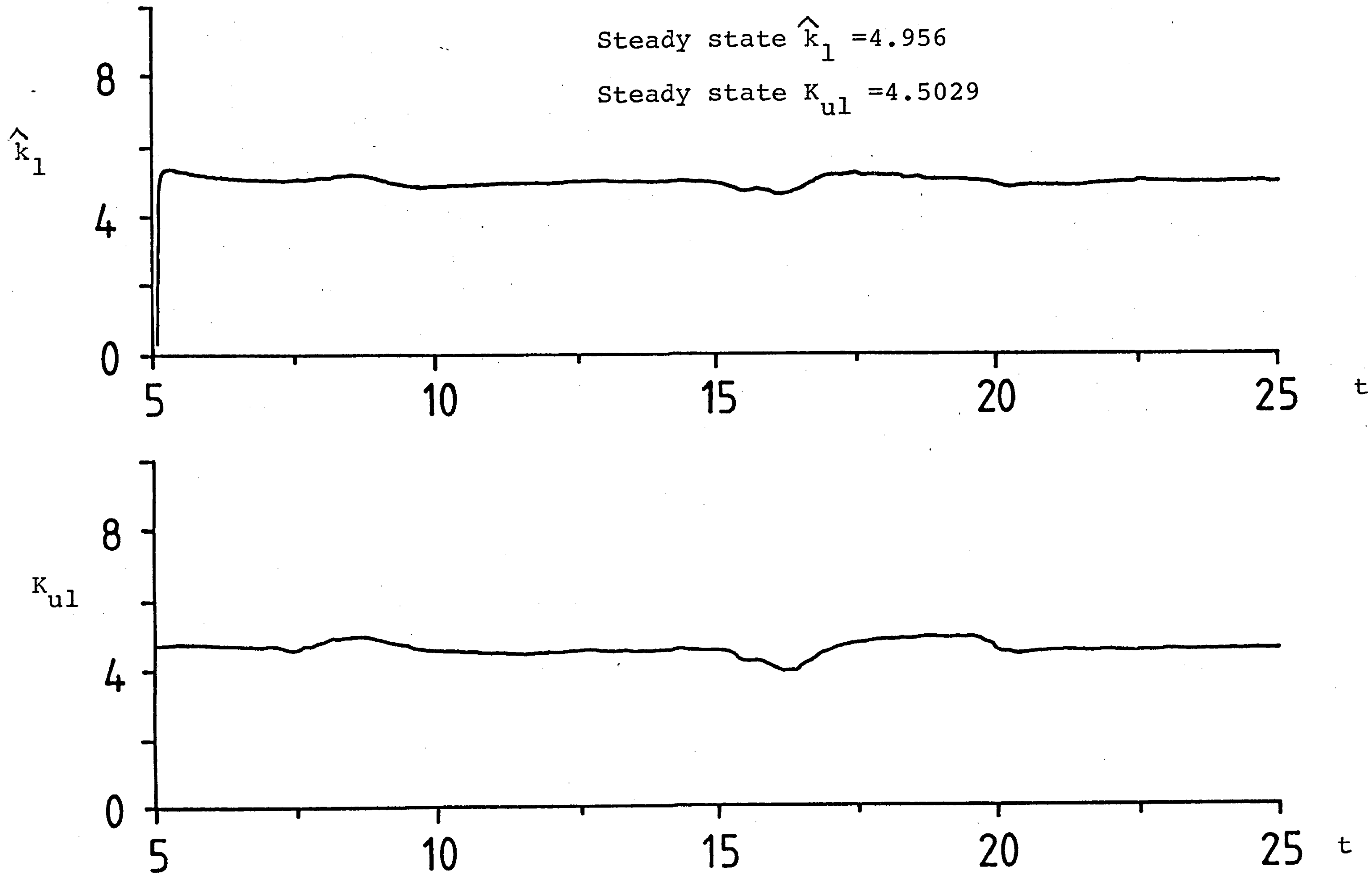


Figure 4.6 Identification with unmodelled modes.

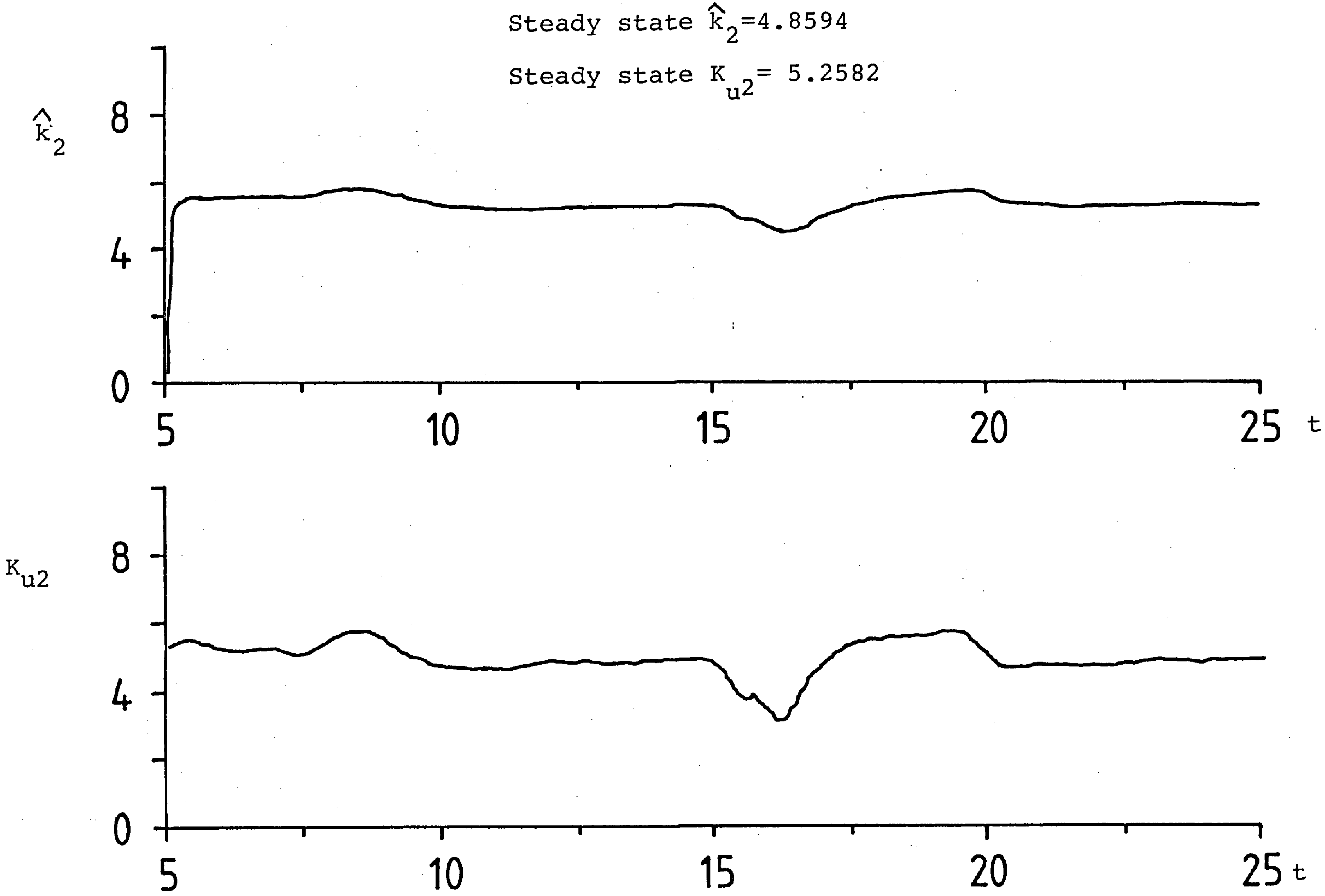


Figure 4.7 Identification with unmodelled modes.

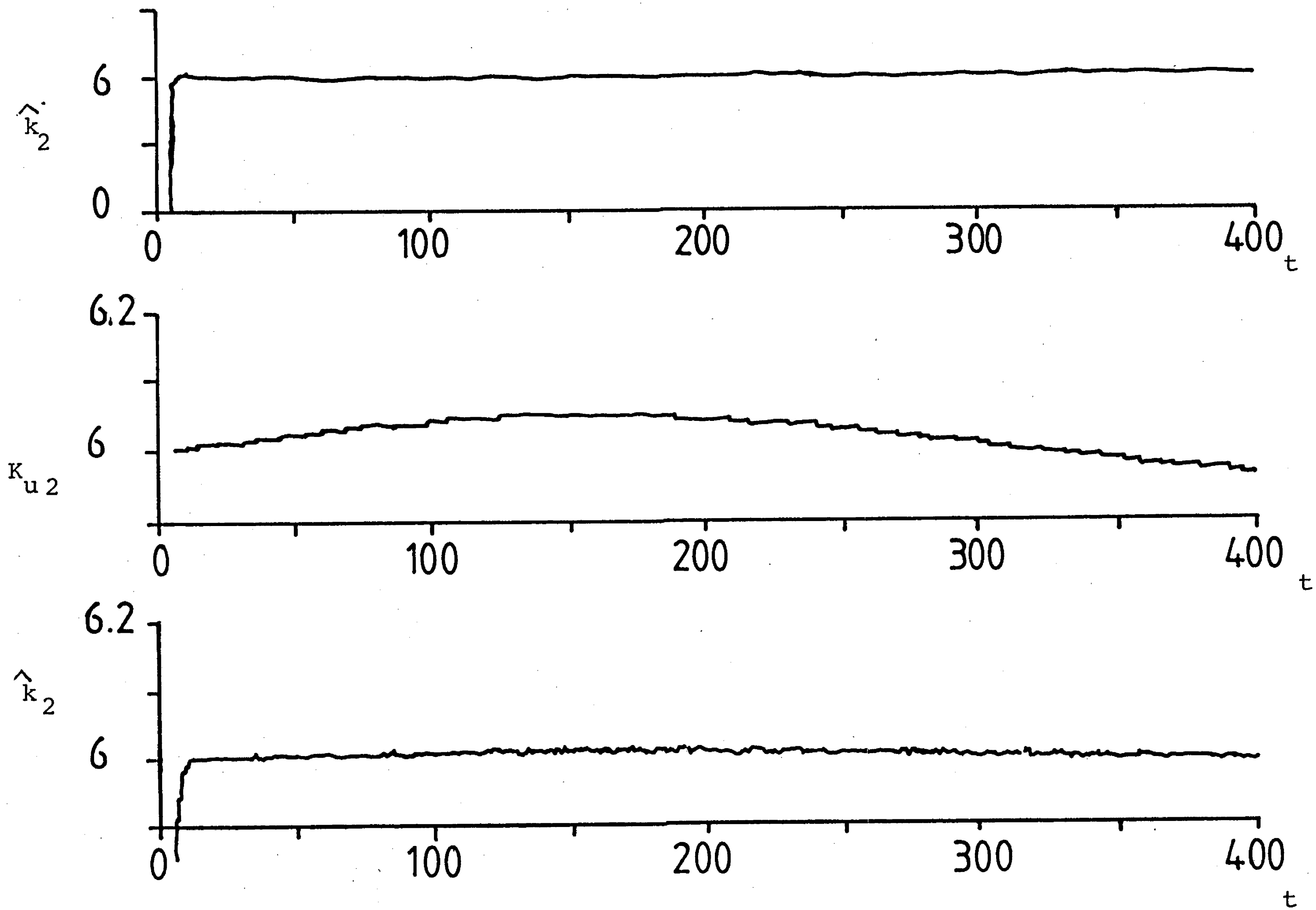


Figure 4.8 Sinusoidal drift in the K_{ui}.

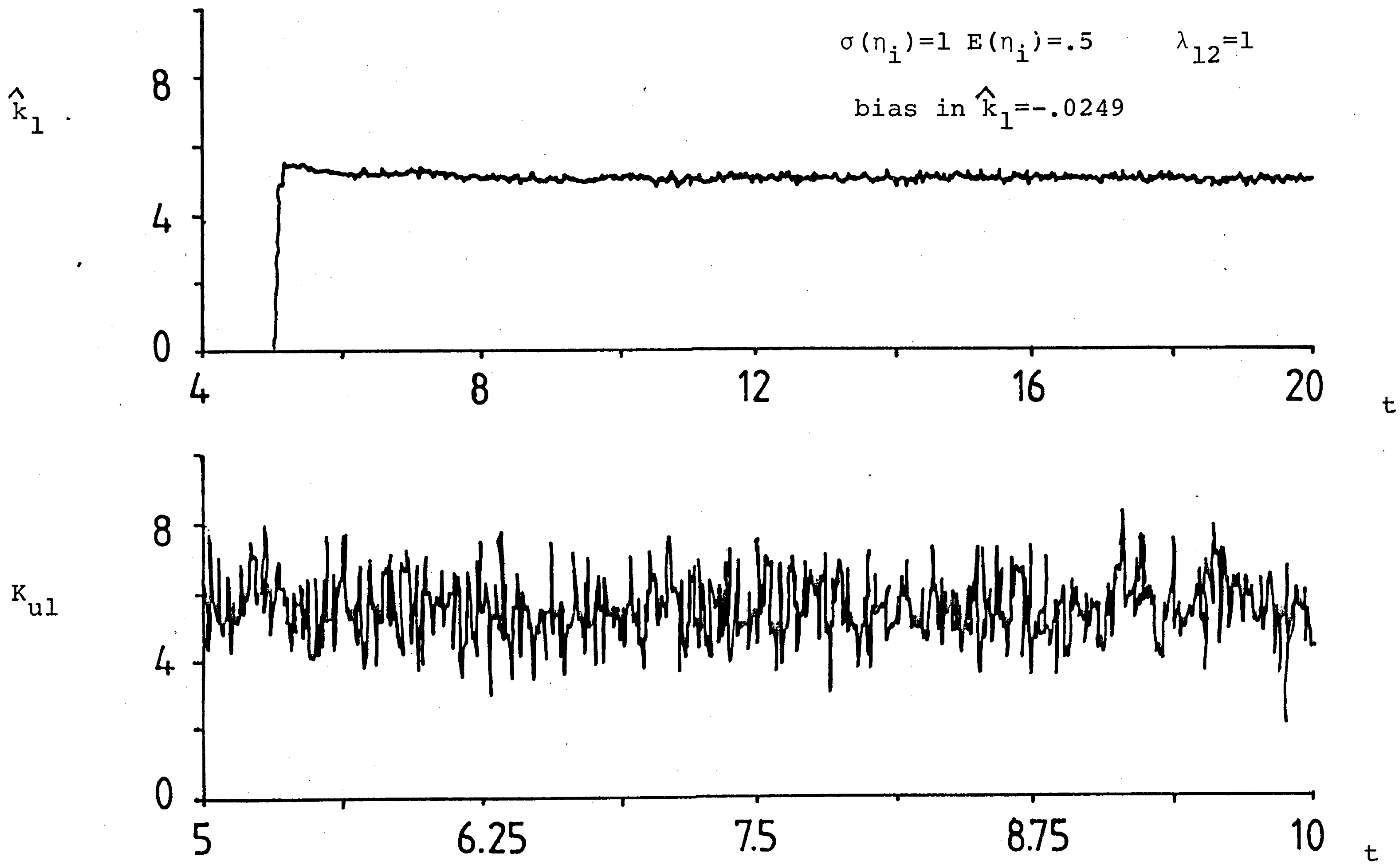


Figure 4.9 Biased noise in K_u.

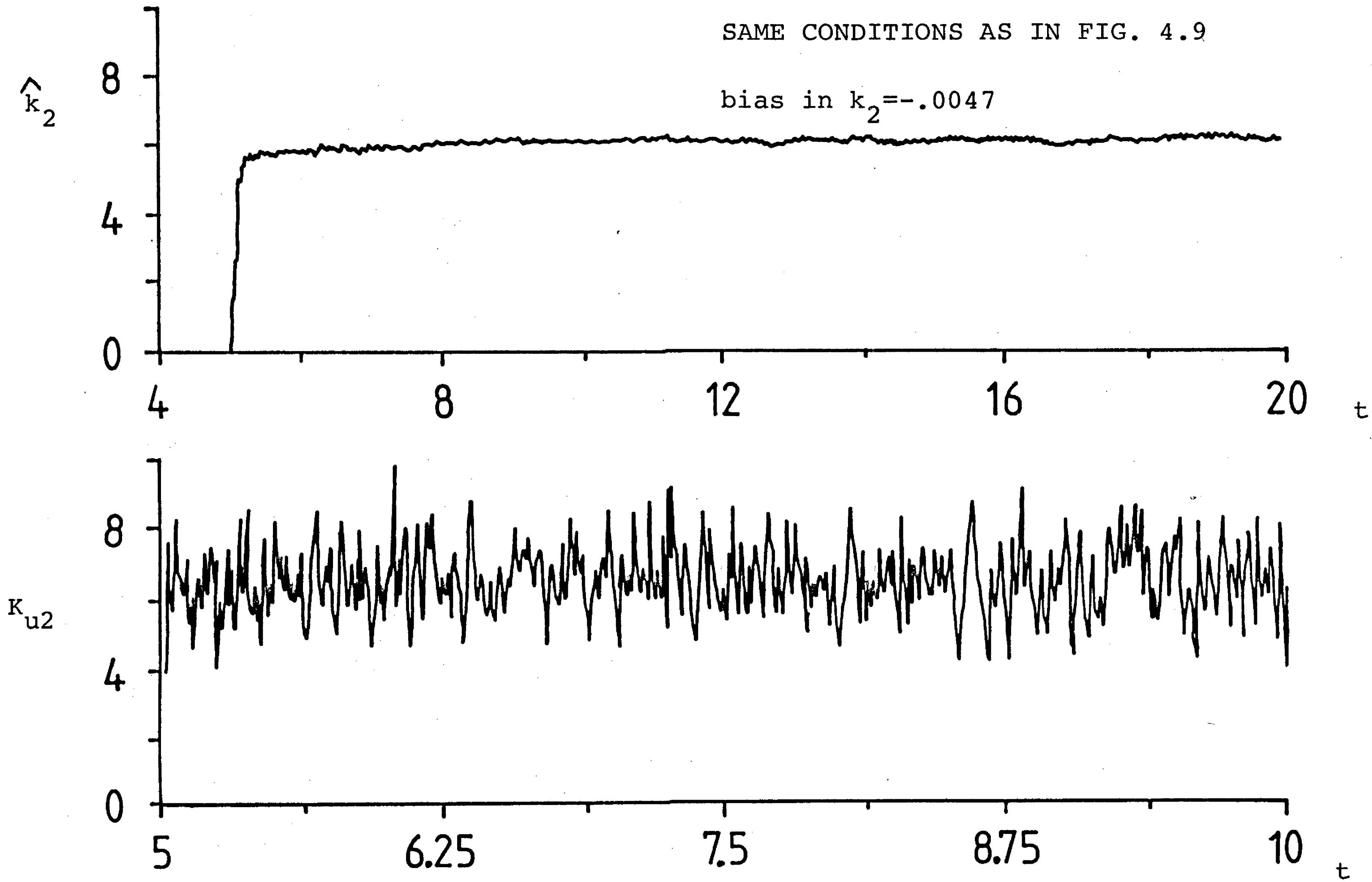


Figure 4.10 Biased noise in K_u.

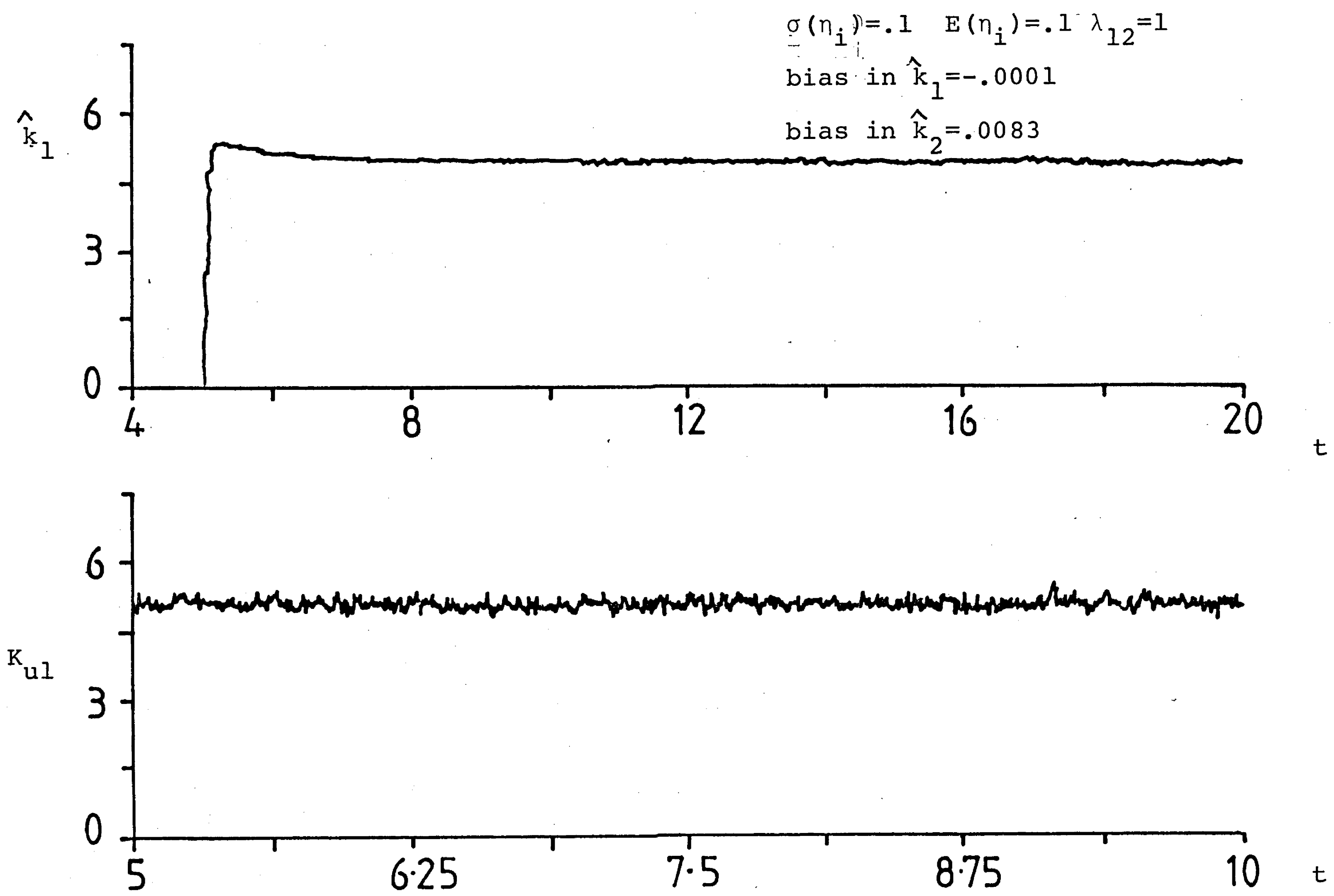


Figure 4.11 Biased noise in Ku

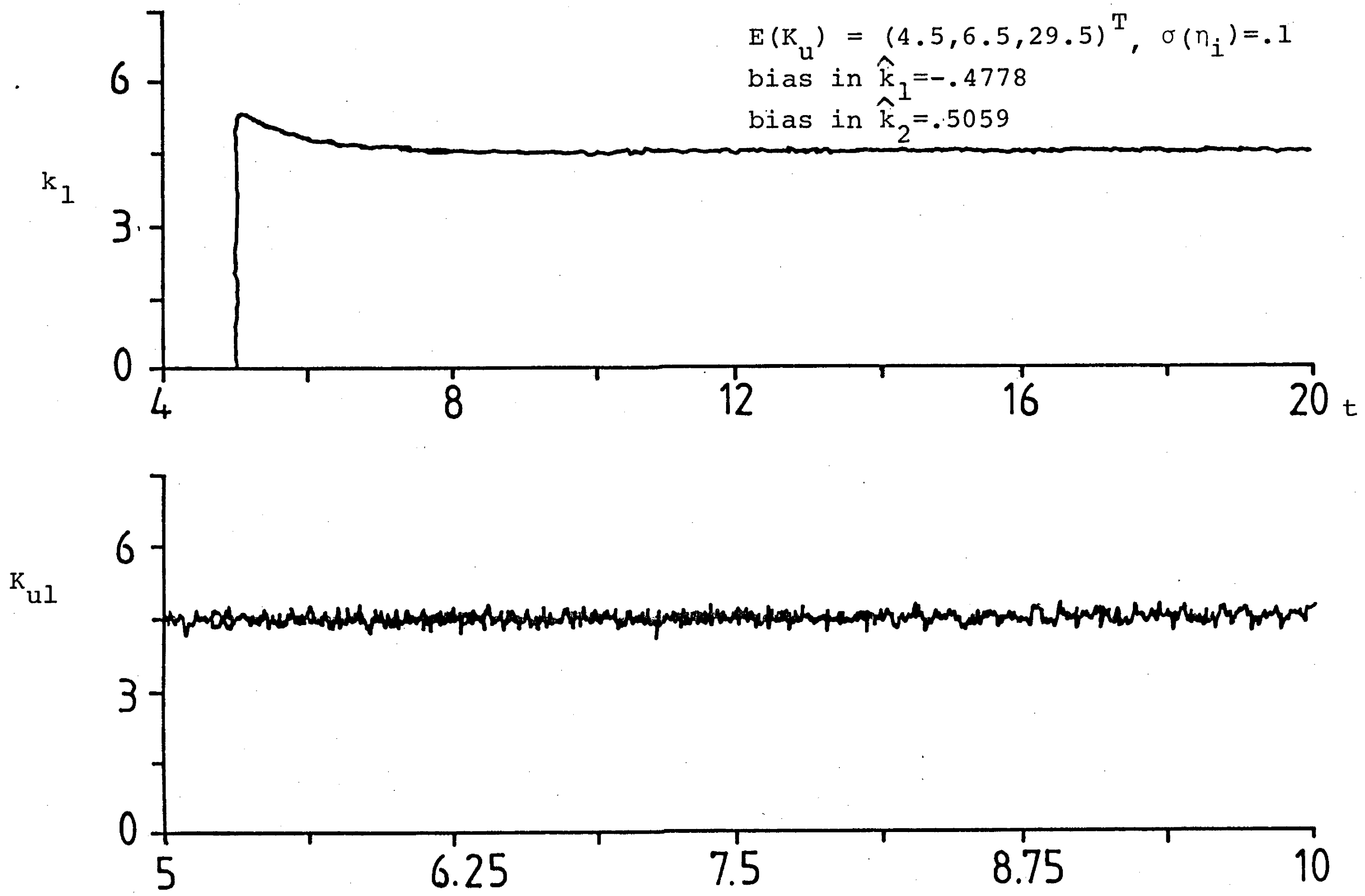


Figure 4.12 Effect of Bias in K_u.

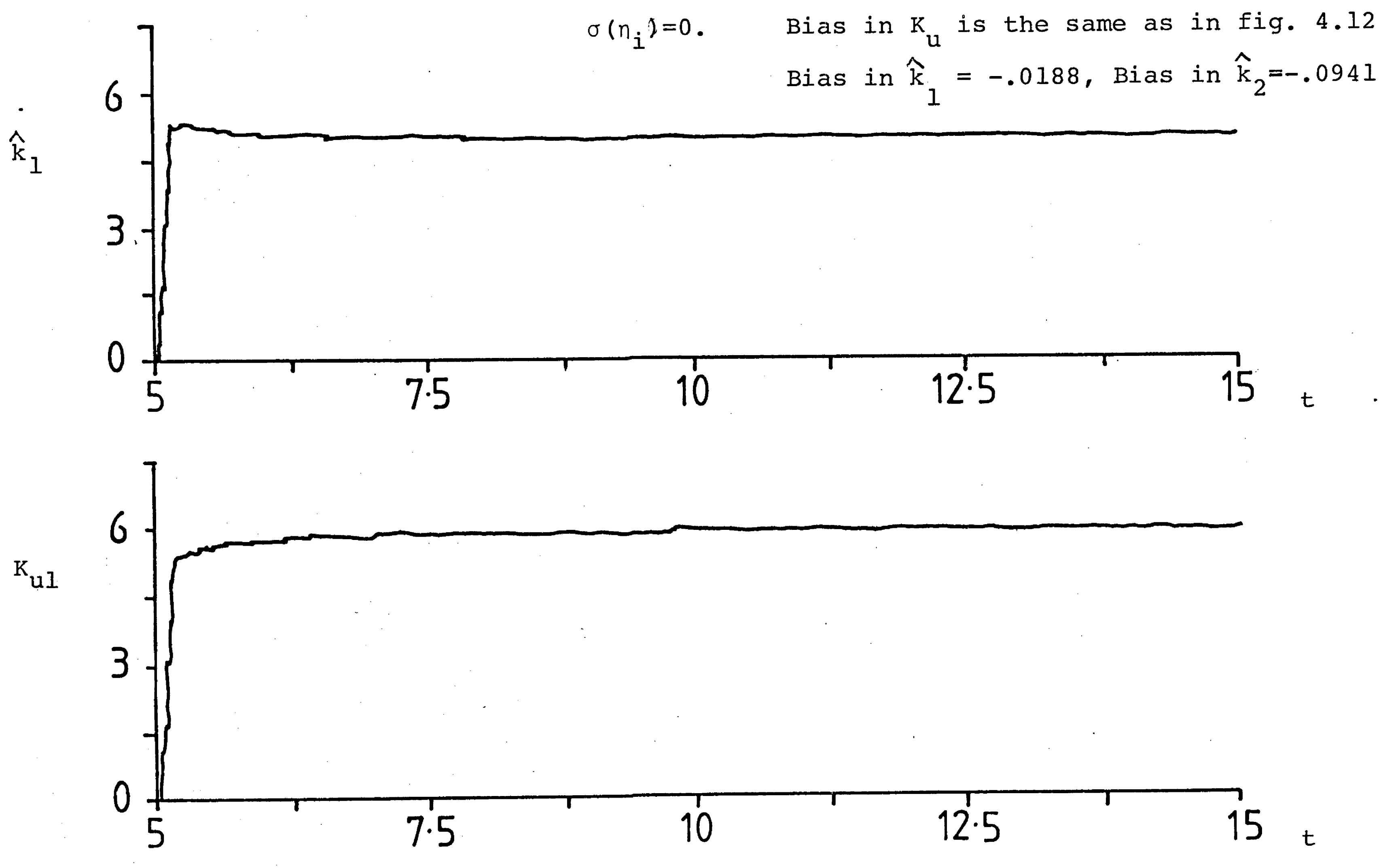


Figure 4.13 Effect of bias in K_u

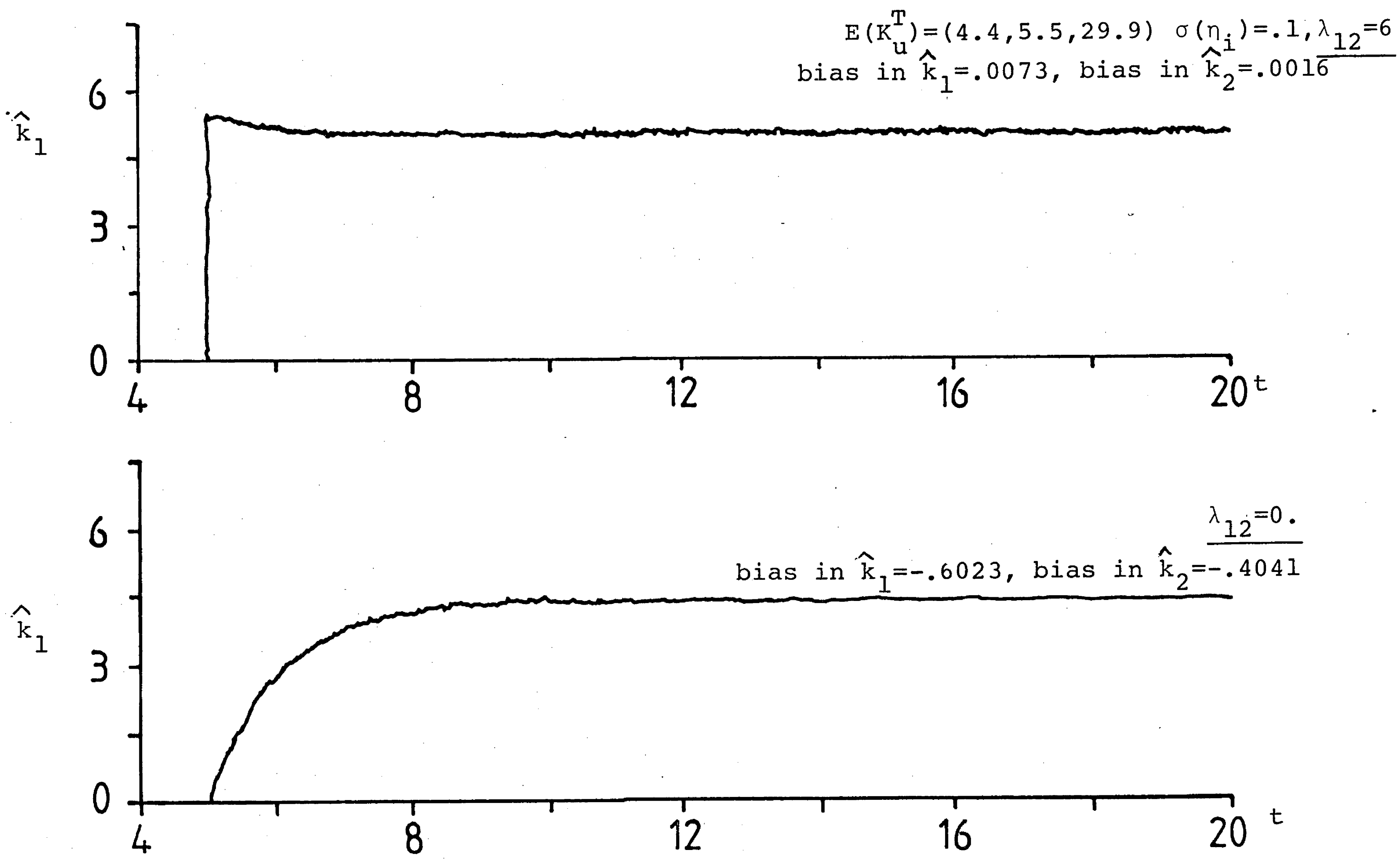


Figure 4.14 The role of λ_{12} in bias reduction.

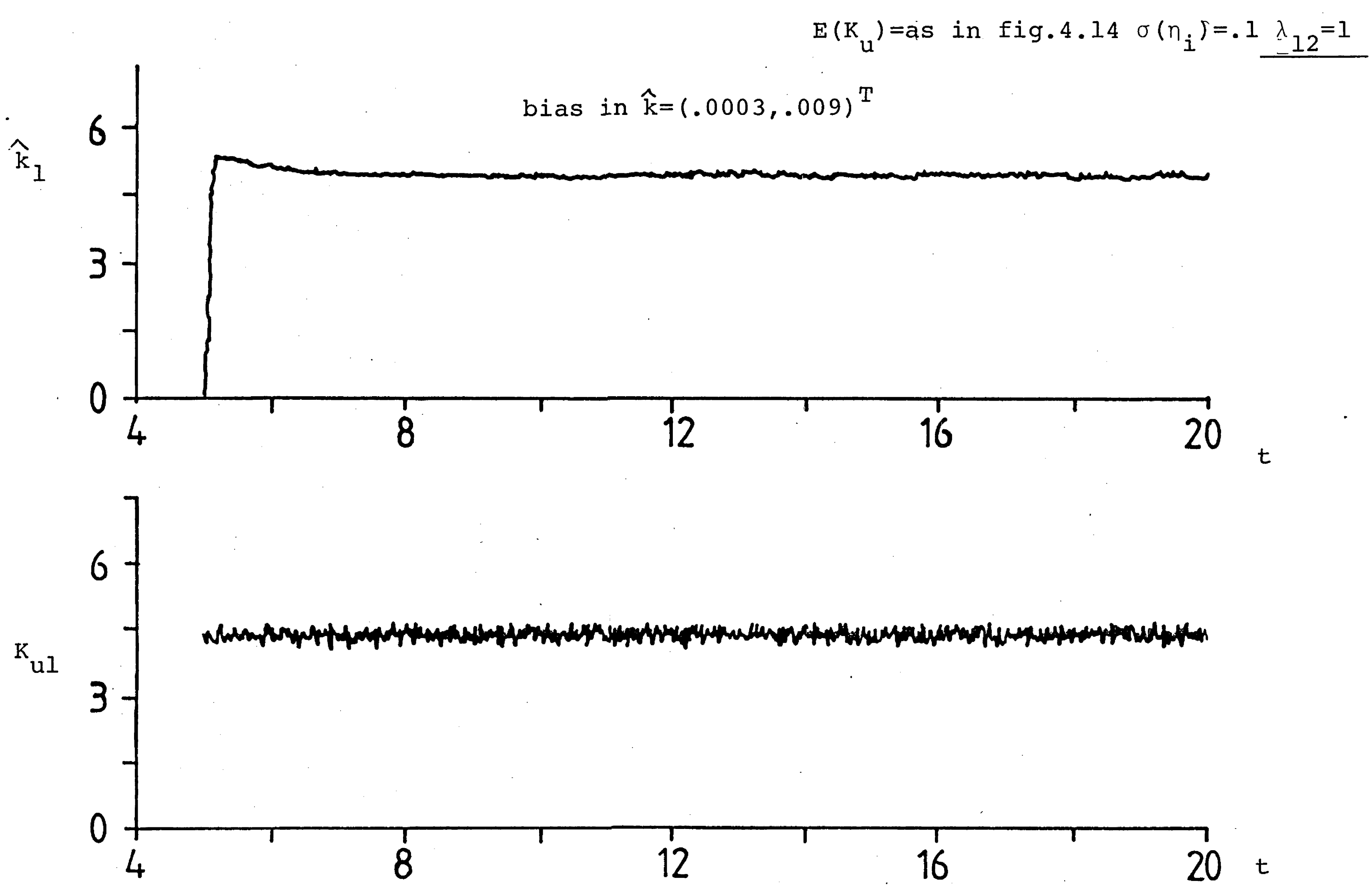


Figure 4.15 Role of λ_{12} in bias reduction.

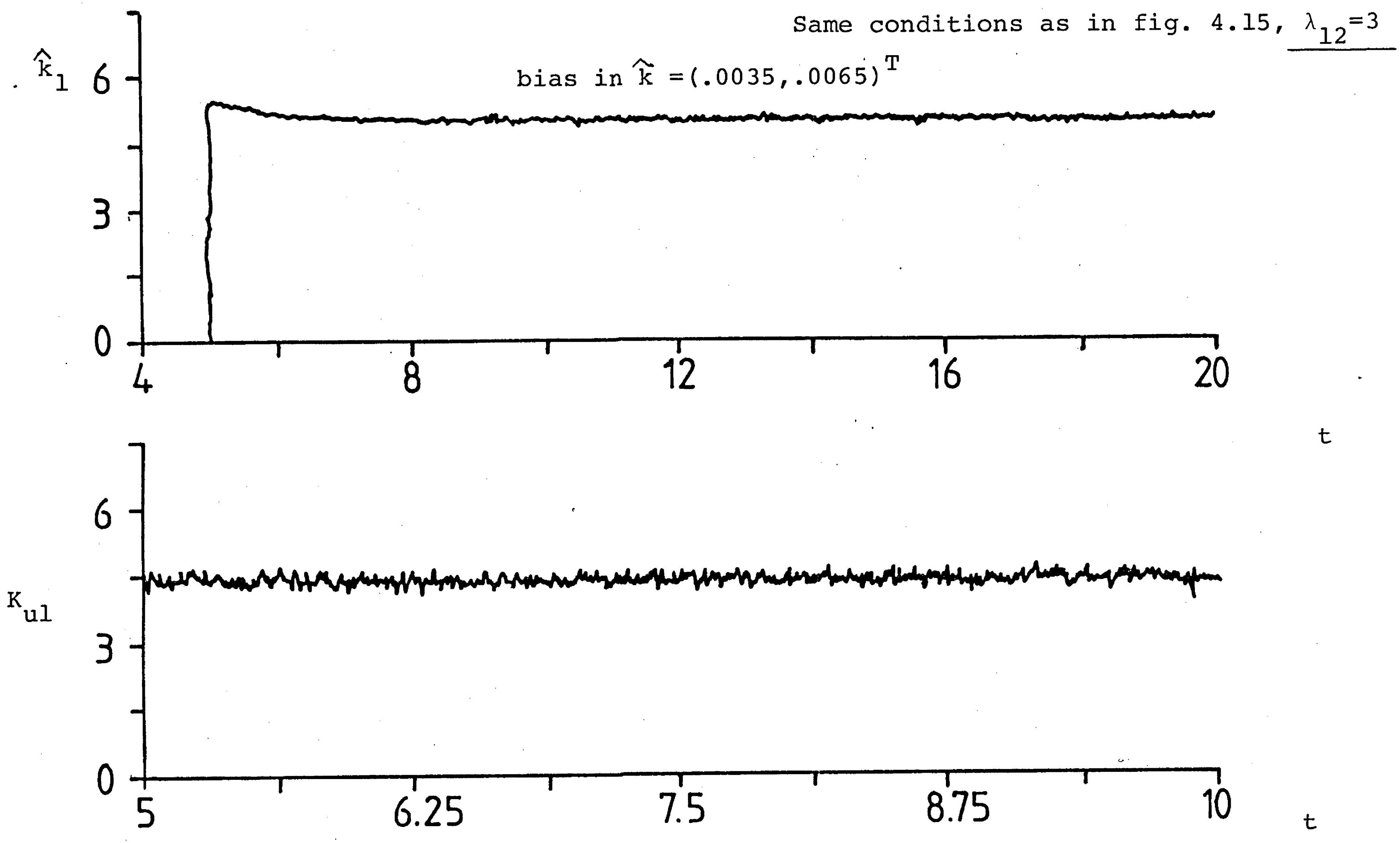


Figure 4.16 Effect of λ_{12} in noise reduction.

References for Chapter 4

1. P.M.Lion, "Rapid identification of linear and nonlinear systems", AIAA Journal, vol 5, Oct 1967, pp 1835-1842.

2. K.S. Narendra and P. Kudva, "Stable adaptive schemes for system identification and control - Parts I and II", IEEE Transactions Syst Man Cybern, vol SMC-4, pp 542-560, Nov 1974.

3. G. Luders and K.S. Narendra, "Stable adaptive schemes for estimation and identification of linear systems", IEEE Transactions Auto Contr, vol AC-19, pp841-847, Dec 1974.

4. R.L. Carroll and D.P. Lindorff, "An adaptive observer for single input single output linear systems." IEEE Trans Auto Contr, AC-18, No. 5, pp 428-435,1973.

5. B.D.O. Anderson,"An approach to multivariable system identification", Automatica, vol 13, pp 401- 408, 1977.

6. G. Kreisselmeier, "Adaptive observers with exponential rate of convergence", IEEE Trans Auto Cont, AC-22, Feb 1977, pp 2-8.

7. G. Kreisselmeier,"Algebraic separation in realizing a linear state feedback control law by means of adaptive observers". IEEE Trans Auto Cont, AC-25, April 1980.

8. S. Dasgupta, B.D.O. Anderson and R.J. Kaye,"Robust identification of partially known systems". Proc. 22nd CDC San Antonio, Texas, vol 1, pp 1510-1514, Dec 1983.

9. S. Dasgupta, B.D.O. Anderson and R.J. Kaye,"Identification of economically parameterized systems".Preprints 9th IFAC World Congress, Hungary, vol X, pp 96-100, July 1984.

10. J.M. Ortega and W.C. Rheinboldt, Iterative Solutions of nonlinear equations in several variables. Academic Press, 1970.

11. N.N. Krasovskii, Stability of motion, Stanford University Press, 1963.

12. G. Kreisselmeier, "An approach to stable indirect adaptive control." Submitted for publication.

§5 Output Error Identification

In this chapter we formulate output error algorithms for the identification of systems described in Chapters 2 and 4. The difference between output and equation error identification is illustrated in figure 1.1. In equation error the system output y, enters the adjustable model twice, directly and through the output error, Δy. Consequently, terms like $y^2(t)$ appear in the error model. Thus unbiased measurement noise could yet result in biased parameter estimates. In output error this difficulty is avoided by allowing the output to enter the adjustable model via Δy only.

Exponential convergence of output error algorithms, however, require that a certain transfer function be strictly positive real. Unfortunately, this transfer function depends on the unknown system parameters, and the condition cannot always be checked. A detailed exposition on output error identification can be found in [1].

The two output error algorithms presented here conform to the two step structure outlined in Chapter 4. Thus while the first steps in the two algorithms are different (one is gradient descent and the other recursive least squares) the second step is still given by equation (4.68). The convergence proofs of this chapter require a key lemma, given in the appendix, which extends a result by Boyd and Sastry [2], to the unbounded signal case. All our results require that magnitude bounds on the parameters be known, a requirement which, as emphasized before, is easy to satisfy given the parametrisations under consideration.

Sections 5.1 and 5.2 give the two output error algorithms together with their convergence proofs. The contents of this chapter are the subject of [3].

5.1 A gradient descent output error algorithm

Let the asymptotically stable unknown system be described by

$$Q(s,k)y(t) = P(s,k)u(t) \tag{5.1}$$

where k is the N dimensional unknown parameter vector, $m = \delta[Q(s,k)] \geq \delta[P(s,k)]$ and

$$Q(s,k) = q_o(s) + \sum_{r \subset S} \left(\prod_{i \varepsilon r} k_i \right) q_r(s) \ ;$$

$$P(s,k) = p_o(s) + \sum_{r \subset S} \left(\prod_{i \varepsilon r} k_i \right) p_r(s) \ ;$$

$S \triangleq \{1,2, \ldots N\}$; $\delta[q_o(s)] \geq \delta[q_r(s)] \ \forall r \subset S$ and $p_r(s)$, $q_r(s)$ known.

As before we shall abuse notation by referring to y(s) as the Laplace transform of y(t). Likewise,

$$y_i(t) = s^i y(t)/(s+\alpha)^n$$

will denote the solution of the differential equation

$$(p+\alpha)^n \ y_i(t) = p^i y(t)$$

with $p \triangleq d/dt$, and with arbitrary finite initial conditions. When referring to vectors of the form

$$W^T \triangleq [u(t), \ \frac{1}{s+\beta} \ u(t) \ \ldots, \ \frac{1}{(s+\beta)^n} u(t)]$$

however, the initial conditions will be assumed to be non-zero.

Let K be a vector containing the multilinear combinations of the k_i and K_u an unconstrained estimate of K. Let $\beta(s)$ be a polynomial of degree n Hurwitz in s. Then for the following adjustable model:

$$\left\{\frac{q_o(s)}{\beta(s)} + \sum_{r\subset S} K_{ur}(t)\,\frac{q_r(s)}{\beta(s)}\right\}\hat{y}(t) = \left\{\frac{p_o(s)}{\beta(s)} + \sum_{r\subset S} K_{ur}(t)\,\frac{p_r(s)}{\beta(s)}\right\}u(t) \tag{5.2}$$

the lemma below relates the output error $\Delta y(t) = \hat{y}(t) - y(t)$ to the parameter error $\Delta K_u(t) = K_u(t) - K$. Note $K_{ur}(t)$ are elements of $K_u(t)$.

Lemma 5.1

Define $\hat{V}(t)$ as the vector whose elements are

$\frac{p_r(s)}{\beta(s)}u(t) - \frac{q_r(s)}{\beta(s)}u(t)$. $\forall r\subset S$. Then

$$\Delta y(t) = \frac{\beta(s)}{Q(s,k)}\{\hat{V}^T(t)\,\Delta K_u(t)\} \tag{5.3}$$

Proof:

Equation (5.1) can be re-expressed as

$$\left\{\frac{q_o(s)}{\beta(s)} + \sum_{r\subset S}\left(\prod_{i\varepsilon r} k_i\right)\frac{q_r(s)}{\beta(s)}\right\}y(t)$$

$$= \left\{\frac{p_o(s)}{\beta(s)} + \sum_{r\subset S}\left(\prod_{i\varepsilon r} k_i\right)\frac{p_r(s)}{\beta(s)}\right\}u(t) \tag{5.4}$$

Subtracting (5.4) from (5.2) we get

$$\frac{q_o(s)}{\beta(s)}\Delta y(t) + \sum_{r\subset S}\left(K_{ur}(t)\frac{q_r(s)}{\beta(s)}\hat{y}(t) - \left(\prod_{i\varepsilon r} k_i\right)\frac{q_r(s)}{\beta(s)}y(t)\right)$$

$$= \sum_{r\subset S}\left(K_{ur}(t) - \prod_{i\varepsilon r} k_i\right)\frac{p_r(s)}{\beta(s)}u(t)$$

$$\Leftrightarrow \frac{Q(s,k)}{\beta(s)}\Delta y(t) = \sum_{r\subset S}\left(K_{ur} - \prod_{\alpha\varepsilon r} k_i\right)\left\{\frac{p_r(s)}{\beta(s)}u(t) - \frac{q_r(s)}{\beta(s)}\hat{y}(t)\right\}$$

$$= \hat{V}^T(t)\ \Delta K_u(t)$$

∇∇∇

<u>Remarks</u>:

(5.1) The adjustable system can be implemented by re-writing (5.2) as

$$\hat{y}(t) = \frac{\beta(s)-q_o(s)}{\beta(s)}\hat{y}(t) + \sum_{r\subset S} K_{ur}(t)\left[\frac{p_r(s)}{\beta(s)}u(t) - \frac{q_r(s)}{\beta(s)}\hat{y}(t)\right].$$

Of the following assumptions 5.1 will be in force throughout this chapter while 5.2 will hold for this section.

<u>Assumption 5.1</u>:

The bounds on the magnitude of the unknown parameters,

$$m_i \leqslant k_i \leqslant M_i\ , \qquad \forall i\varepsilon S$$

are known *a priori*, with $m_i > 0$.

<u>Assumption 5.2</u>

The transfer function $\beta(s)/Q(s,k)$ is strictly positive real (SPR).

Remarks

(5.2) The positiveness of m_i can be ensured by a suitable translation of the k_i. The rank-1 property (see chapter 2) is, of course, independent of such a translation. The problem of selecting a $\beta(s)$, which ensures that Assumption 5.2 is satisfied for all k belonging to the prescribed region, remains an open question. However, as opposed to the situation with the more conventional parametrizations, determination of $\beta(s)$ may well be easier here as a great deal is known about $Q(s,k)$.

The proposed two step algorithm, then, is

$$K_u(t) = -\hat{V}(t)\Delta y(t); \qquad \prod_{i\varepsilon r} m_i \leqslant K_{ur}(0) \leqslant \prod_{i\varepsilon r} M_i, \ \forall r \subset S \tag{5.5}$$

$$\dot{\hat{k}} = -\left[\frac{\partial \hat{K}(t)}{\partial \hat{k}(t)}\right]^T \Lambda \ (\hat{K}(t) - K_u(t)) - \Gamma \ \Psi \ (\hat{k}(t)), \tag{5.6}$$

$$m_i \leqslant \hat{k}_i(0) \leqslant M_i; \ \forall i \varepsilon S$$

Here $\hat{K}$ is a constrained estimate of K (see Section 4.1.1) $\hat{k}$ is the corresponding estimate of k, while Λ is a diagonal matrix whose first N diagonal elements are positive while the rest are non-negative. The term $\Gamma \ \Psi \ (\hat{k}(t))$ is introduced to prevent the $\hat{k}_i(t)$ from becoming negative. In particular the i-th element of $\Psi(\hat{k}(t))$ is given by

$$\Psi_i(\hat{k}(t)) = \begin{cases} \hat{k}_i - M_i & \text{when } \hat{k}_i(t) > M_i \\ 0 & \text{when } m_i \leqslant k_i(t) \leqslant M_i \\ \hat{k}_i - m_i & \text{when } m_i \leqslant \hat{k}_i(t) \end{cases} \qquad (5.6a)$$

To understand the choice of the diagnonal weighting matrix Γ_2 assume, for the moment, that $\|K_u(t)\|_2$ is bounded by a known number M. Later, we shall show, through theorem 5.1 how, M can be determined irrespective of the choice of Γ. Select a set of numbers $\tilde{M}_i$ such that for each $i\varepsilon S$, $M_i \leqslant \tilde{M}_i$. Our choice of Γ will be such that

$$0 \leqslant \hat{k}_i(t) \leqslant \tilde{M}_i \qquad \forall t\varepsilon R_+,\ i\varepsilon S \qquad (5.6b)$$

Due to the initial conditions in (5.6), the inequalities in (5.6b) are satisfied at t = 0. Consider the case when at any instant for some $j\varepsilon S$ $\hat{k}_j(t) = 0$, with all other $\hat{k}_i(t)$ satisfying the condition (5.6b)

$$\dot{\hat{k}}_j(t) = + \sum_{\substack{r\subset S \\ j\varepsilon r}} \lambda_r \Big(\prod_{\substack{i\varepsilon r \\ i\neq j}} \hat{k}_i(t)\Big) K_{ur}(t) + \gamma_j m_j$$

where $\Gamma \triangleq \text{diag}\{\gamma_1,\ldots,\gamma_N\}$.

Since all $\hat{k}_i(t)$ are less than $\tilde{M}_i$ and $\|K_u(t)\| < M$, we have that

$$\dot{\hat{k}}_j(t) \geqslant \gamma_j m_j - M \sum_{\substack{r\subset S \\ j\varepsilon r}} \lambda_r \Big(\prod_{\substack{i\varepsilon r \\ i\neq j}} \tilde{M}_i\Big)$$

Thus if

$$\gamma_j > \frac{M}{m_j} \sum_{\substack{r \subset S \\ j \varepsilon r}} \lambda_r \left(\prod_{\substack{i \varepsilon r \\ i \neq j}} \tilde{M}_i \right)$$

Then $\dot{\hat{k}}_j(t) > 0$ whenever $\hat{k}_j(t) = 0$. Similarly if any $\hat{k}_j$ equals $\tilde{M}_j$ then

$$\dot{\hat{k}}_j(t) = \sum_{\substack{r \subset S \\ j \varepsilon r}} \lambda_r \left(\prod_{\substack{i \varepsilon r \\ i \neq j}} \hat{k}_i(t) \right) \left\{ \prod_{i \varepsilon r} \hat{k}_i(t)) - K_{ur}(t) \right\} - \gamma_j M_j$$

$$\geqslant \sum_{\substack{r \subset S \\ j \varepsilon r}} \lambda_r \left(\prod_{\substack{i \varepsilon r \\ i \neq j}} \tilde{M}_i \right) \left\{ \left(\prod_{i \varepsilon r} \tilde{M}_i \right) + M \right\} - \gamma_j M_j$$

Thus if

$$\gamma_j > \frac{1}{M_j} \sum_{\substack{r \subset S \\ j \varepsilon r}} \lambda_r \left(\prod_{\substack{i \varepsilon r \\ i \neq j}} \tilde{M}_i \right) \left\{ \left(\prod_{i \varepsilon r} \tilde{M}_i \right) + M \right\}$$

then $\hat{k}_j(t) < 0$ whenever $\hat{k}_j(t) = \tilde{M}_j$. Thus choosing γ_j as the

$$\max \left\{ \frac{M}{m_j} \sum_{r \subset S} \lambda_r \left(\prod_{\substack{i \varepsilon r \\ i \neq j}} \tilde{M}_i \right), \frac{1}{M_j} \sum_{\substack{r \subset S \\ j \varepsilon r}} \lambda_r \left(\prod_{\substack{i \varepsilon r \\ i \neq j}} \tilde{M}_i \right) \left\{ \left(\prod_{i \varepsilon r} \tilde{M}_i \right) + M \right\} \right\} \quad (5.6c)$$

will ensure that (5.6b) is satisfied

We now prove the uniform asymptotic convergence of $\hat{k}$ to k in two steps. Theorem 5.1 states that even without persistence of excitation, $\|\Delta K_u\| = \|K_u - K\|$ is bounded and the output error is in $\mathcal{L}^2$. We then show via theorem 5.2 that under persistence of excitation, $\hat{k}$ converges to k, uniformly asymptotically.

Theorem 5.1

Consider the unknown system (5.1), adjustable system (5.2) and the adjustment law (5.5). Suppose (5.1) is asymptotically stable $u(t) \in \Omega_\Delta[0,\infty]$ and $Q(s,k)/\beta(s)$ is S.P.R.. Then

(i) $K_u(t)$ is bounded $\quad \forall t \geq 0 \qquad$ (5.7)

(ii) $$\int_0^\infty \Delta y^2(t)\, dt < \infty \qquad (5.8)$$

(iii) $$\int_0^\infty \|\hat{V}(t) - V(t)\|_2^2\, dt < \infty \qquad (5.9)$$

$$V(t) \triangleq \left[\frac{p_1(s)u(t) - q_1(s)y(t)}{\beta(s)},\ \frac{p_2(s)u(t) - q_2(s)y(t)}{\beta(s)},\ \ldots\right]^T$$

Proof

Note first of all that equation (5.3) needs adjustment:

$$\Delta y(t) = \frac{\beta(s)}{Q(s,k)} \left\{\hat{V}^T(t)\ \Delta K_u(t) + \varepsilon_1(t)\right\} \qquad (5.10)$$

Where $\varepsilon_1(t)$ arises due to initial condition effects and decays exponentially to zero due to the stability of $Q(s,k)$.

However, its exponentially decaying nature implies that $\varepsilon_1(t)$ can be ignored.

Since $\beta(s)/Q(s,k)$ is SPR $x(t)$, A, b, c, d, L and σ such that [6]

$$\dot{x}(t) = Ax(t) + b\,(\hat{V}^T(t)\Delta K_u(t))$$

$$\Delta y(t) = c^T x + d\,(\hat{V}^T(t)\Delta K_u(t)) \qquad (5.11)$$

where

$$A+A^T = -LL^T - 2\sigma I$$

$$b = c - \sqrt{2d}\,L \qquad (5.12)$$

and {A,b,c,d} is time invariant.

Thus with $z(t) = [x^T(t), K_u^T(t)]^T$ we have that

$$\dot{z}(t) = \begin{bmatrix} A & b\hat{V}^T(t) \\ -\hat{V}(t)c^T & -d\hat{V}(t)\hat{V}^T(t) \end{bmatrix} z(t) \qquad (5.13)$$

Selecting the Lyapunov function

$$L_1(t) = z^T(t)z(t)$$

we observe that

$$\dot{L}_1(t) = z^T(t) \begin{bmatrix} A + A^T & b\hat{V}^T(t) - c\hat{V}^T(t) \\ \hat{V}(t)b^T - \hat{V}(t)c^T & -2d\hat{V}(t)\hat{V}^T(t) \end{bmatrix} z(t)$$

$$= -z^T(t) \begin{bmatrix} LL^T+2\sigma I & \sqrt{2d}\,L\hat{V}^T(t) \\ \sqrt{2d}\,\hat{V}(t)L^T & 2d\hat{V}(t)\hat{V}^T(t) \end{bmatrix} z(t)$$

$$= - [x^T(t), K_u^T(t)] \begin{bmatrix} \sqrt{2\sigma}\ I & L \\ 0 & \sqrt{2d}\ \hat{V}(t) \end{bmatrix} \begin{bmatrix} \sqrt{2\sigma}\ I & 0 \\ L^T & \sqrt{2d}\ \hat{V}^T(t) \end{bmatrix} \begin{bmatrix} x(t) \\ \Delta K_u(t) \end{bmatrix}$$

$$= - 2\sigma x^T(t)x(t) - (x^T(t)L + \sqrt{2d}\ \Delta K_u^T(t)\hat{V}(t))^2 \tag{5.14}$$

$$\leqslant 0$$

Thus

$$z^T(t)z(t) \leqslant z^T(0)z(0)$$

$$\Delta K_u^T(t)\Delta K_u(t) \leqslant \Delta K_u^T(0)\Delta K_u(0) + x^T(0)x(0) \tag{5.15}$$

whence (i) follows. Then $\exists$ $\overline{M}_1$ and $\overline{M}_2$ such that

$$\int_0^\infty x^T(t)x(t)dt < \overline{M}_1 \tag{5.16}$$

and

$$\int_0^\infty (x^T(t)L + \sqrt{2d}\ \Delta K_u^T(t)\hat{V}(t))^2 dt < \overline{M}_2 \tag{5.17}$$

Thus

$$\int_0^\infty 2d(\Delta K_u^T(t)\hat{V}(t))^2 dt < \overline{M}_2 + \int_0^\infty \| x(t) \|_2^2\ dt\ \| L \|_2^2$$

$$< \overline{M}_2 + \overline{M}_1\ \| L \|_2^2 \tag{5.18}$$

Thus

$$\int_0^\infty (\Delta K_u^T(t)\hat{V}(t))^2 dt < \overline{M}_3$$

From (5.11)

$$\int_0^\infty \Delta y^2(t)\,dt \leq \|c\|_2^2 \int_0^\infty \|x\|_2^2 dt + d^2 \int_0^\infty |\hat{V}^T \Delta K_u|^2 dt$$

$$\leq \bar{M}_4 < \infty \qquad (5.19)$$

Thus (5.8) is proved. Equation (5.9) is proved by noting that $\hat{V}(t) - V(t) = G(s)\Delta y(t)$

where G(s) is an asymptotically stable, proper transfer function.

∇∇∇

Remark:

(5.3) From (5.15) we see that

$$\{K_u(t) - K\}^T \{K_u(t) - K\} \leq \Delta K_u^T(0)\Delta K_u(0) + x^T(0)x^T(0)$$

(5.19a)

Suppose, a bound on the magnitude of the initial state vector in any minimal realization of (5.1) is known. Then an *a priori* bound, $\bar{M}_5$, on the magnitude of x(0) will also be available. Now if the initial conditions in (5.5) are satisfied then

$$\|K_u(t)\|^2 \leq \|K\|^2 + \sum_{r \subset S} \{\prod_{i \varepsilon r} M_i - \prod_{i \varepsilon r} m_i\}^2 + \bar{M}_5$$

$$\leq \sum_{r \subset S} \{(\prod_{i \varepsilon r} M_i)^2 + \{\prod_{i \varepsilon r} M_i - \prod_{i \varepsilon r} m_i\}^2\} + \bar{M}_5$$

(5.19b)

Thus the M in (5.6c) should equal the square root of the right hand side of (5.19b). Thus with Γ and M chosen as in (5.6c) and (5.19b) respectively $\hat{k}_i(t)$ is always non-negative.

(5.4) Equation (5.8) does not ensure that $\lim_{t\to\infty} \Delta y(t) = 0$ as $\hat{y}(t)$ may not be bounded.

We now prove our main result using a key Lemma given in the Appendix.

<u>Theorem 5.2</u>

For the equation (5.1)(5.2), (5.5-6) let $K_{ur} \geqslant 0$ $\forall r \subset S$. Assume that (5.1) is asymptotically stable, assumption 5.2 holds and there exist no θ, such that $\|\theta\| = 1$ and

$$\sum_{r\subset S} \theta_r p_r(s) \equiv \sum_{r\subset S} \theta_r q_r(s) \equiv 0, \qquad (5.20)$$

Then $[x(t)^T, K_u^T(t)-K^T, \hat{k}(t)^T-k^T]$ converges uniform asymptotically to zero if $\alpha_1, \alpha_2, \delta > 0$ such that $\forall \sigma \varepsilon R_+$

$$\alpha_1 I \leqslant \int_{\sigma}^{\sigma+\delta} W(t)W^T(t)\,dt \leqslant \alpha_2 I \qquad (5.21)$$

Here

$$W(t)^T \triangleq [u(t), \frac{1}{s+\gamma} u(t), \ldots, \frac{1}{(s+\gamma)^m} u(t)]$$

for any $\gamma < 0$ and m the highest degree among the polynomials $p_o q_r - q_o p_r$ $\forall r \subset S$ and $p_r q_{\bar{r}} - q_r p_{\bar{r}}$ $\forall r, \bar{r} \subset S$ and $r \neq \bar{r}$. Also $u(t) \varepsilon \Omega_\Delta [0,\infty)$.

Proof

From theorem 4.2 we know that (5.21) and the asymptotic stability of (5.1) implies $\exists\ \beta_3, \beta_4, \delta_2 > 0$ such that $\forall \sigma \in R_+$

$$\beta_3 I \leqslant \int_{\sigma}^{\sigma+\delta_2} V(t)\, V^T(t) dt \leqslant \beta_4 I \tag{5.22}$$

where V has the elements $\dfrac{p_r(s)}{\beta(s)} u(t) - \dfrac{q_r(s)}{\beta(s)} y(t)\ \forall r \subset S.$

Then by (5.9) and Lemma 5.A.1 $\exists\ \beta_5, \beta_6, \delta_3 > 0$ such that $\forall \sigma \in R_+$

$$\beta_5 I \leqslant \int_{\sigma}^{\sigma+\delta_3} \hat{V}(t)\hat{V}^T(t)\, dt \tag{5.23}$$

and

$$\int_{\sigma}^{\sigma+\delta_3} \| \hat{V}(t) \|_2^2\, dt \leqslant \beta_6 \tag{5.24}$$

From (5.13)

$$z(t) = \begin{bmatrix} A & b\hat{V}(t)^T \\ -\hat{V}(t)c^T & -d\hat{V}(t)\hat{V}(t)^T \end{bmatrix} z(t) \tag{5.25}$$

with $z(t) \triangleq [x(t)^T, \Delta K_u^T(t)]^T$, $x(t)$ defined in (5.11) and $d+c^T(sI-A)^{-1}b = \beta(s)/Q(s,k)$. Then by a result in [6] (5.23-24) imply that (5.25) is exponentially, asymptotically stable if $\exists\ \beta_6 > 0, \delta_4 > 0$ such that $\forall \delta \in R_+$

$$\beta_6 I \leqslant \int_{\sigma}^{\sigma+\delta_4} \left[2d\ \hat{V}(t)\hat{V}(t)^T + b^T b \int_{\sigma}^{t} \hat{V}(t) dt \int_{\sigma}^{t} \hat{V}^T(t) dt \right] dt$$

(5.26)

Clearly as $\delta[\beta(s)] = \delta[Q(s,k)] = n$, $d > 0$. Thus (5.22) implies (5.26). Thus (5.25) is exponentially stable. Thus by a result in [4,p86] there exists a Lyapunov function $\tilde{L}_1$ such that for some $\mu_1, \mu_2, \mu_3 > 0$

$$\mu_1 \|\Delta K_u(t)\|^2 + \mu_1 \|x(t)\|^2 \leqslant \tilde{L}_1 \leqslant \mu_2 \|K_u(t)-K\|^2 + \mu_2 \|x(t)\|^2$$

and

$$\dot{\tilde{L}}_1(t) \leqslant -\mu_3 \|K_u(t)-K\|^2 - \mu_3 \|x(t)\|^2.$$

Consider also the Lyapunov function

$$\tilde{L}_2(t) = \frac{1}{2} (\hat{K}(t)-K)^T \Lambda (\hat{K}(t)-K) + \frac{1}{2} \Psi^T(\hat{k}(t)) \Gamma \Psi (\hat{k}(t))$$

Then

$$\frac{\partial \tilde{L}_2(t)}{\partial \hat{k}(t)} = \frac{\partial \hat{K}^T(t)}{\partial \hat{k}(t)} \Lambda (\hat{K}(t)-K) + \Gamma \Psi (\hat{k}(t))$$

By the definition of $\Psi(\hat{k}(t))$ it follows that since (5.6b) is true

$$\ln \left(\frac{\hat{k}_i(t)}{k_i(t)}\right) \Psi_i(\hat{k}(t)) \geqslant 0$$

Thus as in the proof of Lemma 4.2 we find

$$\frac{\partial \tilde{L}_2(t)}{\partial \hat{k}(t)} = 0 \qquad \text{iff } \hat{k}(t) = k.$$

Thus a simple modification of theorem 4.7 shows that the result is true

∇∇∇

Remark:

(5.5) Consider a non-minimal representation of (5.2)

$$\begin{aligned} \dot{x}_1(t) &= A_1(K_u(t))\, x_1(t) + b_1(K_u(t))\, u(t) \\ \hat{y}(t) &= c_1(K_u(t))\, x_1(t) + d_1(K_u(t))\, u(t) \end{aligned} \qquad (5.27)$$

Then as $K_u(t) \to K$ exponentially and $A(K_u(t))$ is asymptotically stable, a result in [5] shows that

$$\dot{x}_1 = A_1(K_u(t))\, x_1$$

is exponentially stable. Thus $\hat{y}$ is bounded. Also by arguing as in remark (4.12) one can show that $\hat{k}(t)-k$ decays exponentially to zero.

5.2 A recursive least squares formulation

Consider the unknown system and the adjustable model defined by (5.1) and (5.2) respectively and the update scheme

$$\dot{\hat{R}}(t) = -\alpha\hat{R}(t) + \hat{V}(t)\hat{V}^T(t) , \quad \forall t \geq 0 ; \quad \hat{R}(0) = 0 \qquad (5.28)$$

$$\dot{X}(t) = \alpha X(t) - X(t)\hat{V}(t)\hat{V}^T(t)X(t), \quad X(t_0) = \hat{R}^{-1}(t_0) \quad \forall t \geq t_0 \qquad (5.29)$$

$$\dot{K}_u = -X(t)\hat{V}(t)\Delta y(t) \quad \forall t \geq t_0 \qquad (5.30)$$

$$\dot{\hat{k}}(t) = -\frac{\partial\hat{K}^T(t)}{\partial\hat{k}(t)}\Lambda[\hat{K}(t) - K_u(t)] - \Gamma\Psi(\hat{k}(t)) , \quad \forall t \geq t_0 \qquad (5.31)$$

$$m_i \leq \hat{k}_i(t_0) \leq M_i \quad \forall i \in S$$

where t_0 is the first time instant at which $\hat{R}(t)$ becomes well conditioned, and Ψ and Γ are defined in (5.6a,c). The choice of the bound M on $K_u(t)$ will be explained at a later stage, but assume for the time being that M is such that (5.6b) is always satisfied.

One should note that (5.30) is an unconstrained least squares output error algorithm [1].

The following result shows that the infinite memory associated with (5.28) and (5.29) ensures that X(t) is the inverse of $\hat{R}(t) \forall t \geq t_0$. Its proof is omitted as it is trivial.

Lemma 5.2

Suppose $\hat{R}(t)$ and X(t) are as defined in (5.28) and (5.29) and that $\exists$ $t_0 > 0$ such that $\hat{R}^{-1}(t_0)$ exists. Then

$$\hat{R}(t)X(t) = X(t)\hat{R}(t) = I \qquad \forall t \geq t_0 \qquad (5.32)$$

$\nabla \nabla \nabla$

The convergence analysis proceeds on similar lines to the previous section. The SPR condition, however, needs adjustment.

We require here that $\beta(s)/Q(s,k) - \frac{1}{2}$ be SPR. Then theorem 5.3 shows that the output error is in L^2 as long as $\hat{R}(t_0)$ is invertible for some $t_0 > 0$.

Theorem 5.3

For the unknown system (5.1), adjustable system (5.2) and adaptive law (5.30), the following are true as long as $\beta(s)/Q(s,k) - \frac{1}{2}$ is SPR, $u(t) \in \Omega_\Delta[0,\infty)$ and $\exists t_0 > 0$ such that

$$\hat{R}(t_0) \geq \alpha_1 I \ .$$

(i) $\Delta K_u^T(t)\hat{R}(t)\Delta K_u(t)$ is bounded (5.33)

(ii) $\int_0^\infty \Delta y^2(t)dt < \infty$ (5.34)

(iii) $\int_0^\infty \|\hat{V}(t) - V(t)\|^2 dt < \infty$ (5.35)

(iv) $\hat{R}(t) \leq \alpha_2 I$ for some finite α_2 and all $t \geq 0$. (5.36)

Proof:

By lemma 5.1 and the SPR nature of $\beta(s)/Q(s,k) - \frac{1}{2}$ we have that $\exists\ x(t), A, b, c, d, L, \sigma$ such that

$$\left.\begin{aligned} \dot{x}(t) &= Ax(t) + b(\hat{V}(t)^T\Delta K_u(t)) \\ \Delta y(t) &= c^T x(t) + d(\hat{V}^T(t)\Delta K_u(t)) \end{aligned}\right\} \quad (5.37)$$

$$\left.\begin{aligned} A + A^T &= -LL^T - 2\sigma I \\ b &= c - \sqrt{2d-1}\ L \end{aligned}\right\} \quad (5.38)$$

Thus with

$$z^T = [x^T(t), \Delta K_u^T(t)]$$

we find

$$\dot{z}(t) = \begin{bmatrix} A & b\hat{V}^T(t) \\ -X(t)\hat{V}(t)c^T & -dX(t)\hat{V}(t)\hat{V}^T(t) \end{bmatrix} z(t) \tag{5.39}$$

Choose a Lyapunov like function

$$L_2 = z^T(t)\begin{bmatrix} I & 0 \\ 0 & \hat{R}(t) \end{bmatrix} z(t)$$

Then

$$\dot{L}_2 = -z^T(t)\begin{bmatrix} A+A^T & (b-c)\hat{V}^T(t) \\ \hat{V}(t)(b-c)^T & -2d\hat{V}(t)\hat{V}^T(t) \end{bmatrix} z(t) + z^T(t)\begin{bmatrix} 0 & 0 \\ 0 & \dot{\hat{R}}(t) \end{bmatrix} z(t)$$

$$= z^T(t)\begin{bmatrix} -2\sigma I - LL^T & -\sqrt{2d-1}\, L\hat{V}^T(t) \\ -\sqrt{2d-1}\,\hat{V}(t)L^T & -2d\hat{V}(t)\hat{V}^T(t) \end{bmatrix} z(t) + z^T(t)\begin{bmatrix} 0 & 0 \\ 0 & -\alpha\hat{R}(t) \end{bmatrix} z(t)$$

$$+ z^T(t)\begin{bmatrix} 0 & 0 \\ 0 & \hat{V}(t)\hat{V}^T(t) \end{bmatrix} z(t) \tag{5.40}$$

$$\leq z^T(t)\begin{bmatrix} -2\sigma I - LL^T & -\sqrt{2d-1}\, L\hat{V}^T(t) \\ -\sqrt{2d-1}\,\hat{V}(t)L^T & -(2d-1)\hat{V}(t)\hat{V}^T(t) \end{bmatrix} z(t)$$

$$= -z^T(t)\begin{bmatrix} \sqrt{2\sigma}\, I & L \\ 0 & \sqrt{2d-1}\,\hat{V}(t) \end{bmatrix} \begin{bmatrix} \sqrt{2\sigma}\, I & 0 \\ L^T & \sqrt{2d-1}\,\hat{V}^T(t) \end{bmatrix} z(t)$$

$$\leq 0 .$$

Thus

$$\Delta K_u^T(t)\hat{R}(t)\Delta K_u(t) \leq \Delta K_u^T(t_0)\hat{R}(t_0)\Delta K_u(t_0) + x^T(t_0)x^T(t_0) \tag{5.40}$$

Moreover, equations (5.34 and 35) follow from arguments in theorem 5.1. Let $\bar{M}_1$ be such that

$$\int_0^\infty \|\hat{V}(t) - V(t)\|^2 dt < \bar{M}_1$$

As (5.1) is a.s. and $u(t) \in \Omega_\Delta[0,\infty)$

$$\int_0^t e^{-\alpha(t-\tau)} \|V(\tau)\|^2 d\tau \le \bar{M}_2 < \infty \qquad \forall t \in R_+$$

Thus

$$\left[\int_0^t e^{-\alpha(t-\tau)} \|\hat{V}(\tau)\|^2 d\tau\right]^{\frac{1}{2}} \le \left[\int_0^t e^{-\alpha(t-\tau)} \|\hat{V}(\tau) - V(\tau)\|^2 d\tau\right]^{\frac{1}{2}}$$

$$+ \left[\int_0^t e^{-\alpha(t-\tau)} \|V(\tau)\|^2 d\tau\right]^{\frac{1}{2}}$$

$$\le \sqrt{\bar{M}_1} + \sqrt{\bar{M}_2}$$

whence for any unit vector $\theta \in R^{2^N-1}$,

$$\theta^T \hat{R}(t)\theta = \int_0^t e^{-\alpha(t-\tau)} \left(\theta^T \hat{V}(\tau)\right)^2 d\tau$$

$$\le \int_0^t e^{-\alpha(t-\tau)} \|\hat{V}\|^2 d\tau$$

$$\le \left(\sqrt{\bar{M}_1} + \sqrt{\bar{M}_2}\right)^2 .$$

Thus (5.36) is proved.

∇∇∇

Remark:

(5.6). The condition under which the nonsingularity of $R(t_0)$ at some time t_0, can be guaranteed is identical to the input conditions given in Theorem 5.2. This is because, up to $t=t_0$, the adjustable system is constant. Thus by considerations similar to theorems 3.7 and 4.2, it can be shown that $\exists\ \alpha_3,\ \delta>0$ such that $\forall\sigma \le t_0-\delta$

$$\alpha_3 I \le \int_{\sigma}^{\sigma+\delta} \hat{V}(t)\hat{V}^T(t)dt \tag{5.41}$$

Moreover, as the following extension of theorem 4.1 shows, this is enough to ensure the nonsingularity of $R(t_0)$. It should be noted $\hat{V}(t)$ need not be bounded, whereas Theorem 4.1 did require it to be bounded.

Theorem 5.4

Suppose $\exists\ \alpha_3, t_1>0$ such that $\forall\sigma$ and some $T<t_1$

$$\int_{g(\sigma-T)}^{\sigma} \hat{V}(\tau)\hat{V}^T(\tau)d\tau \ge \alpha_3 I \tag{5.42}$$

where

$$g(\tau) = \begin{cases} \tau & \tau>0 \\ 0 & \text{elsewhere} \end{cases}$$

Then $\exists\ \alpha_2>0$ such that $\forall\sigma>\sigma_0$ and some σ_0

$$R(\sigma) = \int_0^{\sigma} e^{-\alpha(\sigma-\tau)}\ \hat{V}(\tau)\hat{V}^T(\tau)d\tau \ge \alpha_2 I \tag{5.43}$$

Proof

Suppose (5.43) is violated. Then for arbitrary $\varepsilon>0$ $\exists$ a unit θ such that for some σ

$$\theta^T R(\sigma)\theta \le \varepsilon^2$$

$\Rightarrow$

$$\int_0^{\sigma} e^{-\alpha(\sigma-t)}\{\theta^T\hat{V}(t)\}^2 dt < \varepsilon.$$

Thus

$$\int_{g(\sigma-t_1)}^{\sigma} e^{-\alpha(\sigma-t)}\{\theta^T \hat{V}(t)\}^2 dt < \varepsilon$$

whence by the definition of $g(\sigma-t_1)$

$$\int_{g(\sigma-t_1)}^{\sigma} \{\theta^T H(t)\}^2 dt < \varepsilon e^{\alpha t_1}$$

$\Rightarrow$ (5.42) is violated. Thus (5.42) $\Rightarrow$(5.43).

∇∇∇

Theorem 5.5, below, shows that $\Delta K_u \to 0$ exponentially fast whenever u(t) is p.e. This in turn implies that $\hat{k}$ converges uniformly asymptotically to k as long as the K_{ur}'s are all positive. In proving this theorem we need to appeal to the notion of uniform complete observability (u.c.o.) used in [6] and defined below.

<u>Definition 5.1:</u>

Let $F(.)$: $R_+ \to R^{n\times n}$ and $H(\cdot)$: $R_+ \to R^{n\times r}$ be regulated (i.e. one-sided limits exist for all $t\in R_+$). Let $\Phi(\cdot,\cdot)$ be the transition matrix associated with $F(\cdot)$. Then [F,H] is uniformly completely observable if the following three conditions hold (any two implying the third [7]) for some positive $\beta_1 - \beta_4$ and δ and $\forall s,t\in R_+$

$$\beta_1 I \le N(s,s+\delta) \le \beta_2 I$$

$$\beta_3 I \le \Phi^T(s,s+\delta)N(s,s+\delta)\Phi(s,s+\delta) \le \beta_4 I$$

$$\|\Phi(t,s)\| \le \alpha_5(|t-s|)$$

where

$$N(s,s+\delta) = \int_s^{s+\delta} \Phi^T(t,s)H(t)H^T(t)\Phi(t,s)dt$$

and $\alpha_5(\cdot)$ is bounded for bounded arguments.

Theorem 5.5

For the equations (5.1 - 2) and (5.29 - 31) suppose assumption 4.1 holds, (5.1) is asymptotically stable and $\beta(s)/Q(s,k) - \frac{1}{2}$ is SPR. Then $\Delta K_u(t)$ converges exponentially to zero if $\exists\, \alpha_1, \alpha_2, \delta > 0$ such that

$$\alpha_1 I \le \int_\sigma^{\sigma+\delta} W(t)W^T(t)dt \le \alpha_2 I \tag{5.44}$$

$\forall \sigma \epsilon R_+$. Here $W(t) \triangleq [u(t), \frac{1}{s+\gamma} u(t), \ldots, \frac{1}{(s+\gamma)^m} u(t)]^T$

for any positive γ and m = highest degree among the polynomials $p_o q_r - q_o p_r$ $\forall r \subset S$ and $p_r q_{\bar{r}} - q_r p_{\bar{r}}$ $\forall r, \bar{r} \subset S, r \neq \bar{r}$, $u(t) \in \Omega_\Delta[0,\infty)$. Moreover, if $K_{ur} > 0$ $\forall r \subset S$ then $[\hat{k}^T(t) - k^T, x^T(t), K_u^T(t) - K^T]$ converges uniformly asymptotically to zero.

Proof

As in theorem 5.2. Assumption 4.1, (5.44), (5.35) and Lemma 5.A.1 imply the existence of $\bar{\delta}, \alpha_3, \alpha_4 > 0$ such that

$$\int_\sigma^{\sigma+\bar{\delta}} \|\hat{V}(t)\|^2 dt < \alpha_4 \tag{5.45}$$

and

$$\int_{\sigma}^{\sigma+\bar{\delta}} \hat{V}(t)\hat{V}^T(t)dt > \alpha_3 I \qquad (5.46)$$

$\forall \sigma \in R_+$.

Thus by theorem 5.4 $\exists\ \alpha_5 > 0$ such that

$$\hat{R}(t) \geq \alpha_5 I \qquad \forall t > t_1, \quad \text{some } t_1 > 0 . \qquad (5.47)$$

Thus L_2 given by

$$L_2 = z^T(t)\begin{bmatrix} I & 0 \\ 0 & \hat{R}(t) \end{bmatrix} z(t)$$

with $z(t)$ defined by (5.39), is a Lyapunov function.

Suppose

$$F(t) = \begin{bmatrix} A & b\hat{V}^T(t) \\ -X(t)\hat{V}(t)c^T & -dX(t)\hat{V}(t)\hat{V}^T(t) \end{bmatrix}$$

$$H(t) = \begin{bmatrix} \sqrt{2\sigma}\ I & L \\ 0 & \sqrt{2d-1}\ \hat{V}(t) \end{bmatrix}$$

Then let $\bar{H}(t)$ be defined by

$$\begin{bmatrix} 0 & 0 \\ 0 & \alpha\hat{R}(t) - \hat{V}(t)\hat{V}^T(t) \end{bmatrix} = F^T(t)\begin{bmatrix} I & 0 \\ 0 & \hat{R}(t) \end{bmatrix} + \begin{bmatrix} I & 0 \\ 0 & \hat{R}(t) \end{bmatrix} F(t) + \bar{H}(t)\bar{H}^T(t$$

$$= -H(t)H^T(t) - \begin{bmatrix} 0 & 0 \\ 0 & \hat{V}(t)\hat{V}^T(t) \end{bmatrix} + \bar{H}(t)\bar{H}^T(t) \qquad \text{(by 5.38)}$$

$$\Longleftrightarrow \bar{H}(t)\bar{H}^T(t) = \begin{bmatrix} 0 & 0 \\ 0 & \alpha\hat{R}(t) \end{bmatrix} + H(t)H^T(t)$$

$$= \bar{R}(t) + H(t)H^T(t) \qquad (5.47a)$$

Thus by [6], (5.39) is exponentially asymptotically stable (e.a.s.) if $[F,\bar{H}]$ is u.c.o. We now show that $[F,H]$u.c.o. $\Rightarrow$ $[F,\bar{H}]$u.c.o. With N defined as in definition 5.1 and $\bar{N}$ similarly defined, we have to show that if $\exists$ $\beta_1 - \beta_4$ and δ' such that $\forall$ $\sigma,t \in R_+$

$$\beta_1 I \leq N(\sigma,\sigma+\delta') \leq \beta_2 I \tag{5.48}$$

and

$$\beta_3 I \leq \Phi^T(\sigma,\sigma+\delta')N(\sigma,\sigma+\delta')\Phi(\sigma,\sigma+\delta') \leq \beta_4 I \tag{5.49}$$

then $\exists$ $\bar{\beta}_1,\ldots,\bar{\beta}_4$ and δ'' such that $\forall \sigma,t \in R_+$

$$\bar{\beta}_1 I \leq \bar{N}(\sigma,\sigma+\delta'') \leq \bar{\beta}_2 I \tag{5.50}$$

and

$$\bar{\beta}_3 I \leq \Phi^T(\sigma,\sigma+\delta'')\ \bar{N}(\sigma,\sigma+\delta'')\Phi(\sigma,\sigma+\delta'') \leq \bar{\beta}_4 I \tag{5.51}$$

We shall show that (5.48) implies (5.50), the proof of the other implication being similar. Let the lower bound of (5.50) be violated. Then for arbitrary $\varepsilon>0$ $\exists$ a unit θ and a $\sigma \in R_+$ such that

$$\int_\sigma^{\sigma+\delta} \theta^T\ \Phi^T(t,\sigma)\ \bar{H}(t)\bar{H}^T(t)\Phi(t,\sigma)\theta dt < \varepsilon$$

$$\Rightarrow \int_\sigma^{\sigma+\delta} \theta^T\ \Phi^T(t,\sigma) \begin{bmatrix} 0 & 0 \\ 0 & \alpha\hat{R}(t) \end{bmatrix} \Phi(t,\sigma)\ \theta dt$$

$$+ \int_\sigma^{\sigma+\delta} \theta^T\ \Phi^T(t,\sigma)\ H(t)H^T(t)\Phi(t,\sigma)\ \theta dt < \varepsilon\ .$$

As $\hat{R}$ is positive definite

$$\theta^T \int_{\sigma}^{\sigma+\delta} \Phi^T(t,\sigma)\ H(t)H^T(t)\Phi(t,\sigma)\ \theta\, dt < \varepsilon$$

whence lower bound of (5.48) is violated. Thus lower bound of (5.48) implies the lower bound of (5.50). Also with $\bar{R}$ defined as in (5.47a)

$$\left[\int_{\sigma}^{\sigma+\delta} \theta^T\Phi^T(t,\sigma)\ \bar{H}(t)\bar{H}^T(t)\Phi(t,\sigma)\ \theta\, dt\right]^{\frac{1}{2}}$$

$$\leq \left[\int_{\sigma}^{\sigma+\delta} \theta^T\ \Phi^T(t,\sigma)\ H(t)H^T(t)\Phi(t,\sigma)\ \theta\, dt\right]^{\frac{1}{2}}$$

$$+ \left[\int_{\sigma}^{\sigma+\delta} \theta^T\ \Phi^T(t,\sigma)\ \bar{R}(t)\Phi(t,\sigma)\ \theta\, dt\right]^{\frac{1}{2}}$$

whence by (5.33) the upperbound of (5.50) follows. Thus z is e.a.s. if [F,H] is u.c.o. Using arguments similar to [6] it can be shown that (5.45) and (5.46) imply that [F,H] is u.c.o. Thus z is e.a.s. and by arguments similar to theorem 5.2 we have that $[\hat{k}^T - k^T, K_u^T - K^T x^T]$ converges u.a. to zero.

∇∇∇

Remarks

(5.7). Suppose $\exists\ \alpha_1, \alpha_2 > 0$ such that $\alpha_1 I \leq R(t)\ \forall t$ and $R(t_0) \leq \alpha_2 I$. Thus (5.40a) ⇒

$$\Delta K^T(t)\Delta K^T(t) \leq \frac{\alpha_2}{\alpha_1}\ \Delta K^T(t_0)\Delta K(t_0) + \frac{\bar{M}}{\alpha_1}$$

where $\bar{M}$ is the bound on $x^T(0)x(0)$. As $\alpha_1 > 0$ one can choose a sufficiently large u(t) to ensure that $\alpha_1 > 1$, i.e.

$$\Delta K^T(t)\Delta K(t) \leq \alpha_2\, \Delta K^T(t_0)\Delta K(t_0) + \bar{M}$$

Also

$$\alpha_2 \leq \frac{1}{\alpha} \max_{t\in[0,t_0]} \|\hat{V}(t)\|^2 \doteq \bar{M}_3\ .$$

Thus for high enough u(t)

$$\Delta K^T(t)\Delta K(t) \le \bar{M} + \bar{M}_3 .$$

It is reasonable to expect that conservative estimates of both $\max_{t\in[0,t_0]}\|\hat{V}(t)\|^2$ and $\bar{M}$ would be known. Thus this can be used to employ the shift technique explained earlier so that $K_{ur}(t) > 0 \;\; \forall t \ge 0$.

(5.8). As in remark (5.5) it is possible to show that $\Delta y(t)$ is bounded and that $\hat{k}(t)$ approaches k exponentially fast.

Appendix 5.A A Key Lemma

Lemma 5.A.1 is an extension of a result in [2] to unbounded signals. All norms considered are two norms.

Lemma 5.A.1

Consider $V:R_+ \to R^n$ and $\hat{V}:R_+ \to R^n$. Suppose $\exists\ M_1, \alpha_1, \alpha_2, \delta > 0$ such that $\forall \sigma \in R_+$

$$\int_0^\infty \|V - \hat{V}\|^2\, dt < M_1 \tag{5.A.1}$$

$$\alpha_1 I \le \int_\sigma^{\sigma+\delta} VV^T\, dt \tag{5.A.2}$$

and

$$\int_\sigma^{\sigma+\delta} \|V\|^2 dt < \alpha_2\ . \tag{5.A.3}$$

Then $\exists\ \alpha_3, \alpha_4, \bar{\delta} > 0$ such that $\forall \sigma \in R_+$

$$\alpha_3 I \le \int_\sigma^{\sigma+\bar{\delta}} \hat{V}\,\hat{V}^T dt \tag{5.A.4}$$

and

$$\int_\sigma^{\sigma+\delta} \|\hat{V}\|^2\, dt \le \alpha_4 \tag{5.A.5}$$

$\forall \sigma \in R_+$.

Moreover, if (5.A.3) does not hold, then (5.A.4) still holds as long as $\hat{V}$ and V do not have finite escape times.

Proof

(i) We first show that (5.A.1) and (5.A.3) imply (5.A.5). Equation (5.A.1) shows

$$\int_\sigma^{\sigma+\bar{\delta}} \|V - \hat{V}\|^2\, dt < M_1 \qquad \forall \sigma, \bar{\delta} \in R_+\ .$$

Thus (5.A.3) and the Minkowskii's inequality shows

$$\left\{\int_{\sigma}^{\sigma+\bar{\delta}} \|\hat{V}\|^2 dt\right\}^{\frac{1}{2}} \le \sqrt{M_1} + \left\{\int_{\sigma}^{\sigma+\bar{\delta}} \|V\|^2 dt\right\}^{\frac{1}{2}} \le \sqrt{M_1} + \sqrt{p\alpha_2 + \alpha_2}$$

where p = Integer part of $\bar{\delta}/\delta$. This proves (5.A.5).

(ii) Consider any arbitrary vector θ of unit magnitude. Then (5.A.2) ⇒

$$\int_{\sigma}^{\sigma+\delta} (\theta^T V)^2 dt \ge \alpha_1 \qquad \forall \sigma \in R_+$$

⇒

$$\int_{\sigma}^{\sigma+\delta p} (\theta^T V)^2 dt \ge \alpha_1 p \qquad \forall \sigma \in R_+ \text{ and } p \in Z_+ . \tag{5.A.6}$$

Suppose $\bar{\delta} = \delta p$. If $\hat{V}$ and V do not have finite escape times or since (5.A.5) holds, Minkowskii's inequality is applicable. Thus (5.A.1) and (5.A.2) ⇒

$$\left\{\int_{\sigma}^{\sigma+\bar{\delta}} (\theta^T \hat{V})^2 dt\right\}^{\frac{1}{2}} = \left\{\int_{\sigma}^{\sigma+\bar{\delta}} (\theta^T \hat{V} - \theta^T V + \theta^T V)^2 dt\right\}^{\frac{1}{2}}$$

$$\ge \left\{\int_{\sigma}^{\sigma+\bar{\delta}} (\theta^T V)^2 dt\right\}^{\frac{1}{2}} - \left\{\int_{\sigma}^{\sigma+\bar{\delta}} (\theta^T \hat{V} - \theta^T V)^2 dt\right\}^{\frac{1}{2}}$$

$$\ge \sqrt{\alpha_1 p} - \left\{\int_{\sigma}^{\sigma+\bar{\delta}} \|\hat{V} - V\|^2 dt\right\}^{\frac{1}{2}}$$

$$\ge \sqrt{\alpha_1 p} - \sqrt{M_1} .$$

Thus if $p > M_1/\alpha_1$ and $\bar{\delta} \ge p\delta$, (5.A.4) holds.

∇∇∇

References for Chapter 5

1. Y.D. Landau, Adaptive control: the model reference approach, Marcel Dekker, Inc., 1979.

2. S. Boyd and S. Sastry, "On parameter convergencein adaptive control", Systems & Control Letters, pp 311-319, Dec 1983.

3. S.Dasgupta, B.D.O. Anderson and R.J. Kaye, "Output error identification of physical element values in structured systems", Under preparation.

4. N.N. Krasovskii, Stability of motion, Stanford University Press, 1963.

5. G. Kreisselmeier, "An approach to stable indirect adaptive control", Submitted for publication.

6. B.D.O. Anderson, "Exponential stability of linear equations arising in adaptive identification", IEEE Trans Auto Contr, AC-22, Feb 1977, pp 83-88.

7. R.E. Kalman, "Contributions to the theory of optimal control", Bol. Soc. Mat. Mexicana, vol 5, 1960, no. 2 pp 102-119.

§6 Adaptive Control

This chapter considers the indirect adaptive control of partially known systems described in Chapter 2. The indirect approach involves the estimation of the plant parameters, based on which the controller parameters are adjusted. As opposed to this in direct adaptive control the controller parameters are directly evaluated.

A major attraction of the indirect approach is that the control and identification phases can be decoupled and to an extent analysed independent of each other. In [1] Egardt and Samson put forward an approach involving a specific control algorithm and a general identifier satisfying certain assumptions. Kreisselmeier [2] on the other hand considers a general controller which is coupled with a specific identifier. In this chapter the latter approach is adopted. Two algorithms are formulated, each having the same general contoller, but differing in the nature of the identifier. In the first a two step modified constrained least squares algorithm is used, while the second employs a two step gradient descent algorithm. Both parallel closely the corresponding identification algorithms presented in Chapter 4.

Although global stability of several direct adaptive controllers has been shown [3-6] , very few such results exist for indirect adaptive control. Kreisselmeier [2] , has proved global stability for his control algorithm under the assumption that the extent of a convex region containing the true parameter values is known. This region has the added property that when the plant parameter values lie in it the plant can be well controlled in the sense of being both stabilizable and detectable. For the first algorithm of this chapter global stability is proved by assuming that the plant is completely controllable and completely observable at the true parameter values only. For the second algorithm global stability is proved under the same assumptions as those made in [2] . As we have remarked earlier the assumption of known magnitude bounds for the unknown parameters is reasonable, given their direct physical significance. As in [2] , the convergence analysis for the second algorithm includes the derivation of a link between the identification error and the stability of the closed loop system.

Furthermore, under the assumption of a completely controllable and completely observable plant, persistence of excitation conditions on the reference inputs are presented. These conditions guarantee the global uniform asymptotic stability of the parameter estimates, a property which, as has been argued in Chapter 3 and [7,8] , ensures the robust behaviour of the adaptive controllers.

Section 6.1 and 6.2 respectively present and analyse the two algorithms proposed while Section 6.3 presents simulation results. The contents of this chapter will appear in [9] .

6.1 Adaptive control using modified least squares identifier :

Let the plant be described by the strictly proper transfer function

$$T_p(s) = \frac{p_o(s) + \sum_{r \subset S} (\prod_{i \in r} k_i) p_r(s)}{q_o(s) + \sum_{r \subset S} (\prod_{i \in r} k_i) q_r(s)} ; \; S \triangleq \{1,2,\ldots N\} \tag{6.1}$$

$$= \frac{P(s,k)}{Q(s,k)}$$

with $\delta[Q(s,k)] = \delta[q_o(s)] = n > \delta]q_r(s)] \; \forall r \subset S \; , \; k \in R^N$

the unknown parameter vector and the $p_i(s)$, $q_i(s)$, known a priori. Suppose K is a vector containing the multilinear combinations of the k_i and K_u an unconstrained estimate of K. Assume $P(s,k)$ and $Q(s,k)$ are coprime in the sense that they have no non-trivial common factors which are polynomials in s. Then there exists the following minimal state variable realization of (6.1) :

$$\begin{aligned} \dot{x}_p &= (F + g_1(K)e_1^T)x_p + g_2(K)u \\ y &= e_1^T x_p \end{aligned} \tag{6.2}$$

where $f(s) = \det(sI - F)$ is Hurwitz, $e_1^T = [1,0 \ldots,0]$

$$g_1(K) = G_1K + g_{10} \tag{6.3}$$

$$g_2(K) = G_2K + g_{20} \tag{6.4}$$

$$F = \begin{bmatrix} -f_1 & \\ \cdot & I_n \\ \cdot & \\ \cdot & \\ -f_n & 0 \end{bmatrix} \tag{6.5}$$

and G_1, g_{10}, G_2 and g_{20} are defined as follows :

If

$$G_1 = [g_{11},\ldots,g_{1123..N}]$$

and

$$G_2 = [g_{21},\ldots,g_{2123..N}]$$

then the g_{1r} and g_{2r} are defined by

$$f(s) - [s^{n-1},\ldots,1]g_{10} = q_0(s) \qquad (6.6)$$

$$[s^{n-1},\ldots,1]g_{20} = p_o(s) \qquad (6.7)$$

$$[s^{n-1},\ldots,1]g_{1r} = q_r(s) \qquad (6.8)$$

$$[s^{n-1},\ldots,1]g_{2r} = p_r(s) \qquad (6.9)$$

Let the controller be

$$\begin{aligned} \dot{x}_c &= C_1(\hat{k})x_c + C_2(\hat{k})y + C_3(\hat{k})r \\ u &= C_4^T(\hat{k})x_c + C_5(\hat{k})y + C_6(\hat{k})r \end{aligned} \qquad (6.10)$$

where $\hat{k} \in R^N$ is an estimate of k and

$\hat{K} \in R^{2^N-1}$ is a constrained estimate of K in the sense discussed in Chapter 4; e.g. for $N = 2$ $\hat{K} = [\hat{k}_1,\hat{k}_2,\hat{k}_1\hat{k}_2]^T$, the $\hat{k}_i$ being the elements of $\hat{k}$. We now make the following assumptions on the plant and the controller.

Assumption 6.1

When $\hat{k} = k$ the closed loop system is stable, with no unstable pole-zero cancellations.

Assumption 6.2

The functions $C_1(\cdot),\ldots,C_6(\cdot)$ are piecewise continuous and finite for finite $\hat{k}$. Moreover, $\exists$ an $\varepsilon_1 > 0$ such that they are Lipschitz whenever $\|\hat{k} - k\| < \varepsilon_1$

Assumption 6.3

There exists an $m_1 > 0$ such that

$$\|C_3(\hat{k})\| \; \|C_2(\hat{k})\| \; \|C_4(\hat{k})\| + |C_5(\hat{k})|\,|C_6(\hat{k})| > m_1 .$$

$\forall\hat{k}$, except on sets of measure zero. The inequality must also hold at $\hat{k} = 0$.

Remark: (6.1) Assumptions 6.1 and 6.2 ensure the existence of a neighbourhood of $\hat{k} = k$, of non-trivial but unknown extent for which the closed loop system (6.2), (6.10) is stable.

(6.2) Our formulation excludes the possibility of unstable pole-zero cancellation in the frozen closed loop system at $\hat{k} = k$ and therefore in a non-trivial neighbourhood of it. It does allow stable pole-zero cancellation in the frozen closed loop system everywhere, including $\hat{k} = k$. The coprimeness condition of course precludes stable pole-zero cancellation in the plant itself.

(6.3) Assumption 6.3 ensures that the frozen transfer functions from $r(t)$ and $y(t)$ to $u(t)$ are not zero for almost all values of $\hat{k}$. The requirement that this be true at $\hat{k} = 0$ arises because, in the identifier to be proposed, $\hat{k}(0) = 0$.

Identifier.

Consider the following identifier :

$$\dot{x}_1 = F^T x_1 + e_1 y \tag{6.11}$$

$$\dot{x}_2 = F^T x_2 + e_1 u \tag{6.12}$$

$$h_o(t) = g_{20}^T x_2 - g_{10}^T x_1 \tag{6.13}$$

$$H(t) = G_2^T x_2(t) - G_1^T x_1(t) \tag{6.14}$$

(G_1 and G_2 defined as in (6.3) and (6.4))

$$\dot{R}(t) = -\alpha R(t) + H(t) H^T(t) , R(0) = 0 , t \leq t_o \tag{6.15}$$

$$\frac{d}{dt}[R^{-1}(t)] = \alpha R^{-1}(t) - R^{-1}(t)H(t)H^T(t)R^{-1}(t) \qquad t \geq t_o \tag{6.16}$$

$$\dot{r}_o(t) = -\alpha r_o(t) + h_o(t)H(t) \tag{6.17}$$

$$\dot{\hat{k}}(t) = \begin{cases} 0 & t \leq t_o \\ -\dfrac{\partial \hat{K}^T}{\partial \hat{k}} \Lambda \; [R^{-1}(t)r_o(t) + \hat{K}] & t > t_o \end{cases} \qquad \hat{k}(0) = 0 \tag{6.18}$$

where t_o is the first time instant at which $R(t)$ becomes well conditioned, $\alpha > 0$, and $\Lambda = \text{diag}[\lambda_1, \lambda_2, \ldots \lambda_{123\ldots N}]$, $\lambda_1, \ldots \lambda_N > 0$, $\lambda_{12}, \ldots \lambda_{123\ldots N} \geq 0$. Note that beyond $t = t_o$ only (6.16) as opposed to (6.15) need be implemented and that $\hat{k}(t)$ is constant for all $t \leq t_o$. Equations (6.15 - 6.18) represent the two step least squares identifier formulated in Chapter 4.

In Theorem 6.1 below we show that the existence of t_o, for which $R(t_o)$ is nonsingular, is enough to assure parameter convergence and signal boundedness for the closed loop system defined by (6.2 - 6.18). Theorem 6.2 gives a p.e. condition on the reference input $r(t)$ which ensures that $R(t)$ does indeed become positive definite while $\hat{k}(t)$ is constant. Theorems 6.3 and 6.4 in a sense show that this same condition on $r(t)$ ensures that the overall scheme is robust, as $R(t)$ is uniformly positive definite even beyond $t = t_o$. A further discussion on this robustness property is postponed until after Theorem 6.4.

Theorem 6.1

If $\exists$ a $t_o \in R_+$ such that $R(t_o)$ is invertible then for the system defined by (6.2 - 6.18) and with $r(\cdot) \in \Omega_\Delta[0,\infty)$

(i) $\hat{k}(t)$ converges exponentially to k.

(ii) $\exists$ an M such that

$$\int_0^\infty \|\hat{k}(t) - k\|_2 \, dt < M\|\hat{k}(0) - k\|_2 \tag{6.19}$$

and

(iii) the state $[x_p^T, x_c^T]^T$ and hence all signals appearing in the system are bounded.

Proof:

(i) Some algebraic manipulation shows that

$$h_o(t) + K^T H(t) \equiv 0 \qquad \forall\, t \geq 0 \tag{6.20}$$

and

$$R^{-1}(t_0) r_o(t_0) = -K \tag{6.21}$$

Now consider

$$\frac{d}{dt}[R^{-1}(t) r_o(t)] = R^{-1}(t)\dot{r}_o(t) + \frac{d}{dt}\{R^{-1}(t)\} r_o(t)$$

$$= R^{-1}(t)[-\alpha r_o(t) + h_o(t) H(t)]$$

$$+ [\alpha R^{-1}(t) - R^{-1}(t) H(t) H^T(t) R^{-1}(t)] r_o(t)$$

$$= R^{-1}(t) H(t)[h_o(t) - H^T(t) R^{-1}(t) r_o(t)]$$

$$= R^{-1}(t) H(t)\ \nu(t) \tag{6.22}$$

where $\nu(t) = h_o(t) - H^T(t) R^{-1}(t) r_o(t)$.

From (6.20) and (6.21) it is clear that $\nu(t_o) = 0$. Thus

$$\frac{d}{dt}[R^{-1}(t) r_o(t)] = 0 \qquad \forall\, t \geq t_o$$

whence

$$R^{-1}(t) r_o(t) = R^{-1}(t_0) r_o(t_0) = -K$$

$$\forall\, t \geq t_o \tag{6.23}$$

Thus by Theorem 4.5 of section 4.3.1.2. and the remarks following it one can see that $\hat{k}(t)$ converges to k

exponentially.

(ii) Follows as a consequence of (i).

(iii) Equations (6.2) and (6.10) combine to give

$$\begin{bmatrix} \dot{x}_p \\ \dot{x}_c \end{bmatrix} = \begin{bmatrix} F+g_1(k)e_1^T+g_2(k)C_5(\hat{k})e_1^T & g_2(k)C_4^T(\hat{k}) \\ C_2(\hat{k})e_1^T & C_1(\hat{k}) \end{bmatrix} \begin{bmatrix} x_p \\ x_c \end{bmatrix}$$

$$+ \begin{bmatrix} g_2(k)\,C_6(\hat{k}) \\ C_3(\hat{k}) \end{bmatrix} r(t) \qquad (6.23)$$

$$= \begin{bmatrix} F+g_1(k)e_1^T+g_2(k)\;C_5(k)e_1^T & g_2(k)C_4^T(k) \\ C_2(k)e_1^T & C_1(k) \end{bmatrix} \begin{bmatrix} x_p \\ x_c \end{bmatrix}$$

$$+ \begin{bmatrix} g_2(k)\{C_5(\hat{k}) - C_5(k)\}e_1^T & g_2(k)\{C_4^T(\hat{k}) - C_4^T(k)\} \\ \{C_2(\hat{k}) - C_2(k)\}\,e_1^T & C_1(\hat{k}) - C_1(k) \end{bmatrix} \begin{bmatrix} x_p \\ x_c \end{bmatrix}$$

$$+ \begin{bmatrix} g_2(k)\,C_6(\hat{k}) \\ C_3(\hat{k}) \end{bmatrix} r(t) \qquad (6.24)$$

Thus (6.24) can be written as

$$\begin{bmatrix} \dot{x}_p \\ \dot{x}_c \end{bmatrix} = [A(k) + \Delta A(\hat{k})] \begin{bmatrix} x_p \\ x_c \end{bmatrix} + b(\hat{k})r(t) \qquad (6.25)$$

with $A(k)$, $\Delta A(\hat{k})$ and $b(\hat{k})$ obviously defined. By assumption 6.1, $A(k)$ has eigenvalues in the open left half plane only. Moreover, by assumption 6.2 and (i) it follows that $\Delta A(\hat{k})$ tends to zero exponentially and that $b(\hat{k})$ is bounded. Thus

$$\dot{x} = [A(k) + \Delta A(\hat{k})]x$$

is exponentially stable whence $[x_p^T, x_c^T]^T$ is bounded as $r(t)$ is bounded.

∇∇∇

We now present a p.e. condition under which $R(t)$ is invertible for some $t = t_o$. By Theorem 6.1 this is enough to guarantee the stability of the adaptive law. This result relies on the fact that for $t \le t_o$ the closed loop system is time invariant.

<u>Theorem 6.2</u>:

Consider the closed loop system (6.2) and (6.10) with $\hat{k}$ constant. Suppose that assumptions 4.1 and 6.3 hold, $r(t) \in \Omega_\Delta[0,\infty)$ and $\exists\ \alpha_1, \alpha_2, \delta > 0$ such that $\forall\ \sigma \in R_+$

$$\alpha_1 I \le \int_\sigma^{\sigma+\delta} \eta^{(2n+m)}(t)\eta^{T(2n+m)}(t)\ dt \le \alpha_2 I \tag{6.26}$$

where n is the dimension of (6.2), m is the highest degree among the polynomials $p_o q_r - q_o p_r\ \forall r \subset S$ and $p_r q_{\bar{r}} - q_r p_{\bar{r}}\ \forall r, \bar{r} \subset S$.

$$\eta^{(2n+m)}(t) \triangleq [r(t), \frac{r(t)}{s+\beta}, \ldots, \frac{r(t)}{(s+\beta)^{2n+m-1}}]^T \; ; \; \beta > 0 \; .$$

Then $\exists \; \alpha_3 > 0$ such that $R(t)$ defined in (6.15) obeys

$$R(t) > \alpha_3 I \qquad (6.27)$$

for some $t = t_o > 0$.

Proof.

For $t \le t_o$, $\hat{k} = 0$, and the closed loop system is time invariant.

From (6.15)

$$R(t) = \int_0^t e^{-\alpha(t-\tau)} H(\tau)H^T(\tau)d\tau$$

Thus by Theorem 5.4 one can see that (6.27) is satisfied if $\exists \; \alpha_4$, $\delta_1 > 0$ such that $\forall \; \sigma \in R_+$,

$$\int_\sigma^{\sigma+\delta_1} H(t)H^T(t)dt > \alpha_4 I \; .$$

The overall system is representable as in Fig. 3.3 with the block A' (relating $r(t)$ to $\mathbf{u}(t)$) being strictly proper, time invariant non-minimal and with degree $2n$, n being the degree of the plant. Then combining technique used in the proofs of Theorems 3.6 - 3.8 and Theorem 4.2, the result can be proved.

∇∇∇

Remarks.

(6.4) When $r(t)$ is a linear combination of sinusoids, (6.26) demands that $(2n+m-1)/2$ different frequencies be present. As argued in Chapter 4, the regression vector $H(t)$ is p.e. if $u(t)$ has $m/2$ frequency components. Moreover the non-minimal system relating $u(t)$ to $r(t)$ may have as many as $2n-1$ zeros on the imaginary axis every two of which may cancel out one frequency component. This accounts for the number $(2n+m-1)/2$.

(6.5) Assumption 6.3 ensures that the transfer function from r to u is not zero. We have shown that a reference input satisfying (6.26) guarantees the global stability of the proposed adaptive controller. Theorem 6.3 shows that this condition also ensures the uniform positive definiteness of $R(t)$ beyond $t = t_o$ as long as (6.19) holds; i.e. the parameter error is in L^2. Theorem 6.4 presents a stronger result by showing that even if (6.19) is not satisfied, but the parameter variations are small over all intervals of a fixed length, then $R(t)$ will still be uniformly positive definite. As discussed in Chapter 4, such a positive definiteness will impart to the algorithm the form of robustness we have mentioned.

Theorem 6.3:

Consider the closed loop system defined by (6.2) and (6.10) with $r(t) \in \Omega_\Delta[0,\infty)$. Suppose that assumptions (4.1) and (6.3) hold, conditions (6.19) is satisfied, all signals are bounded and $\exists\ \alpha_1, \alpha_2, \delta > 0$ such that $\forall\ \sigma \in R_+$

$$\alpha_1 I \leq \int_{\sigma}^{\sigma+\delta} \eta(t)^{(2n+m)} \eta(t)^{(2n+m)} dt \leq \alpha_2 I \qquad (6.28)$$

$\eta(.)^{(.)}$ defined in Theorem 6.2.

Then $\exists\ \alpha_3,\ t_o > 0$ such that $\forall\ t \geq t_o$

$$R(t) > \alpha_3 I \qquad (6.29)$$

Proof:

By Theorem 5.4 , the existence of $\alpha_4,\ \delta_1 > 0$ such that $\forall\ \sigma \in R_+$

$$\int_{\sigma}^{\sigma+\delta_1} H(t)H^T(t)dt > \alpha_4 I \qquad (6.30)$$

guarantees (6.29) . As shown in Theorem 6.1 the closed loop system can be represented as

$$\dot{x} = [A(k) + \Delta A(\hat{k})]x + [B(k) + \Delta B(\hat{k})]r(t) \qquad (6.31)$$

where, by (6.19) and assumption 6.2 , $\int_0^{\infty} \| \Delta A(\hat{k})\| dt$ and $\int_0^{\infty} \| \Delta B(\hat{k})\| dt$ are finite.

Let x^* be defined by

$$x^* = A(k)x^* + B(k)r(t) \qquad (6.32)$$

with H^* , y^* and u^* obviously defined. Then by (6.19) $\exists\ \varepsilon_1$ and $T > 0$ such that $\| \hat{k}(t) - k\|_2 < \varepsilon_1$, ε_1 defined in assumption 6.2 . Thus $\exists\ \overline{M}_1 < \infty$

$$\int_T^{\infty} \|x - x^*\|_2\ dt < \overline{M}_1 \qquad (6.33)$$

As A(k) has eigenvalues in the open left hand plane (Assumption 6.1)

$$\int_0^T \| x - x^* \|_2 \, dt < \bar{M}_2$$

for some finite $\bar{M}_2$.

Thus $\exists$ $\bar{M}_3 > 0$ such that

$$\int_0^\infty \| x - x^* \|_2 < \bar{M}_3$$

Hence

$$\int_0^\infty \| H - H^* \|_2 < \bar{M}_4 \qquad (6.34)$$

Since H^* is the output of a time-invariant system, arguments similar to those used in Theorem 6.2 show that (6.28) implies the existence of α_5 , $\delta_2 > 0$

$$\int_\sigma^{\sigma+\delta_2} H^*(t)H^{*T}(t)\,dt > \alpha_5 I \ .$$

Thus by Lemma 5.A.1 the result follows. ∇∇∇

Remark:

(6.6) The hypotheses of Theorem 6.3, in particular the satisfaction of (6.19) and the boundedness of system signals, follow from Theorem 6.1, whenever the idealizing assumptions, of no noise, no modelling error, no time variation, apply. The case where departures occur from these assumptions is dealt with after Theorem 6.4.

Theorem 6.4

For the closed loop system (6.2) and (6.10), with $r(t) \in \Omega_\Delta[0,\infty)$ let the following be true :

(i) assumptions (4.1), (6.1), (6.2) and (6.3) hold;

(ii) for some $T > 0 \;\exists\; M_5, \delta > 0$ such that $\forall\; i=1,\dots 6$ and $t \geq T$

$$\| C_i(\hat{k}(t+\delta)) - C_i(\hat{k}(t)) \|_2 \leq \overline{M}_5 \; ; \qquad (6.35)$$

(iii) $\exists\; \overline{M}_6 > 0$ such that the closed loop system state $[x_p^T, x_c^T]$ obeys

$$\| [x_p^T(t), x_c^T(t)]^T \| \leq \overline{M}_6 \qquad \forall\; t \in R_+ \qquad (6.36)$$

(iv) $\exists\; \alpha_5, \alpha_6$, $\delta' > 0$, α_5, depending on $\overline{M}_5$, $\overline{M}_6$ and β (defined in Theorem 6.2) such that $\forall\; \sigma > T$

$$f(\overline{M}_5, \overline{M}_6) < \alpha_5 I \leq \int_\sigma^{\sigma+\delta} \eta^{(2n+m)}(t) \, \eta^{(2n+m)T}(t) \, dt \leq \alpha_6 I$$

(6.37)

The function $f(\overline{M}_5, \overline{M}_6)$ is a non-decreasing function of $\overline{M}_5$ and $\overline{M}_6$, which is zero when $\overline{M}_5$ equals zero. The vector $\eta^{(.)}(.)$ is defined in Theorem 6.2.

Then $\exists\; \alpha_7 > 0$ and $T_1 > T$ such that

$$R(t) \geq \alpha_7 I \qquad (6.38)$$

for all $t > T_1$. ∇∇∇

Proof:

By arguing as in Theorem 5.4 one can see that (6.38) is satisfied whenever $\exists\ \alpha_8$, T, $\delta > 0$ for which

$$\int_{\sigma}^{\sigma+\delta} H(t)\, H^T(t)dt \geq \alpha_8 I \tag{6.39}$$

$\forall\ \sigma > T$. Then the result can be proved by arguing as in Theorems 3.9 and 3.6. ∇∇∇

Remarks:

(6.7) The quantity $\overline{M}_5$ is a measure of the extent of time variation in the system. The nonsingularity of R(t) demands that α_5 be "large enough" to overcome this time variation. For time-invariant systems, clearly, α_5 can be any positive constant. In the ideal case (6.19) guarantees the existence of T defined in (6.35).

(6.8) Theorems 6.1 and 6.2 show that if r(t) is "sufficiently rich", viz. has at least 2n+m-1 frequencies, with m defined in the theorem statement, then $R(t_o)$ will become nonsingular, as for $t \leq t_o$ we have in effect a time invariant albeit possibly unstable system. Thus under ideal circumstances this condition on r(t) guarantees the exponential convergence of the parameter estimates. To assess the robustness of the algorithm consider first the properties of adaptive control algorithms which exist in the literature. These turn out to give bounded signals under ideal settings. But in them tracking error convergence is not exponential and while parameter variation does die down, the parameter estimates do not converge to their true values. However, given that the signals are bounded under ideal conditions, one can show, that in some of these schemes once parameter variations fall below a certain value, "sufficiently rich" inputs guarantee exponential parameter convergence.

As opposed to this, Theorems 6.1 and 6.2 show, that our algorithm is exponentially convergent given "sufficiently rich" r(t) and ideal conditions. Also Theorem 6.3 shows that non-singularity of R(t) follows for all $t \geq t_o$ and the overall algorithm is thus equipped with the robustness characteristics mentioned in Chapter 4.

Further, in face of modest departures from ideality the following mechanisms will guarantee robustness :

(1) For "small" errors while $\hat{k}$ may not converge to k , it will in a finite time, reach that neighbourhood of k , mentioned in remark 6.1 , where the closed loop system is stable. Thus even with small departures from ideality the system signals will retain boundedness.

(2) More importantly Theorem 6.4 shows that this very same r(t) guarantees non-singularity of R(t) once the time variations fall below a certain value. Thus while large $\dot{\hat{k}}$ immediately after t_o may destroy persistence of excitation, within a finite time $||\dot{\hat{k}}||$ will fall below this required value and R(t) will once again become non-singular thereby imparting a further measure of robustness.

Simulations presented in Section 6.3 demonstrate the robustness of this algorithm. They also demonstrate one other important characteristic. Observe that in the ideal case $R^{-1}(t)r_o(t) = -K \quad \forall\, t \geq t_o$. Thus ideally one could well dispense with (6.18) and use $R^{-1}(t)r_o(t)$ directly to obtain k . The simulation results demonstrate, however, that the second step (6.18) of the identification part of the algorithm results in substantial improvements in the $\hat{k}_i$.

6.2 Adaptive Control Using a Gradient Descent Identifier:

This section considers an adaptive controller using a two step gradient descent algorithm mentioned in Chapter 4 . Consider the plant described by (6.1 - 6.9). Let the controller be

$$\dot{x}_c = C_1(\bar{K})x_c + C_2(\bar{K})y + C_3(\bar{K})r \qquad (6.40)$$

$$u = C_4^T(\bar{K})x_c + C_5(\bar{K})y + C_6(\bar{K})r$$

where $\bar{K}$ is an unconstrained estimate of K , selected in a manner described later. For the remainder of this section the following definitions and assumptions will hold.

<u>Definition 6.1</u>

The convex regions J_1 and J_2 are defined as follows:

$$\text{(i)} \qquad z \in J_1 \quad \Rightarrow \quad \begin{cases} 0 \leq z_i \leq M_{i1} & \forall\, i \in S \\ \\ 0 \leq z_r \leq \prod_{i \in r} M_{i1} & \forall\, r \subset S \end{cases} \qquad (6.41)$$

$$\text{(ii)} \qquad z \in J_2 \Rightarrow \begin{cases} m_i \leq z_i \leq M_{i2} & \forall\, i \in S \\ \prod_{i \in r} m_i \leq z_r \leq \prod_{i \in r} M_{i2} & \forall\, r \subset S \end{cases} \qquad (6.42)$$

where $0 < m_i < M_{i2} < M_{i1}$. ∇∇∇

Note that $J_2 \subset J_1$.

Assumption 6.4

Suppose the plant is denoted by $P(\overline{K})$ and the controller by $C(\overline{K})$. Then the closed loop system $[P(\overline{K}),C(\overline{K})]$ is asymptotically stable, with no unstable pole zero cancellation whenever $\overline{K} \in J_1$.

Remark

(6.9) Assumption 6.4 can always be satisfied by a linear translation on the k_i .

Assumption 6.5

The parameter vector $K \in J_2$.

Assumption 6.6

The functions $C_1(\cdot) \ldots C_6(\cdot)$ are bounded and piecewise continuous. Moreover, in some neighbourhood of J_1 and in J_1 itself they are Lipschitz as well.

Assumption 6.7

There exists a $m > 0$ such that

$$\| C_3(\overline{K}) \| \, \| C_2(\overline{K}) \| \, \| C_4(\overline{K}) \| + |C_5(\overline{K})| \, |C_6(\overline{K})| > m \qquad (6.43)$$

$\forall\ \bar{K}$ except on a set of measure zero. This inequality holds when $\bar{K} \in J_1$.

Identifier

The following two step identifier is used with $H(t)$ defined in (6.11) - (6.14) .

$$\dot{K}_u(t) = -\frac{H(t)(h_o(t)+K_u^T H(t))}{\beta_2 + H^T(t)H(t)} - \Gamma_1 \Psi_1(K_u(t)) \qquad (6.44)$$

$$\beta_2 > 0\ ,\ K_{ur}(0) \in J_1$$

$$\psi_{r1}(K_u(t)) \triangleq \begin{cases} K_{ur} - \prod_{i \in r} m_i & \text{when } K_{ur} \leq \prod_{i \in r} m_i \\ K_{ur} - \prod_{i \in r} M_{i2} & \text{when } K_{ur} \geq \prod_{i \in r} M_{i2} \\ 0 & \text{elsewhere} \end{cases}$$

(6.45)

$$\dot{\hat{k}}(t) = -\frac{\partial \hat{K}^T}{\partial \hat{k}} \Lambda [\hat{K}-K_u] - \Gamma_2 \Psi_2(\hat{k}(t)),\ \hat{K}(0) \in J_1 \qquad (6.46)$$

$$\psi_{i2}(\hat{k}(t)) = \begin{cases} \hat{k}_i(t) - M_{i2} & \text{when } \hat{k}_i > M_{i2} \\ 0 & \text{elsewhere} \end{cases} \qquad (6.47)$$

Here Λ is as defined in (6.18) , K_{ur} is the r^{th} element of K_u , $\Gamma_1 = \text{diag}\{\gamma_1, \gamma_2, \ldots \gamma_{123\ldots N}\}$

with

$$\gamma_r > \frac{\sum_{\bar{r} \subset S} (\prod_{i \in \bar{r}} M_{i1})^2}{\min\{\prod_{i \in r} m_i, \prod_{i \in r} M_{i1} - \prod_{i \in r} M_{i2}\}} \qquad (6.48)$$

and

$$\Gamma_2 = \text{diag}\{\bar{\gamma}_1, \ldots \bar{\gamma}_N\}$$

with

$$\bar{\gamma}_i > \frac{1}{M_{i1} - M_{i2}} \sum_{\substack{r \subset S \\ i \in r}} \lambda_r [\prod_{\substack{\alpha \in r \\ i \neq \alpha}} M_{\alpha 1}][\prod_{\beta \in r} M_{\beta 1}] \quad (6.49)$$

Remarks

(6.11) The role of Ψ_1 is to ensure that all elements of $K_u(t)$ lie in J_1, and as will be shown later, $K_u(t)$ enters J_2 asymptotically. The function Ψ_2 by the same token ensures that $\hat{K}(t)$ is always confined to the region J_1.

Selecting the controller parameter

Equation (6.50) defines the way in which $\bar{K}(t)$, used in the controller (6.40), is selected. Define $\ell = |(\hat{K} - K_u)^T H|$ and let $\bar{m} < \bar{M} \in R_+$

$$\bar{K} = \begin{cases} \hat{K} & \text{when } \ell \leq \bar{m} \\ \dfrac{\bar{M}-\ell}{\bar{M}-\bar{m}} \hat{K} + \dfrac{\ell-\bar{m}}{\bar{M}-\bar{m}} K_u & \bar{m} \leq \ell \leq \bar{M} \\ K_u & \ell \geq \bar{M} \end{cases} \quad (6.50)$$

where $\bar{m}$ and $\bar{M}$ are arbitrary preselected positive numbers.

Remarks:

(6.12) Equation (6.50) shows that the parameter estimate, $\bar{K}$, used in the controller is selected as

$\hat{K}$, the constrained estimate, only when $\hat{K}$ is in some sense close to K_u . Elsewhere either a convex combination of $\hat{K}$ and K_u or K_u itself are used. The parameters $\bar{m}$ and $\bar{M}$ are design parameters which determine the extent to which one choice of $\bar{K}$ is preferred over the others.

Theorem 6.5 given below is an analysis of the identifier only, independent of controller characteristics and of any p.e. conditions. Theorem 6.6 uses this theorem to establish the boundedness of system signals.

Theorem 6.5

For the system described by (6.1 - 6.9), (6.11 - 6.14), $r(t) \in \Omega_\Delta[0,\infty)$ and the identifier (6.44 - 6.49) the following hold :

(i) $K_u(t),\ \hat{K}(t) \in J_1\ \forall\ t \geq 0$

(ii) $\lim_{t\to\infty} \Psi_1(t) = 0$ (6.51)

(iii) $\bar{K}(t) \in J_1\ \forall\ t > T_1$

(iv) $\lim_{t\to\infty} \frac{\{H^T(K_u - K)\}^2}{\beta_2 + H^T H} = 0$ (6.52)

(v) $\lim_{t\to\infty} \dot{K}_u(t) = 0$ (6.53)

(vi) $\lim_{t\to\infty} \dot{\hat{k}}(t) = 0$ (6.54)

Proof

(i) By (6.20) and (6.44) one has

$$\dot{K}_u(t) = -\frac{H(t)H^T(t)[K_u(t)-K]}{\beta_2 + H^T(t)H(t)} - \Gamma_1 \Psi_1(K_u(t)) \qquad (6.55)$$

Initially $K_u(t)$ and $\hat{K}(t) \in J_1$. Let any K_{ur} equal zero or $\prod_{i \in r} M_{i1}$. Then one can show that the corresponding derivative in (6.55) is positive or negative respectively. Thus $K_u(t) \in J_1$. Thus by analysis similar to lemma 4.2 all elements of $\hat{K}(t)$ are positive. Moreover, if the ith element of $\hat{k}$ equals M_{i1} then $\dot{\hat{k}}_i < 0$. Thus $\hat{K}(t) \in J_1$.

(ii) (iii) and (iv)

Let

$$L_1 = \frac{1}{2}[K_u - K]^T[K_u - K]$$

then

$$\dot{L}_1 = -\frac{[K_u - K]^T HH^T [K_u - K]}{\beta_2 + H^T H} - [K_u - K]^T \Gamma_1 \Psi_1(K_u)$$

(6.56)

which by the definition of Ψ_1 is negative semi-definite. Integration of (6.56) on $[t_o, \infty)$, $\forall\, t_o \geq 0$ yields

$$\int_{t_o}^{\infty} \frac{\{(K_u - K)^T H\}^2}{\beta_2 + H^T H}\,dt + \int_{t_o}^{\infty} (K_u - K)^T \Gamma_1 \Psi_1(K_u)\,dt \qquad (6.57)$$

$$= L_1(t_o) - L_1(\infty) < \infty$$

It can be shown that the derivative of the integrands in (6.57) are bounded and that both integrands are non-negative. Thus, both integrands converge to zero whence (ii), (iii) and (iv) follow.

(v) It follows from (iii) and (iv) that $\lim_{t\to\infty} \dot{K}_u(t) = 0$.

(vi) Consider the non-negative function

$$L_2(t) = \frac{1}{2}[\hat{K}(t) - K_u(t)]^T \Lambda \, [\hat{K}(t) - K_u(t)]$$

$$+ \frac{1}{2} \Psi_2^T(\hat{k}) \, \Gamma_2 \, \Psi_2(\hat{k}) \qquad (6.58)$$

As Γ_2 is diagonal

$$\frac{1}{2} \frac{\partial}{\partial \hat{k}} \{\Psi_2^T(\hat{k}) \quad \Gamma_2 \, \Psi_2(\hat{k})\} = \Gamma_2 \Psi_2(\hat{k})$$

Thus

$$\dot{L}_2(t) = - [\frac{\partial L_2}{\partial \hat{k}}]^T [\frac{\partial L_2}{\partial \hat{k}}] + [\hat{K} - K_u]\dot{K}_u$$

$$\leq -\| \dot{\hat{k}}(t) \|^2 + \alpha_2(t)$$

where $\lim_{t\to\infty} \alpha_2(t) = 0$, as $\hat{K} - K_u$ is bounded and $\dot{K}_u \to 0$.

Thus for arbitrary $\varepsilon_2 > 0 \; \exists \; T_3$ such that $\alpha_2(t) < \varepsilon_2$ $\forall t > T_3$. Thus $\forall t > T_3$

$$\dot{L}_2(t) \le - \| \dot{\hat{k}}(t) \|^2 + \varepsilon_2$$

Since L_2 is non-negative, for arbitrary $\varepsilon_3 > \varepsilon_2$ a T_4 such that

$$\| \dot{\hat{k}}(t) \|^2 < \varepsilon_3 \qquad \forall\ t > T_3 .$$

Thus

$$\lim_{t \to \infty} \| \dot{\hat{k}}(t) \|^2 = 0$$

whence (vi) follows. ∇∇∇

Remark

(6.13) We have thus shown that $\bar{K}$ is always in J_1 , that its variations decay asymptotically to zero and that K_u converges to J_2 . Then the following result shows that all system signals are bounded.

Theorem 6.6

For the closed loop system (6.1 - 6.9), (6.11 - 6.14) and (6.44 - 6.50) all signals are bounded, whenever $r(t) \in \Omega_\Delta[0,\infty)$ and assumptions (6.4 - 6.6) hold.

Proof

From (6.2 - 6.9) and (6.11 - 6.14) one can see that

$$y(t) = g_2^T(K)x_2 - g_1^T(K)x_1$$

$$= g_2^T(\bar{K})x_2 - g_1^T(\bar{K})x_1 - (\bar{K} - K)^T H \qquad (6.59)$$

Thus the plant can be re-expressed as

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ y \end{bmatrix} = \begin{bmatrix} F^T x_1 + e_1^T y \\ F^T x_2 + e_1^T u \\ g_2^T(\overline{K}) x_2 - g_1^T(\overline{K}) x_1 - (\overline{K} - K)^T H \end{bmatrix}$$

with $x^T = [x_1^T, x_2^T, x_c^T]$ we have the closed loop system expressible as

$$\dot{x}(t) = A(\overline{K}) x(t) - \Delta b(\overline{K}) (\overline{K} - K)^T H(t) + r'(\overline{K},t) \qquad (6.60)$$

where

$$A(\overline{K}) = \begin{bmatrix} F^T - e_1 g_1^T(\overline{K}) & e_1 g_2^T(\overline{K}) & 0 \\ -e_1 g_1^T(\overline{K}) C_5(\overline{K}) & F^T + e_1 g_2^T(\overline{K}) C_5(\overline{K}) & e_1 C_4^T(\overline{K}) \\ -C_2(\overline{K}) g_1^T(\overline{K}) & C_2(\overline{K}) g_2^T(\overline{K}) & C_1(\overline{K}) \end{bmatrix} \qquad (6.61)$$

$$\Delta b(\overline{K}) = \begin{bmatrix} e_1 \\ e_1 C_4(\overline{K}) \\ C_2(\overline{K}) \end{bmatrix} \qquad (6.62)$$

$$r'(\overline{K},t) = \begin{bmatrix} 0 \\ e_1 C_6(\overline{K}) r(t) \\ C_3(\overline{K}) r(t) \end{bmatrix} \qquad (6.63)$$

By assumption (6.4) $A(\overline{K})$ is asymptotically stable when

$\bar{K} \in J_1$.

Also by (6.50) $0 \leq \alpha_3(t) \leq 1$ such that

$$(\bar{K}-K)^T H(t) = \alpha_3(t)(\hat{K}(t)-K)^T H(t) + (1-\alpha_3(t))(K_u(t)-K)^T H(t)$$

$$= \alpha_3(t)(\hat{K}(t) - K_u(t))^T H(t) + (K_u(t)-K(t))^T H(t)$$

where a nonzero $\alpha_3(t)$ implies that

$$(\hat{K}(t) - K_u(t))^T H(t) \leq M .$$

Thus

$$(\bar{K} - K)^T H(t) = \alpha_4(t) + (K_u(t) - K(t))^T H(t)$$

where

$$|\alpha_4(t)| \leq M \qquad \forall\ t \geq 0 .$$

Thus (6.60) can be re-expressed as

$$\dot{x}(t) = \tilde{A}(\bar{K},x,t)x + \tilde{r}(\bar{K},x,t) \tag{6.64}$$

where

$$\tilde{A}(\bar{K},x,t) = A(\bar{K}) - \frac{(K_u-K)^T H(t)}{\sqrt{\beta_2 + x_c^T x_c + H^T H}} \frac{\sqrt{\beta_2 + x_c^T x_c + H^T H}}{\beta_2 + x^T x} \Delta b(\bar{K}) x^T$$

and

$$\tilde{r}(\bar{K},x,t) = r'(\bar{K},t) - \frac{(K_u-K)^T H(t)\sqrt{\beta_2 + H^T H + x_c^T x_c})\beta_2 \Delta b(\bar{K})}{\sqrt{\beta_2 + x_c^T x_c + H^T H}\ \ (\beta_2 + x^T x)}$$

$$- \alpha_4(t)\ \Delta b(\bar{K})$$

It is possible to show that

$$\frac{\beta + x_c^T x_c + H^T H}{\beta + x^T x}$$

is bounded. Also, by (iii) of Theorem 6.5

$$\lim_{t\to\infty} \frac{H^T(K_u - K)}{(\beta + x_c^T x_c + H^T H)^{\frac{1}{2}}} = 0$$

and

$$\lim_{t\to\infty} \overline{K}(t) = \overline{K}^* \ , \qquad \overline{K}^* \in J_1$$

Thus $\tilde{r}$ is bounded, and $\lim_{t\to\infty} \tilde{A}(\overline{K},x,t) = A(\overline{K}^*)$.

Thus by a result in [2] x is bounded, whence, u, x_c, y are bounded. Thus in (6.2) as F is asymptotically stable, x_p is bounded. ∇∇∇

<u>Remark</u>

(6.14) From (6.60) one can see that the closed loop system is driven by the output error $(\overline{K}-K)^T H$. The exploitation of this feature can also be found in [2] .

We have thus established that as long as the reference input, $r(t)$, is bounded the closed loop system will be stable, under ideal settings. But, for the algorithm to be robust we need more. We need to show that $\overline{K}$ converges to K at a rate which is exponential. Theorem 6.7, stated below, shows that this is guaranteed if $\exists\ T > 0$

such that the gramian

$$\int_{\sigma}^{\sigma+\delta} H(t)H^T(t)dt \ , \qquad (6.65)$$

is uniformly positive definite for all $\sigma > T$ and if $H(\cdot)$ is bounded. By Theorem 6.5, $\dot{\bar{K}}(t)$ converges to zero asymptotically. Thus there exists a $T > 0$ such that $||\dot{\bar{K}}(t)||$ becomes small for all $t > T$. Thus considerations similar to the arguments in the proofs of Theorems 3.9 and 6.4 show that the satisfaction of (6.26) ensures that (6.65) is uniformly positive definite for all $t > T$.

Theorem 6.7

Consider the differential equations (6.44) and (6.46). Suppose there exist $T, \delta, \alpha_6, \alpha_7 > 0$ such that $\forall \sigma > T$

$$\alpha_6 I \le \int_{\sigma}^{\sigma+\delta} H(t)H^T(t)dt \le \alpha_7 I \qquad (6.66)$$

Then $\bar{K}(t)$ converges to K at an exponential rate.

Proof.

Consider the Lyapunov function

$$L_3 = \frac{1}{2}[\hat{K}(t) - K]^T \Lambda \ [\hat{K}(t) - K] + \frac{1}{2} \Psi_2^T(\hat{k})\Gamma_2\Psi_2(\hat{k})$$

Recall also the Lypaunov function

$$L_1 = (K_u(t) - K)^T(K_u(t) - K)$$

defined in Theorem 6.5.

Now,

$$\frac{\partial L_3}{\partial \hat{k}(t)} = \frac{\partial \hat{K}^T(t)}{\partial \hat{k}(t)} \Lambda \ (\hat{K}(t) - K) + \Gamma_2 \ \Psi_2(\hat{k})$$

By the definition of $\Psi_2(\hat{k})$ it follows that since $\hat{K}(t) \in J_1$

$$\ell n \ \frac{\hat{k}_i}{k_i} \ \{\frac{\hat{k}_i}{k_i}\} \ \Psi_{2i}(\hat{k}) \geq 0$$

Thus as in the proof of lemma 4.2 we have that

$$\frac{\partial L_3}{\partial \hat{k}(t)} = 0 \quad \text{iff} \quad \hat{k} = k \ , \text{ as long as } \quad \hat{K} \in J_1 \ .$$

By (6.44) , (6.45) , (6.56) and the boundedness of $H(t)$

$$\dot{L}_1 \leq -\beta_1 (K_u(t) - K)^T HH^T (K_u(t) - K$$

for some $\beta_1 > 0$.

Whence by an extension of a result in [10] ,

$$|| K_u(t) - K|| \leq \beta_3 || K_u(T) - K|| \ e^{-\beta_4 (t-T)} \qquad \forall t \geq T$$

and some $\beta_3, \beta_4 > 0$. As $K_u(t)$ is bounded, there exists an M such that $|| K_u(T) - K|| < M$.

Thus by a result in [11,p86] there exists a Lyapunov function L_4 such that for some $\mu_1, \mu_2, \mu_3 > 0$

$$\mu_1 || K_u(t) - K||^2 \leq L_4 \leq \mu_2 || K_u(t) - K||^2$$

and

$$- \dot{L}_4 \leq \mu_3 || K_u(t) - K||^2$$

$\forall \ t \geq T$. Moreover, as $\hat{k}(t)$ is bounded $\overline{M}$ such that

$$\left\| \frac{\partial \hat{K}^T(t)}{\partial \hat{k}(t)} \Lambda \right\| \leq \overline{M} < \infty$$

Define

$$L = \frac{1}{2} L_1 + \beta_5 L_4$$

where

$$\beta_5 = 1 + \frac{\overline{M}^2}{2\mu_3}$$

Clearly L is positive definite in $[\hat{k}^T - k^T, K_u^T - K^T]^T$.

Also

$$\dot{L}(t) = - \frac{\partial L_3(t)}{\partial \hat{k}(t)} \{ \frac{\partial \hat{K}^T(t)}{\partial \hat{k}(t)} \Lambda \; (\hat{K}(t) - K) + \frac{\partial \hat{K}^T(t)}{\partial \hat{k}(t)} \Lambda \; (K - K_u(t))$$

$$+ \Gamma_2 \Psi_2(\hat{k}) \} - \beta_5 \mu_3 \| K_u(t) - K \|^2$$

$$\leq - \left\| \frac{\partial L_3(t)}{\partial \hat{k}(t)} \right\|^2 + \overline{M} \left\| \frac{\partial L_3(t)}{\partial \hat{k}(t)} \right\| \| K_u(t) - K \|$$

$$- \beta_5 \mu_3 \| K_u(t) - K \|^2$$

$$\leq - \frac{1}{2} \left\| \frac{\partial L_3(t)}{\partial \hat{k}(t)} \right\|^2 - \mu_3 \| K_u(t) - K \|^2$$

which is negative definite in $[\hat{k}^T(t) - k^T, \; K_u^T(t) - K^T]^T$, as $\frac{\partial L_3}{\partial \hat{k}}$ has no explicit dependence on t. Thus as in the proofs of Theorem 4.5 and 6.1, $\hat{k}(t)$ converges to k exponentially. Thus $\overline{K}(t)$ converges to K exponentially. ∇∇∇

6.3 Simulation Results

In this section we present simulation results using adaptive controllers typified by our first algorithm. The plant being considered has transfer function

$$T_p(s) = \frac{s+1}{(s+3(s-1)}$$

so modelled as to ensure that $k_1 = 3$ and $k_2 = -2$. The object is to place the closed loop poles at -3 and -4. A standard pole placement algorithm is used, with controller parameters being computed as a function of the parameter estimates $\hat{k}_1, \hat{k}_2$. The controller parameters are held constant over the region where the Sylvester determinant [12, p142] has magnitude less than .1 . The state variable filters used have transfer function $\frac{1}{(s+5)(s+6)}$. The identifier (6.18) has $t_o = 3$.

Figure 6.1 gives the results in the absence of modelling inadequacies. In the first plot the lines 1 and 2 refer to $\hat{k}_1(t)$ and $\hat{k}_2(t)$, which converge rapdily to 3 and -2, respectively. The second plot depicts the output, which after $t = 3$, is quickly stabilized.

Next consider the presence of an unmodelled mode at $s = -60$.

Figure 6.2 lists the output y, which is stabilized despite the unmodelled mode. The first and second plots of figure 6.3 depict the constrained estimate $\hat{k}_i$ and the first two elements of the unconstrained estimate K_u, respectively. The steady state values are

$$\hat{k}_1 = 3.1074 \qquad K_{u1} = 3.4637$$

$$\hat{k}_2 = -1.8653 \qquad K_{u2} = -2.4614$$

Clearly $\hat{k}_i$ track the respective k_i values more closely. Also the $\hat{k}_i$ convergence is significantly smoother. Another feature to note is that at $t = 3$, as $\hat{k}(t)$ starts to be updated the unmodelled mode is excited and this causes the orthant condition to be violated, significantly. Yet, the algorithm is robust enough to withstand this violation and to force the $\hat{k}_i$ to return to the correct orthant.

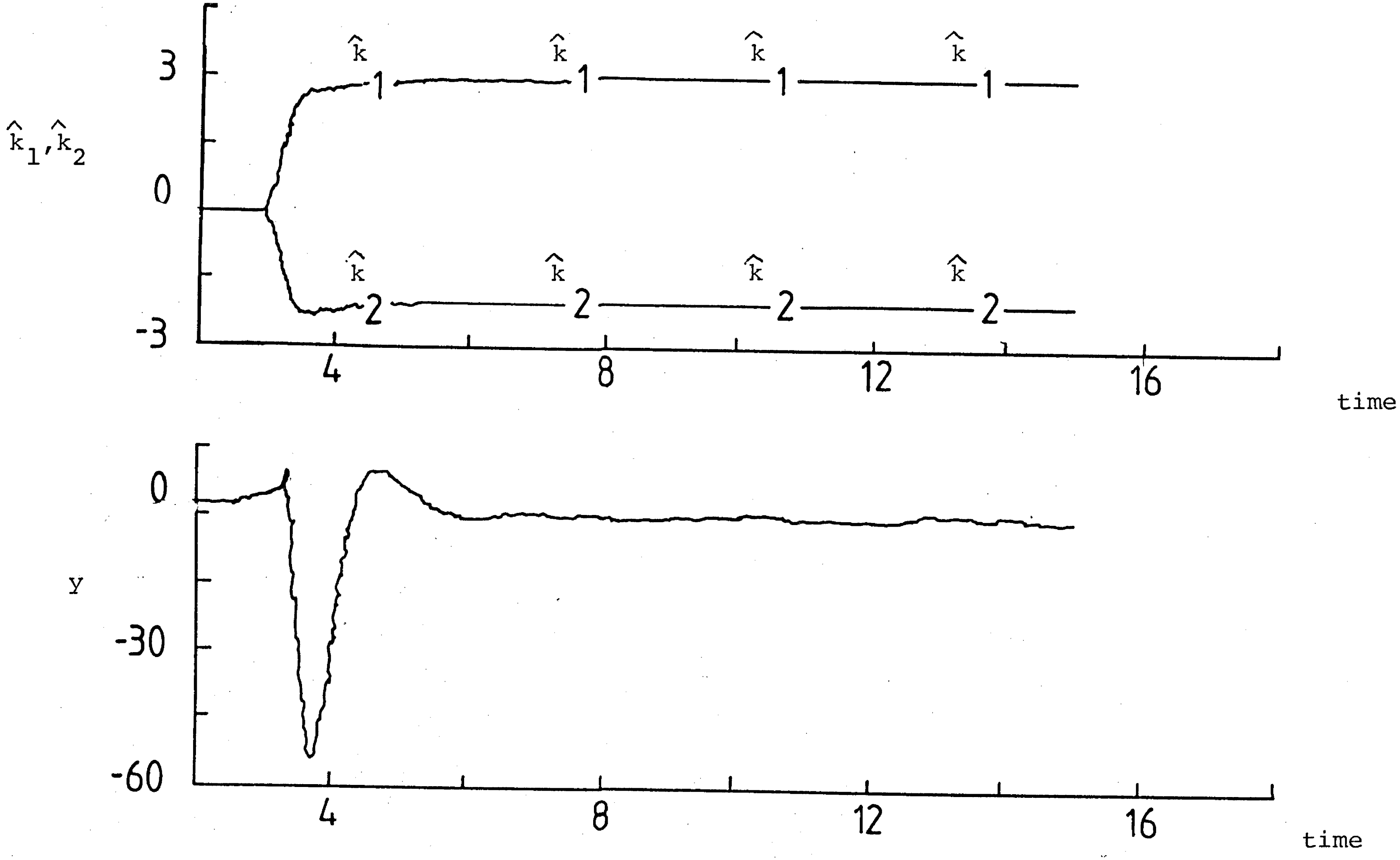


Figure 6.1 Results of an adaptive pole placement algorithm.

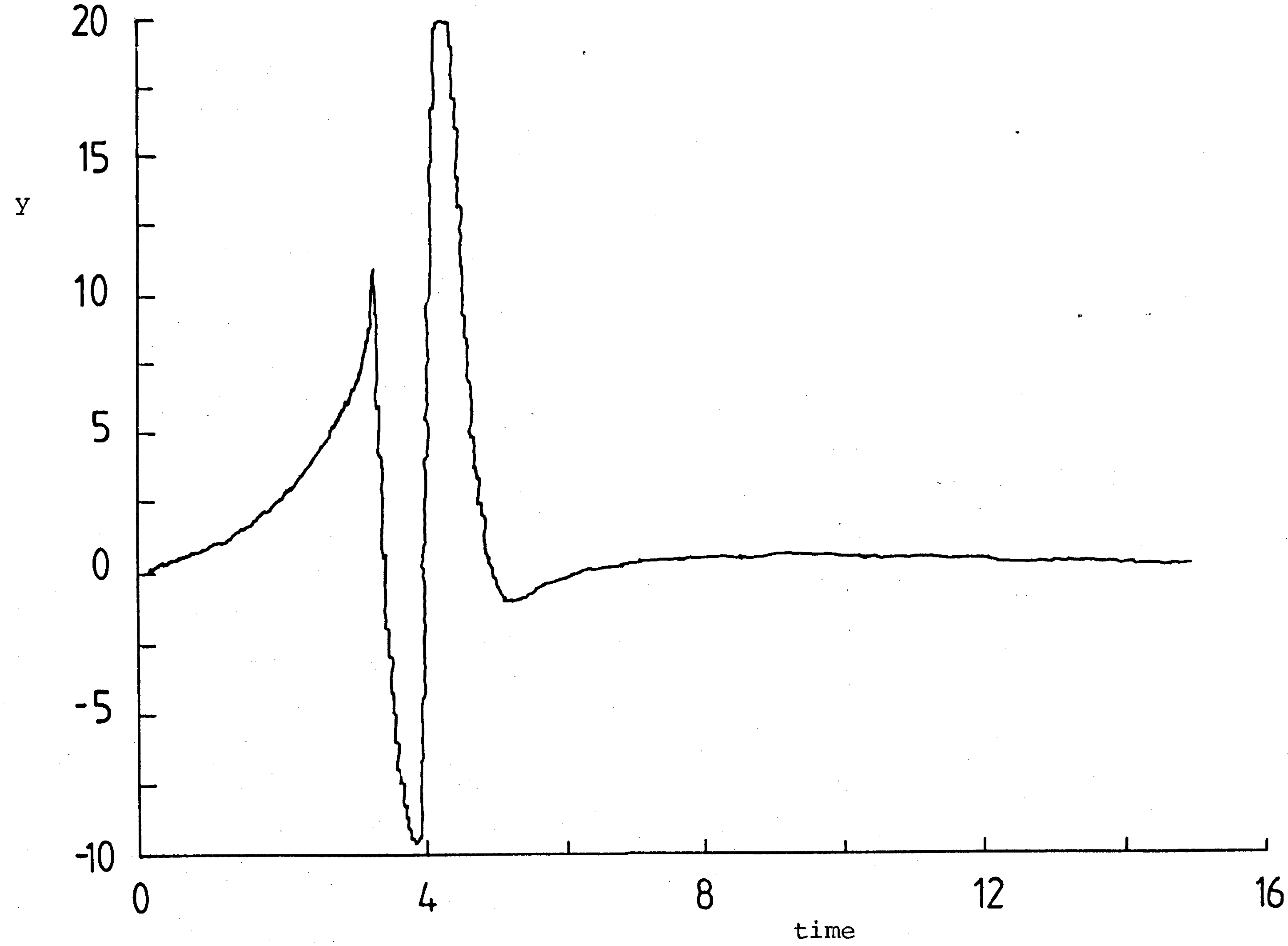


Figure 6.2 Adaptive control with unmodelled modes.

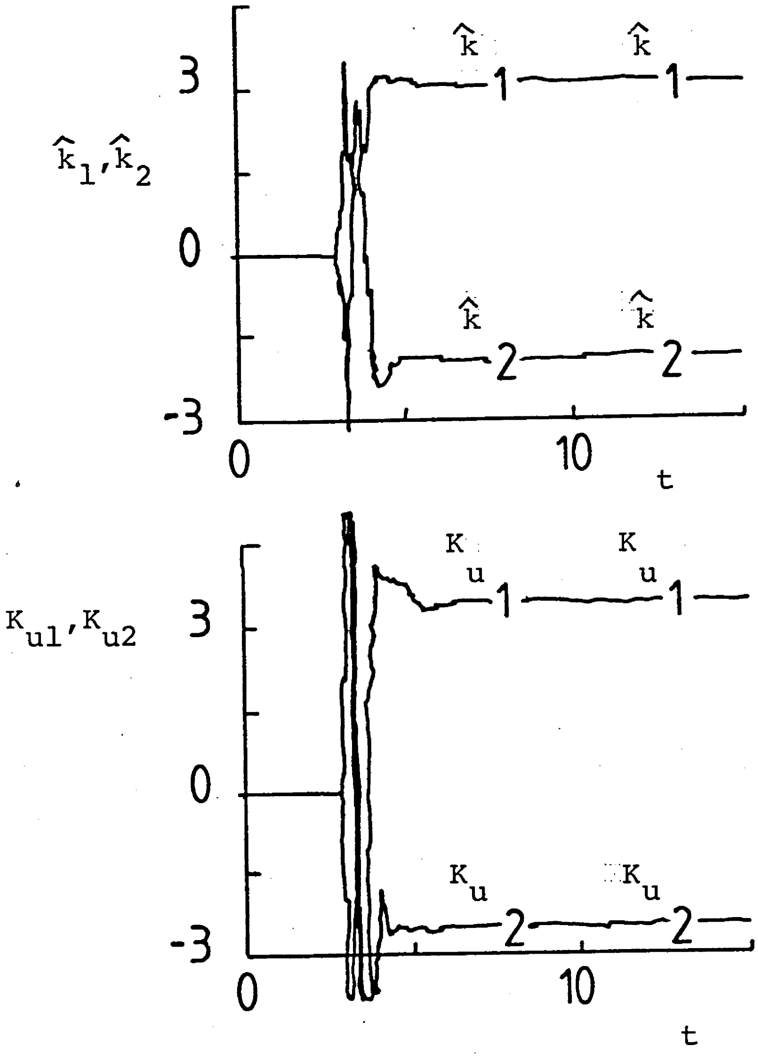


Figure 6.3 Behaviour of constrained and unconstrained estimates in adaptive control with unmodelled modes.

References for Chapter 6

1. B. Egardt and C. Samson,"Stable adaptive control of nonminimum phase systems",Systems & Control Letters, vol 2, pp137-144, Oct 1982.

2. G. Kreisselmeier, "An approach to stable indirect adaptive control", Submitted for publication.

3. A.S. Morse, "Global stability of parameter adaptive control systems", IEEE Trans Auto Cont, AC-25, pp 433-439,June 1980.

4. K.S. Narendra, Y.H. Lin, L.S. Valavani, "Stable adaptive controller design - Part II: Proof of stability", IEEE Trans Auto Cont , AC-25, pp 440-448, June 1980.

5. G.C. Goodwin, P.J. Ramadge and P.E. Caines, "Discrete-time multivariable adaptive control", IEEE Trans Auto Cont, AC-25 pp 449-456, June 1980.

6. G.C. Goodwin, C.R. Johnson, Jr. and K.S. Sin, "Global convergence for adaptive one step ahead optimal controllers based on input matching", IEEE Trans Auto Cont, AC-26, pp 1269-1273, Dec. 1981.

7. B.D.O. Anderson, "Adaptive systems, lack of persistence of excitation and bursting phenomena". Submitted to Automatica.

8. B.D.O. Anderson and R.M. Johnstone, "When will adaptive systems really adapt? The robustness issue." Proc. 2nd Conference on Control Eng, IE Australia, 1982, pp 59-66.

9. S. Dasgupta, B.D.O. Anderson, G. Kreisselmeier and R.J. Kaye "Adaptive control of systems with unknown physical element values", Under preparation.

10. B.D.O. Anderson, "Exponential stability of linear equations arising in adaptive identification", IEEE Trans Auto Contr, AC-22 , Feb 1977, pp 83-88.

11. N.N. Krasovskii, Stability of motion,Stanford University Press, 1963.

§7 Conclusion

This thesis considers the adaptive identification and control of single input - single output, continuous time systems whose unknown parameters are directly related to physical parameter values.

At the outset it is shown that in many cases such a parametrisation leads to state variable realizations which have a rank-1 dependence on the unknown parameters. This in turn results in minimal transfer function descriptions, whose denominator and numerator polynomials are multilinear in the unknown parameters. The adaptive algorithms formulated here exploit this multilinearity.

General tools for deriving persistence of excitation conditions on the inputs of the unknown system, are developed. The results derived include those involving unstable systems with bounded or unbounded inputs and time varying systems with bounded time variations and system signals. These tools are subsequently applied, in specialized forms, to derive p.e. conditions which guarantee the convergence of the algorithms formulated here.

Both equation and output error identification are considered. Of the three equation error algorithms, the first two, respectively called the two step least squares and generalized gradient descent algorithm, involve a two step procedure. The first step generates an unconstrained parameter estimate by ignoring the intrinsic multilinearity. The second obtains a constrained estimate which is in some sense, the "nearest" to this unconstrained estimate. The least squares two step algorithm uses a least squares approach in generating the unconstrained estimate; the generalized gradient descent one uses any gradient descent algorithm for the first step. The second step for both is the same, being a steepest descent algorithm minimizing a quadratic of the difference between the constrained and unconstrained estimates.

The least squares algorithm is implementable only when the input satisfies a p.e. condition. Once implementable, it is uniformly asymptotically convergent as long as the parameter estimates are initialized to zero. The other two are u.a.s. as long as the input is p.e. and the parameter estimates never leave the orthant containing the true parameter value. Several ways of ensuring the latter are presented and involve the knowledge of parameter magnitude bounds. Given the physical significance of the unknown parameters such bounds can be easily ascertained.

Simulation results are presented, primarily to show that the second step in the aforementioned two step approach, serves to improve the accuracy of the parameter estimates.

Two output error algorithsm are formulated, both involving the two step structure outlined above. They are shown to be uniformly asymptotically stable whenever the input is p.e., a certain transfer function involving the unknown parameters is strictly positive real (SPR) and the parameter magnitude bounds are known. The convergence proof utilizes a lemma which extends a result derived by Boyd and Sastry [1] , to the unbounded signal case.

Two indirect adaptive controllers are formulated. Both involve a general controller, satisfying certain assumptions, but differ in the identifiers employed. The first is shown to be stable under the assumption of a stabilizable and detectable plant. In the second the knowledge of a convex region containing the true parameter value is assumed. This region has the following added property. Suppose the plant is described by $P(\bar{k})$ and the controller by $C(\bar{k})$, $\bar{k}$ being the parameter estimate. Then the closed loop system is asymptotically stable whenever $\bar{k}$ lies in this region. In both the algorithms uniform asymptotic convergence of parameter estimates is established under the assumption of p.e. reference inputs.

7.1 Areas of further work

Several open questions remain.

(i) We have not established the multivariable extension of the result which states that SISO systems with a rank-1 dependence on the system parameters have minimal transfer function descriptions, whose numerator and denominator polynomials are multilinear in the parameters.

(ii) For systems having more than one parameter, the problem of relating transfer functions with the multilinear property, to rank-1 state variable realizations has not been addressed.

(iii) In the two step algorithms, simulation results show a marked improvement in the parameter estimates as a result of the second step. A comprehensive theoretical explanation of this improvement has not been given.

(iv) The least squares two step algorithm is implementable only when the input satisfies a p.e. condition. Non-satisfaction of this condition, however, is not always tantamount to lack of system identifiability. Thus modifications of this algorithm to cope with such a situation is desirable. Here, we have only considered modifications for a restricted subclass of these situations.

(v) Output error convergence requires that a polynomial $\beta(s)$ be known such that $\beta(s)/Q(s,k)$ is SPR , with $Q(s,k)$ the denominator polynomial of the system transfer function. The determination of a $\beta(s)$ which satisfies this for all k belonging to a known convex region is a worthwhile problem to solve.

Reference:

1. S. Boyd and S.Sastry, "On parameter convergence in adaptive control",Systems and Control Letters, pp 311-319, Dec 1983.

ADDENDUM

"ADAPTIVE IDENTIFICATION AND CONTROL"
S. Dasgupta

CORRECTIONS

Page 1 — Equation (1.1): $y^{(n)}(t)$ as opposed to $y^n(t)$.

Page 24 — Equation (2.6): The second $y_1(s)$ should read $u_1(s)$.

Page 25 — First formula should read:

$$\frac{\partial H(s,k_1)}{\partial k_1} = \frac{h_{12}(s)h_{21}^T(s)}{(1+k_1h_{22}(s))^2}$$

Page 33 — After the last line add:

$$D_3 = \begin{vmatrix} -\bar{M}_{21} \\ -\bar{M}_{31} \end{vmatrix} \; ; \quad D_4 = \begin{vmatrix} \bar{M}_{24} & \bar{M}_{25} \\ \bar{M}_{34} & \bar{M}_{35} \end{vmatrix} \begin{vmatrix} \Lambda_{C2} & 0 \\ 0 & \Lambda_{L2} \end{vmatrix} \; ; \quad D_5 = \begin{vmatrix} \bar{M}_{14} & \bar{M}_{15} \end{vmatrix}^T$$

Page 39 — Line after equation (2.15): α to be changed to α_1. Following line: α to be changed to a.

Pages 43/44 — Page 44 should come before Page 43.

Page 88 — Page 88a is now included after this page.

Page 103 — Last equation:

$$||x(t)||$$

Page 144 — Page 144 is now added.

Page 153 — Equation (4.20) should read:

$$\dot{\hat{k}}(t) = -\frac{\partial \hat{K}^T(t)}{\partial \hat{k}} \Lambda \, [R(t)^{-1}r_0(t) + \hat{K}(t)] \; ; \quad t \geqq t_0 \qquad (4.20)$$

Page 154 — Equation (4.25) should read:

$$K = -R(t)^{-1}r_0(t) \; , \quad \text{for } t \geqq t_0 \qquad (4.25)$$

Page 179 — Equations (4.59) and (4.60): > should be replaced by ≥.

Page 236 — Line 3 from bottom: Delete "non".

Page 237 — Line above Equation (5.3) should read:

$$\frac{p_r(s)}{\beta(s)} u(t) - \frac{q_r(s)}{\beta(s)} \hat{y}(t) \; , \; \forall \; r \subset S. \quad \text{Then}$$