

# DYNAMICS OF ADAPTIVE CONTROL

Iven M.Y. Mareels

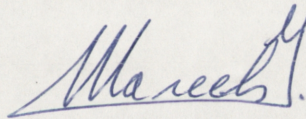
August 1986

A thesis submitted for the degree of  
Doctor of Philosophy  
of The Australian National University.

The doctoral studies were conducted under the supervision of  
Dr. R.R. Bitmead and Prof. Dr. B.D.O. Anderson.

I hereby certify that the research reported in this thesis, unless explicitly  
and otherwise stated in the text, is based on personal and original work.

Canberra, August 1986

A handwritten signature in blue ink, appearing to read 'Mareels' with a stylized flourish at the end.

Iven M.Y. Mareels

Department of Systems Engineering

Research School of Physical Sciences

The Australian National University

G.P.O. Box 4

A.C.T. 2601

Canberra

Australia

"The fear of the Lord is the beginning of all knowledge"

Proverbs 1,7

## ACKNOWLEDGMENTS

In retrospect it seems hard to do justice to all those who have contributed in a stimulating way to make this thesis reality. But in the spirit of "a drop of wisdom is worth an ocean of knowledge", I want to thank first those who supported me in spirit and then my teachers whose keen minds I have been allowed to analyse. Of all those, this thesis is an integrated reflection, I hope it honours them.

My family, parents, brother and sister have always been an invaluable support to me; although we are now separated by 20,000 km, we are one; thanking them appropriately is impossible. My dear wife, Sylvia, who encouraged me unselfishly and no doubt must feel married to a man who is married to his study, cannot be thanked with words.

My study time has been a pleasure, due to my excellent teachers. Especially, to Bob (Dr R.R. Bitmead) whose visionary ideas put me on the subject of the dynamics of adaptive control, I would like to pay tribute. Prof. Brian Anderson, my co-supervisor, whose energetic enthusiasm has been most stimulating, made it possible for me to meet and to work with some of the best researchers in the systems engineering community. Working with both my supervisors was a delight. Realising I am not capable of righteously thanking all those who gave me the opportunity to learn from their methods and to gain insight from their knowledge, I will simply list them alphabetically, in a modest recognition of their excellence: Dr R.K. Boel, Dr M. Gevers, Dr C.R. Johnson Jr, Dr P.V. Kokotovic, Dr R.L. Kosut, Dr J.B. Moore, Dr E. Noldus, Dr V. Kučera, Dr S.S. Sastry, Dr J.L. Willems, Dr D. Williamson. My fellow students, especially M. Green, have been most helpful with their inherent criticism and sharp minds, they are highly appreciated.

I thank the Australian National University and the National Fund for Scientific Research Belgium, for financial support.

Heather Wood deserves the best of praise for the excellent work she did in typing the manuscript.

In my life I have always been guided by the Bible, and I believe that the



following scriptures are particularly relevant in the present situation:

"For in much wisdom is much grief, and he that increaseth knowledge,  
increaseth sorrow"

Ecclesiastes: 1,18.

"My fruit (I=knowledge) is better than gold, yea than fine gold, and my  
revenue than choice silver"

Proverbs: 8,19.

In my studies I have been able to experience both, to all those who made this  
possible: Thank you.

August, 1986

A handwritten signature in black ink, appearing to read 'Mareels', with a stylized flourish at the end.

Iven Mareels

## PREFACE

The material covered in this thesis reports only part of the research with which I have been involved during my postgraduate studies. Although it would have been possible - at the expense of a considerable amount of work - to produce a magnum opus containing the whole body of research in one coherent work, I opt to present only this material which discusses the dynamics of adaptive control from the parameter space point of view. In doing so, I believe I do both the reader and myself a service.

The research not presented in this thesis concerns:

1. The application and extension of averaging techniques and singular perturbation methods to determine the stability properties of adaptive systems. This research is presented in detail in [1,2,3]. Some of these techniques are used in Chapter 2.
2. The characteristics of persistency of excitation and sufficiently rich inputs for linear time invariant systems in the light of its necessity for establishing exponential stability of adaptive systems, as well as a discussion of the robustness of persistency of excitation with respect to time variations in the dynamics of the plant. This research is reported in [1,2,4,5,6,7,8]. The notion of persistency of excitation is used throughout the thesis.
3. The study of modifications of basic adaptive algorithms and their effects on the stability properties and our perception of sufficiently rich input signals. This has been reported in [9,10,11].
4. An initial attempt to deal with the transient behaviour of adaptive algorithms in a more general way. This is presented in [12].

The material treated in the thesis is an extension and/or a more detailed presentation of research presented in [13-20].

## References

- [1] B.D.O. Anderson, R.R. Bitmead, C.R. Johnson Jr., P.V. Kokotovic, R.L. Kosut, I.M.Y. Mareels, L. Praly, B.D. Riedle, 'Stability of Adaptive Systems: Averaging and Passivity Analysis', MIT Press, Boston, 1986.
- [2] R.L. Kosut, B.D.O. Anderson, I.M.Y. Mareels, 'Stability theory for adaptive systems: methods of averaging and persistency of excitation', Accepted for publication in IEEE Trans Automatic Control, 1986. also in Proc. of 24th IEEE Conf. on Dec. and Cont. Ft. Lauderdale, Florida, Dec. 11-13, 1985, pp. 478-484.
- [3] M. Bodson, S.S. Sastry, B.D.O. Anderson, I.M.Y. Mareels, R.R. Bitmead, 'Nonlinear Averaging Theorems, and the Determination of Parameter Convergence Rates in Adaptive Control', Systems & Control Letters, Vol. 7, No. 3, pp. 145-159.
- [4] I.M.Y. Mareels, M. Gevers, 'Persistency of excitation criteria', Proc. 25th IEEE Conf. on Decision and Control, Athens, Dec. 1986, to appear.
- [5] I.M.Y. Mareels, M. Gevers, 'Persistency of excitation criteria for linear multivariable time-varying systems', IEEE Trans. on Auto. Control, submitted for publication.
- [6] I.M.Y. Mareels, R.R. Bitmead, M. Gevers, C.R. Johnson, R.L. Kosut and M.A. Poubelle, 'How exciting can a signal really be?', Systems and Control Letters, to appear
- [7] I.M.Y. Mareels, 'Sufficiency of excitation', Systems and Control Letters, Vol.5, 1984, pp.159-163.
- [8] W.A. Sethares, I.M.Y. Mareels, B.D.O. Anderson, C.R. Johnson Jr. and R.R. Bitmead, 'Excitation conditions for convergence of signed regressor LMS', IEEE Trans. on Circuits and Systems, submitted for publication.
- [9] I.M.Y. Mareels, 'A Simple Selftuning Controller for Stably Invertible Systems', Systems&Control Letters, Vol 4, No1, Feb 1984, pp.5-17.
- [10] I.M.Y. Mareels, 'Parameter Estimation by the Least Squares method', Journal A, Vol 25, No2, pp.99-104, 1984.

- [11] B.D.O. Anderson, I.M.Y. Mareels, W.A. Sethares, C.R. Johnson Jr, 'Signed Regressor, Signed Error Adaptation and Averaging', Submitted to IEEE Trans. Circuits and Systems.
- [12] R.L. Kosut, I.M.Y. Mareels, B.D.O. Anderson, R.R. Bitmead, C.R. Johnson, 'Transient Analysis for Adaptive Systems', submitted to 10th IFAC World Congress, Munich, 1986.
- [13] I.M.Y. Mareels, B.D.O. Anderson, R.R. Bitmead, M. Bodson, S.S. Sastry, 'Revisiting the Mit-rule for Adaptive Control', Proc. of the 2nd IFAC workshop on Adaptive Systems in Control and Signal Processing, Lund Sweden), July 1986, pp. 67-72.
- [14] B.D.O. Anderson, I.M.Y. Mareels, 'Marrying Frequency Domain Intuition and Adaptive Control', Proc. of the 3rd Conf. on Contr. Engineering, May 1986, Sydney Australia, pp.66-72
- [15] I.M.Y. Mareels, R.R. Bitmead, 'Bifurcation Effects in Adaptive Control', submitted to 10th IFAC World Congress, Munich, 1986.
- [16] I.M.Y. Mareels, R.R. Bitmead, 'Nonlinear Dynamics in Adaptive Control: Chaotic and Periodic Stabilization', Automatica (to appear 1986), also Proc. of the 2nd IFAC workshop on Adaptive Systems in Control and Signal Processing, Lund (Sweden), July 1986, pp. 245-250
- [17] R.R. Bitmead, I.M.Y. Mareels, 'Nonlinear Dynamics of Adaptive Control', Proc. of the 3rd Conf. on Contr. Engineering, Sydney Australia, May 1986, pp.35-39
- [18] I.M.Y. Mareels, R.R. Bitmead, 'On the Dynamics of an Equation Arising in Adaptive Control', submitted to Automatica, also Proc. 25th IEEE Conf. on Decision and Control, Athens, Dec. 1986, to appear.
- [19] A.C. Tsoi, I.M.Y. Mareels, R.R. Bitmead, 'Robust Adaptive Control of Model Mismatched Systems', Proc. of the 3rd Conf. on Contr. Engineering, Sydney Australia, May 1986, pp. 81-87
- [20] I.M.Y. Mareels, 'The M.I.T.-Rule Revisited', submitted to IEEE Student Paper Contest, Region 10, 20 February 1986.



## Table of Contents

### 1. Introduction

1.1 Adaptive Control.....	1
1.2 Feasibility of Adaptive Control.....	3
1.3 Robustness of Adaptive Control.....	4
1.4 Case Studies in the Dynamics of Adaptive Control.....	7
1.5 Our Contribution: Nonlinear Theory for Adaptive Control.....	8
1.6 Organisation of the Material.....	9
1.7 References.....	11

### 2. The M.I.T. Rule For Adaptive Control

2.1 Introduction.....	13
2.2 The M.I.T. Rule.....	13
2.3 Instability Mechanisms.....	18
2.4 Stability Analysis via Averaging.....	26
2.5 Generalizations.....	38
2.6 Historical Overview of the M.I.T. Rule.....	39
2.7 Appendix: Averaging Theorems for Linear Timevarying Systems...	41
2.8 References.....	54

### 3. Bifurcations In Adaptive Model Reference Control

3.1 Introduction.....	56
3.2 Problem Description.....	57
3.3 Basic Properties of the Adaptive Response.....	60
3.4 Some Notions from Global Analysis.....	63
3.5 Adaptive System Dynamics for Constant Reference Input.....	68
3.6 Adaptive System Response for General Reference Input.....	76
3.7 Numerical Experiments.....	85
3.8 Discussion.....	100
3.9 Historical Overview.....	102
3.10 References.....	104

#### 4. Stabilizing Nonlinear Dynamics in Adaptive Control

4.1	Introduction.....	107
4.2	Conventions and Notations.....	109
4.3	Problem Description.....	110
4.4	The Closed Loop Dynamics I; $b = 0$ .....	114
4.5	The Closed Loop Dynamics II; $b < 0$ .....	115
4.6	The Closed Loop Dynamics III; $b > 0$ .....	124
4.7	The Adaptive Control Problem: Discussion.....	142
4.8	Complements.....	150
4.9	Conclusions.....	155
4.10	References.....	156

#### 5. General Discussion

5.1	Introduction.....	158
5.2	What Makes Adaptive Control Adaptive?.....	159
5.3	Dynamics of Adaptive Control ; Global Versus Local Analyses..	161
5.4	Towards a Transient Analysis for Adaptive Control.....	163
5.5	Necessary Conditions for Adaptive Control?.....	164
5.6	Some Final Observations.....	165
5.7	References.....	167

## 1. INTRODUCTION

In this introduction we discuss concisely what we understand by adaptive control, pointing out some of the generic forms in which adaptive control manifests itself. A brief review of the major results about the dynamics of adaptive systems is given, emphasising their importance and their weaknesses. Consequently, we present our approach to obtaining a better understanding of the dynamics of adaptive control. Finally, we outline the organisation of the material covered in this thesis and describe what we see as our contribution to the field of adaptive control.

### 1.1 Adaptive Control

Adaptive control is one approach to control imprecisely known systems. Methods developed to control uncertain systems can be distinguished on the basis of the a priori knowledge assumed about the plant. (In the absence of any information about the system to be controlled, any controller is a priori as good or as bad as an adaptive one.) In most nonadaptive methods, the plant is assumed to be one fixed member of some class  $P$  of causal, time invariant, linear systems. A stabilizing controller designed for one particular fully known system may stabilize a neighbourhood of this nominal system. The larger this class, the more "robust" the controller. Robust control design tries to optimize in some sense this neighbourhood around a nominal plant. Adaptive control has an underlying assumption that the plant at any instant in time belongs to some class  $P$  (generally much larger than just a "neighbourhood" of a nominal system), and can vary "slowly" throughout this class. The aim is then to find a universal controller which is able to stabilize any fixed member in the given class and can track the system. One approach consists of parametrizing the class  $P$  and a corresponding class  $C$  of stabilizing controllers (designed as if the system were known) and to control the system by identifying from the observed input output data which controller of  $C$  to use. If the identification procedure is a two-step method, first identifying the plant (i.e. the parameter corresponding to the plant) and then obtaining from this parameter the controller, one speaks of indirect adaptive control. In direct adaptive control, one tries to identify the controller immediately from the input, output data. If the adaptive algorithm is not based on a coupled parametrization of the classes

P and C, one speaks of non parametric adaptive control.

This automation of the identify-then-control approach, characteristic of classical linear control, underlies most of the adaptive control algorithms. This methodology leads to an overall system consisting of three distinct subsystems:

- (a) the plant to be controlled, which the control designer assumes to belong to the model set, for future reference, represented as

$$y(t) = [H(p)u](t); \quad p \in P \quad (1.1.1)$$

where  $u$  is the (plant) input function,  $y$  the (plant) output,  $H(p)$  the model operator, parametrized by a vector  $p$ , belonging to the set  $P$  characterizing the model set under consideration. (The actual plant is not necessarily a member of the model set!)

- (b) the control law - designed for plants belonging to the model set:

$$u(t) = [G(\theta(p), d_1)(y, r)](t) \quad \theta \in \Theta; \quad d_1 \in D_1 \quad (1.1.2)$$

$G$  is the causal operator (depending on the parameter  $\theta$ , which depend themselves on the parameter  $p$ , and  $d_1$  is a vector of design variables), which from the output measurements of the actual plant output  $y$ , and the reference signal  $r$  (setpoints,...) generates the input to the plant.

- (c) the identification mechanism (adapted to both the model set and the controller structure):

$$\theta(t) = [I(d_2)(u, y, r)](t); \quad d_2 \in D_2 \quad (1.1.3)$$

$I$  is a causal, typically nonlinear operator with memory, producing the control parameters  $\theta$  on the basis of the available data, i.e. input  $u$ , output  $y$  and reference signal  $r$  up to time  $t$ .  $d_2$  is a vector of design variables, (stepsize, gain, filter poles,...).

Classically one assumes both plant and model to be linear, time invariant and finite dimensional, designs  $G$  accordingly as a linear control law (pole placement, LQG, model matching, minimum variance, ...) and estimates  $\theta$  via an identification scheme suited to identifying a linear, time invariant system, using, for example, a least squares or a stochastic gradient approach. All the ingredients are inherited from linear control theory, and one might venture that the adaptively controlled closed loop system would behave basically in a linear way. This is however not the case: the adaptive closed loop system (plant with control law defined in (1.1.2)-(1.1.3)) is fundamentally nonlinear. It is our aim to demonstrate this unequivocally and to unravel (at least as far as practical) these dynamics.



## 1.2 Feasibility of Adaptive Control

The longstanding problem of demonstrating the existence of a universal controller (designed using the methodology outlined above) for the class P of causal, linear time invariant, single input single output (SISO) systems with known number of poles and zeros in a minimal representation, has only recently been solved, both for the parametric approach [1,2,3,4] and the nonparametric approach [5,6]. Most of the possible combinations of deterministic or stochastic and discrete or continuous time SISO systems have been covered. The case of multivariable systems has not yet been fully settled [1,2,7,8]. The generic form of the results is:

C1: "All signals within the adaptive system are bounded, and asymptotically the control objective is achieved."

The hypotheses, under which this result can be established, involve typically:

H1: "The plant belongs to the model set P (as described above (1.1.1))."

H2: "The plant's transfer function (combined with other transfer functions arising in the adaptive systems) satisfies a strict positive real condition."

This is quite a formidable result, as it is indeed amazing that the methodology outlined above works. However, from a practical point of view and from a dynamical systems point of view this result is unsatisfactory. Firstly, the results say very little about transient response - what does "bounded signals" mean?, and the information about the asymptotic dynamics is weak - no uniform convergence. The nonuniformity of the dynamics implies that small changes in the representation of the dynamics may cause fundamental differences, e.g. loss of stability. The absence of the "uniform" qualifier in the result C1 is not due to a careless analysis, but is unfortunately a property of the adaptive system under the hypothesis H1 and H2. Indeed small unmodelled bounded disturbances - e.g. due to round off errors in a computer simulation, or measurement noise - may cause unacceptably large (but bounded) deviations (called bursting phenomena) from the desired response. The control objective cannot be fulfilled anymore [9]! This sensitivity to unmodelled bounded disturbances can be overcome by requiring that:

H3: "The external signals are persistently exciting (PE) so as to guarantee uniform asymptotic identifiability of the control parameter  $\theta$  [10]."

Indeed, under the extra hypothesis H3, one can demonstrate that in addition to C1, the following conclusion holds:

C2: "The control objective is achieved uniformly asymptotically, and the model parameter/control parameter is uniformly asymptotically identified."

A more fundamental shortcoming is the crucial presence of the strict positive real condition and the requirement to know the system's order. These properties are sensitive to slight variations in the transfer function (in frequency response terms). Because it is very likely that the plant does not belong to the model set, the strict positive real condition is generally found not to be satisfied in practice. Moreover, the examples in [2,11] demonstrate that, even in the presence of PE signals, loss of the strict positive real condition can completely destabilize the adaptive closed loop.

These observations do not invalidate the methodology of adaptive control - witnessed by many successful applications of adaptive controllers - but do demonstrate that these results C1 and C2 obtained under the hypotheses H1, H2 and H3 (although of fundamental importance to establish the feasibility of adaptive control), do not provide the necessary information in order to understand the dynamics of adaptive control, nor to make adaptive control work under more realistic conditions.

### 1.3 Robustness and Adaptive Control

A "robust" or "structurally stable" system is a system which preserves some of its qualitative properties under (small) perturbations in its description. The preserved qualitative properties are called structurally stable or, concisely, robust. Our notion of robustness therefore depends on the properties we are interested in and on the perturbations we allow for, in particular on how we quantify (small) perturbations.

In an adaptive control context it is realistic and relevant to consider robustness with respect to the perturbation undermodelling: "the plant does not belong to the model set" (which requires the introduction of an appropriate

notion of smallness). This corresponds to violating both hypotheses H1 and H2. Considering perturbations which cause H3 to be violated, is less realistic, as under normal circumstances we always can enforce this hypothesis and because condition H3 is robust with respect to small signal deviations by construction. The qualitative property of interest is the control performance: only small (asymptotic) deviations from the desired response are acceptable. It is desirable to include transient response considerations in our measure of the control performances, and whenever this is amenable for analysis we will do so.

The property "belongs to the model set P (the class of causal, linear, time-invariant systems with  $m$  zeros and  $n$  poles) and satisfies a strict positive real condition" is not robust with respect to the most commonly used notion of "closely resembling systems". Indeed, in engineering terms, it is natural to regard (stable) systems as close whenever their Bode diagrams are close over the relevant frequency range. (A similar notion of closeness can be defined for unstable systems, using matrix fraction description with stable factors [12].) But in this measure systems of arbitrary order can be arbitrarily close and strict positive realness is completely out of the picture! In linear, classical robust control design ( $H_\infty$ -theory) this is the relevant measure, as systems which are close in the above sense for open loop response will be close for the closed loop response [12,13]. Ideally, we would like our adaptive control algorithm to have this property (if the plants to be controlled are "close", then the adaptive response is "close"). As was pointed out in the previous section, this is not the case for the adaptive system (1.1.1-1.1.3) under the hypotheses H1-H3. Therefore we would like to specify a range of design variables, or specify extra information about the plant such as to guarantee this type of robustness. This is precisely where the "global" results C1 and C2 fall short, the role and the interrelation of the design parameters and external driving signals is largely neglected.

Very useful information about the role of the design variables and the excitation of the external signals or the stability and control performances of adaptive control can be obtained by considering only slow adaptation [14]. The assumption of slow adaptation enforces a time scale separation. The dynamics of the controlled plant (1.1.1-1.1.2) are fast compared with the slowly time varying  $\theta$  governed by (1.1.3). Therefore we can analyse the controlled plant (1.1.1-1.1.2) separately from the update law (1.1.1). On the time scale of the

subsystem (1.1.1-1.1.2) we fix  $\theta$  and analyse the corresponding (by assumption and construction) linear loop, whilst we analyse the identification scheme by considering  $y$  and  $u$  in a regime for the particular  $\theta$  considered. The results yield meaningful design criteria (specifying filter poles, frequency content of external signals, adaptive gain) for good local performance of the adaptive closed loop. The local nature of the results is a direct consequence of the approach. In order to have a regime for  $u$  and  $y$ ,  $\theta$  should at all times specify a stabilizing controller (1.1.2) and the time scale separation, (uniform in time) requires that the encountered regimes of  $u$  and  $y$  are close (in some sense) to the desired behaviour so as to make the driving forces in (1.1.3) small and therefore  $\theta$  slowly time varying. The theory demonstrates that the local dynamics of adaptive control are robust with respect to undermodelling, provided some intuitively appealing design criteria are met (design criteria which are insensitive with respect to modelling errors, of course).

As opposed to this local approach, there exists a tendency to develop global "robust" adaptive control results by redesigning the adaptive law. This approach is initiated in [15], where the effect of unmodelled fast dynamics in the plant on the adaptive response is investigated. The modifications introduced in the adaptive mechanism usually only concern the identification mechanism (rather strange!) and appear as ad hoc "fixes", e.g. normalization, dead zones, exponential forgetting, data-windowing [16,17,18]. The theory is basically nonquantitative and concentrates on bounded input bounded state stability of the adaptive closed loop - often without quantifying the gain from input to state magnitudes. It is our opinion that this is not enough, and we will demonstrate that this coarse theory overlooks some rather remarkable, undesirable effects (such as extreme sensitivity to initial conditions) which can be present if one does not pay careful attention to the selection of the design parameters. This aspect of the adaptive control design has largely been neglected in the global approach.

It is natural that in the search for robustness we have to introduce more information about the plant and its environment and that we pursue a less ambitious control task. In both approaches, discussed above, the balance between the extra information (respectively in the form of design criteria and "fixes") and the relaxed control objective (small deviations from desired response (output property) and a bounded state response) is introduced on an ad hoc



basis.

The apparent discrepancy (local  $\leftrightarrow$  global, good performance  $\leftrightarrow$  bounded errors, parameter selection  $\leftrightarrow$  fixes) clearly indicates a gap in our understanding of the dynamics of adaptive control. We aim to fill this gap at least partially in order to understand what extra information is necessary and how much we need to relax the control task. In doing so, we indicate how the local theory breaks down and what the coarser global theory overlooks, and demonstrate that the dynamical behaviour that adaptive control can exhibit is only limited by the imagination of the beholder.

#### 1.4 Case Studies in the Dynamics of Adaptive Control

We propose to study the dynamics (and therefore the robustness properties) of adaptive control along the following lines. In designing the adaptive controller, we first decide upon a model set  $P$  and construct the controller  $G$  (1.1.2) and the identification scheme  $I$  (1.1.3) accordingly, and then analyse the closed loop when the actual plant does not belong to  $P$ , but to some other set  $P'$ , not necessarily a superset of  $P$ . This corresponds to a real world scenario. The task is then to classify (on the basis of qualitative differences) the possible (global) dynamics over the whole parameter space  $P' \times D$  - including both plant and design parameters. Formulated in such generality this is an immensely difficult task, unlikely to be ever resolved in any great detail. Therefore, and in the spirit of global analysis and bifurcation theory - mathematical disciplines concerned with the classification of the qualitative behaviour of dynamical systems - we present an inductive approach to the problem formulated above based on three fairly complete case studies.

These case studies are "simple" examples along the lines of the above scenario, allowing a detailed analysis. At all times we will be interested in structurally stable phenomena and in the mechanisms which cause them, enabling us to make more general statements.

In Chapter 2 we discuss the M.I.T.-Rule for adaptive feedforward control of a linear time invariant plant. This leads to a linear (in the state) overall system, which simplifies the analysis. In this situation the "local" theory described above provides global results and therefore this example serves to illustrate how the design parameters effect the dynamics. As the linearity restricts the possible dynamics severely, the main aim is to describe the

stability-instability boundary in the parameter space. We identify the different mechanisms which cause instability (high gain, resonance and undermodelling effects) and indicate how to select the design variables to avoid these problems. Some comments are made about the robustness of the adaptively controlled system with respect to other non modelled effects such as exogenous disturbances (measurement noise) and small nonlinearities in the plant.

Chapter 3 deals with a model reference adaptive control algorithm. The plant to be controlled belongs to the class  $P'$  of causal, linear, time invariant second order systems without finite zeros, whilst the controller is designed for systems belonging to the class of  $P$  of causal, linear, time invariant first order systems without finite zeros. The resulting nonlinear system exhibits a large variety of different dynamics for various parameters ranging from uniform asymptotic stability to strange attractors via a sequence of period doubling bifurcations initiated by a Hopf bifurcation. Special attention is paid to the implications these findings have on our understanding of adaptive control in general. Most importantly we argue that it is of crucial importance to select the design variables properly in order to avoid some of these complicated dynamics, more so than to modify the adaptive control law.

A modified version of the selftuning minimum variance regulator is analysed in Chapter 4. Special emphasis is placed on its robustness properties with respect to undermodelling by considering a second order plant controlled by an adaptive law designed for a first order system. A detailed analysis reveals that for a large class of second order plants output regulation is achieved, whilst some internal signals behave either periodically or chaotically depending on the parameters describing the plant. A discussion of the effects of time variations in the dynamics of the plant to be controlled as well as of other non-modelled disturbances is included. This is the only problem where we discuss transient behaviour in detail. As in the previous chapters the implications that this analysis has on adaptive control in general are pointed out.

### 1.5 Our Contribution: Nonlinear Theory for Adaptive Control

The present work is an attempt to understand the dynamics of adaptive control, and to come to grips with the fundamentally nonlinear mechanisms that govern its properties.

It is demonstrated that contrary to common belief, an adaptive controller

(designed along the methodology outlined in section 1.2) only exhibits its linear heritage for slow adaptation, which requires both small adaptive gain and small deviations from desired behaviour. Therefore one should resist the temptation to view and to analyse adaptive control as a linear time-varying control. Specifically robust adaptive control is not just a combination of robust (linear) control and robust (linear) identification. The robustness mechanisms (in the large, not slow adaptation) are essentially nonlinear, and are reminiscent of the robustness properties of nonlinear oscillators, as for example, the Van der Pol oscillator.

The first case study, the M.I.T. rule, demonstrates the efficacy of the perturbation tools. In this instance, the local theory is globally valid.

The next two case studies indicate how the local theory may break down and demonstrate what the global theory appears to overlook. In particular we argue that it is not sufficient only to include boundedness considerations, and that it is an unsound idea to introduce fixes disconnected from the control objective. Global and local theories should therefore be combined - an attempt in this direction can be found in [19].

These case studies lead us to some profound questions into the nature of adaptive control. To what extent is an adaptive controller different from any robust nonlinear controller? Certainly the boundary between parametric, nonparametric and direct, indirect adaptive control, as seen from a dynamical system point of view appears to be extremely vague, and in our opinion serves only an historical purpose.

The two last case studies indicate that adaptive control can lead to extremely complex dynamical behaviour (chaotic dynamics). In our opinion, this is also the first instance where chaotic dynamics and their overpowering structural stability serves a useful purpose - it makes these adaptive controllers work!

### 1.6 Organisation of the Material

Each of the Chapters 2, 3 and 4, containing the bulk of the presented material, has its own introduction where the problem is set up, followed by the analysis and conclusions. An historical overview finishes each chapter. It describes the results, available in the literature, about the particular adaptive law (and its variants) considered in that chapter and situates our results. The

references are located at the end of each chapter.

In each chapter, the important equations are numbered per section, separate from the lemmas and theorems and points of discussion. The latter are numbered as Lemma  $n.m$ , Theorem  $n.m$  and R. $n.m$ , where  $n.m$  stands for the  $m$ th Lemma, Theorem or Remark in the  $n$ th Chapter.



1.7 References

- [1] G.C. Goodwin, K.S. Sin, 'Adaptive Filtering Prediction and Control', Prentice Hall, Englewood Cliffs, N.J. 1984.
- [2] B. Egardt, 'Stability of Adaptive Controllers', Springer Verlag, 1979, Lecture Notes in Control and Information Sciences.
- [3] A.S. Morse, 'Global stability of parameter adaptive control systems', IEEE Trans. Auto Contr., Vol.AC-25, No.3, June 1980, pp. 433-440.
- [4] K.S. Narendra, Y.H. Lin, L. Valavani, 'Stable adaptive controller design, Part II: Proof of stability', IEEE Trans. Auto Contr., Vol.AC-25, No.3, June 1980, pp.440-449.
- [5] R.D. Nussbaum, 'Some remarks on a conjecture in parameter adaptive control', System and Control Letters, Vol.3, Nov.1983, pp.243-246.
- [6] B. Mårtensson, 'The order of a stabilizing regulator is sufficient a priori information for adaptive stabilization', System and Control Letters, Vol.6, July 1985, pp.87-91.
- [7] G.C. Goodwin, P.J. Ramadge, P.E. Caines, 'Discrete time multivariable adaptive control', IEEE Trans. Auto. Contr., Vol.AC-25, No.3, June 1980, pp.449-456.
- [8] H. Elliot, W.A. Wolovich, 'A parameter adaptive control structure for linear multivariable systems', IEEE Trans. Auto. Contr., Vol.Ac-27, No.2, April 1982, pp.340-352.
- [9] B.D.O. Anderson, 'Adaptive systems, lack of persistency of excitation and bursting phenomena', Automatica, Vol.21, No.3, May 1985, pp.247-259.
- [10] B.D.O. Anderson, 'Exponential stability of linear equations, arising in adaptive control', IEEE Trans. Auto. Contr., AC-22, Feb. 1977, pp.83-88.
- [11] C. Rohrs, L. Valavani, M. Athans, G. Stein, 'Robustness of continuous time adaptive control algorithms in the presence of unmodelled dynamics', IEEE Trans.Auto. Contr., AC-30, Sept.1985, pp.881-889.

- [12] M. Vidyasagar, 'Control Systems Synthesis: A Factorization Approach', MIT Press, Boston, MA., 1985.
- [13] H. Kwakernaak, 'Minimax frequency domain performance and robustness optimization of linear feedback systems', IEEE Trans. Auto. Contr., Vol.AC-30, No.10, Oct.1985, pp.994-1004.
- [14] B.D.O. Anderson, R.R. Bitmead, C.R. Johnson Jr, P.V. Kokotovic, R.L. Kosut, I.M.Y. Mareels, L. Praly, B.D. Riedle, 'Stability of Adaptive Systems: Passivity and Averaging Analysis', M.I.T. Press, Boston, MA., 1986.
- [15] P. Ioannou, P.V. Kokotovic, 'Adaptive Systems with Reduced Models', Lecture Notes In Control and Information Sciences, Springer-Verlag, 1983, No.47.
- [16] B.B. Peterson, K.S. Narendra, 'Bounded error adaptive control', IEEE Trans. Auto. Contr., Vol.AC-27, No.6, December 1982, pp.1161-1169.
- [17] G. Kreisselmeier, 'An approach to stable indirect adaptive control', Automatica, Vol.21, No.4, 1985, pp.425-431.
- [18] C.R. Johnstone, 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.
- [19] G.C. Goodwin, D.J. Hill, D.Q. Mayne, R.H. Middleton, 'Adaptive robust control, (Convergence, stability and performance)', Technical Report EE8544, Newcastle University.

## 2. THE M.I.T. RULE FOR ADAPTIVE CONTROL

### 2.1 Introduction

The first case study is chosen to highlight both the influence the design parameters have on the dynamics of an adaptive control scheme and the nonlinear way in which they exert this influence. The M.I.T. rule for adaptive feedforward control is particularly well suited for this purpose as its overall description (e.g. equations 1.1.1-1.1.3) can be presented in the form of a linear, time varying system. The linearity of the dynamics (in state space) allows us to concentrate on the property of asymptotic stability, which can be readily interpreted in terms of good control performance, and specially on how it depends on the design parameters. An equally well suited alternative to the M.I.T. rule is a model reference control algorithm for feedforward control. We discuss this approach only when its behaviour is significantly different from the M.I.T. rule. More details for the model reference approach can be found in [1, Chapter 3].

In Section 2 we introduce the M.I.T. rule using the original, heuristic arguments [2], re-interpret them in the form of the methodology discussed in Chapter 1, Section 1.2 and discuss some of its variants. Next we discuss several instability mechanisms, high gain, resonance phenomena and modelling errors. Then, using the principle of "timescale separation", suggested by the instability analysis, we derive design guidelines for good adaptive control performance. Finally we indicate how our results can be generalized and what they imply for adaptive control in general. An historical overview, in which we situate our contributions, ends the chapter.

### 2.2 The M.I.T. Rule

We refer to Figure 2.1. The plant to be controlled consists of an unknown linear, time invariant plant with strictly stable transfer function  $Z_p(s)$  and a positive, but further unknown, premultiplier/gain  $k_p$ . The gain  $k_p$  is the only parameter in the plant which possibly depends on time. Using a feedforward gain adjustment  $k_c$  for the plant  $k_p Z_p(s)$  the control objective is that the plant output  $y_p(t)$  tracks the model output  $y_m(t)$  prescribed by the parallel model with strictly stable transfer function  $Z_m(s)$  driven by the bounded reference input  $r(t)$ . Whitaker [2] suggested to select  $k_c$  the precompensator gain so as to

minimise the integral squared error:

$$I = \lim_{T \rightarrow \infty} \inf \frac{1}{T} \int_0^T e^2(t, k_C) dt \quad (2.2.1)$$

where  $e(t, k_C)$  is the output error (for fixed  $k_C$ ):

$$e(t, k_C) = y_p(t) - y_m(t) \quad (2.2.2)$$

He proposes to update  $k_C$  as follows:

$$\dot{k}_C \propto - \left( \frac{\partial e^2(t, k_C)}{\partial k_C} \right) \quad (2.2.3)$$

which is equivalent to

$$\dot{k}_C \propto - e(t, k_C) [Z_p(s)(k_p r)](t) \quad (2.2.4)$$

Equation (2.2.4) cannot be implemented as the signal  $[Z_p(s)(k_p r)](t)$  is not available. However, assuming that  $Z_m(s)$  is a good approximation for the plant's transfer function  $Z_p(s)$  and treating  $k_p$  as being constant, (2.2.4) can be approximately implemented as:

$$\dot{k}_C = -g(y_p(t) - y_m(t))y_m(t) \quad (2.2.5)$$

Where we used the approximation:

$$[Z_p(s)(k_p r)](t) \approx [Z_m(s)(r)](t) \cdot k_p \approx k_p y_m(t) \quad (2.2.6)$$

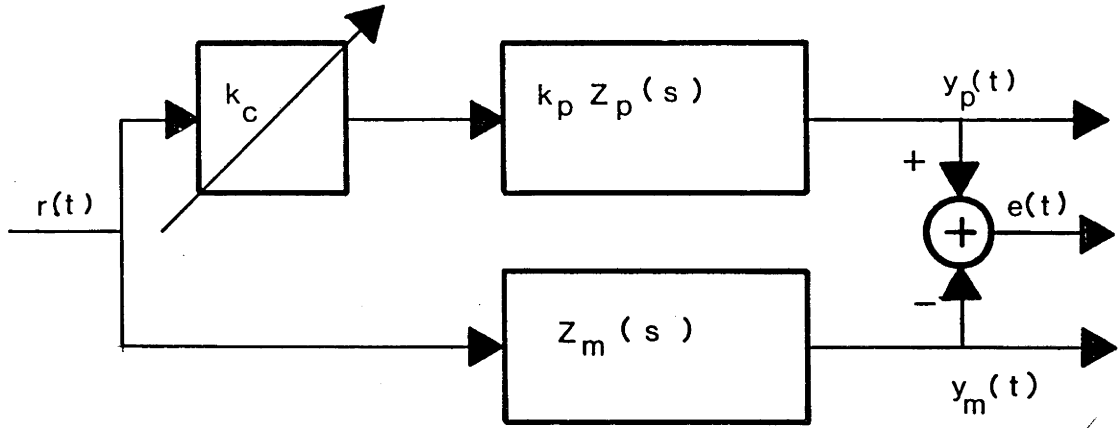
Because  $k_p$  is by assumption positive,  $g$  is a positive scalar constant scaling the adaptation speed. This is the M.I.T. rule for adaptive feedforward control (see Figure 2.2). To emphasize the dependence on  $k_C$  of the right hand side of (2.2.5) and for future analysis, we rewrite (2.5) in the following form:

$$\begin{aligned} \dot{k}_C = & -g[Z_m(s)(r)](t) \cdot [Z_p(s)(k_p k_C r)](t) \\ & + g[Z_m(s)(r)]^2(t) \end{aligned} \quad (2.2.7)$$

---

\* We denote by  $[G(s)(u)](t)$  the output at time  $t$  of a linear time invariant system with transfer function  $G(s)$  driven by the input function  $u(t)$  with zero initial state. Initial condition effects of the state of  $G(s)$  will be included only when necessary.

Figure 2.1 The M.I.T. rule

Remarks:

(R.2.1) In writing equation (2.2.7) we disregarded the effect of initial conditions in plant and model. Because the plant and model are strictly stable these terms decay exponentially fast. Exponentially decaying terms cannot effect the stability properties nor the asymptotic behaviour of the M.I.T. rule [3, Chapter 2], therefore it is permitted to disregard them.  $\square$

(R.2.2) This setup can be reinterpreted in the following way (see Chapter 1, Section 1.2):

- (1) The model set is the set of strictly stable transfer functions:

$$P = \{k_p Z_p(s) \mid k_p > 0; Z_p(s) = Z_m(s)\}$$

- (2) The control objective is to track  $y_m(t) = [Z_m(s)(r)](t)$  which can be achieved exponentially fast for any plant belonging to the model set using a feedforward control:

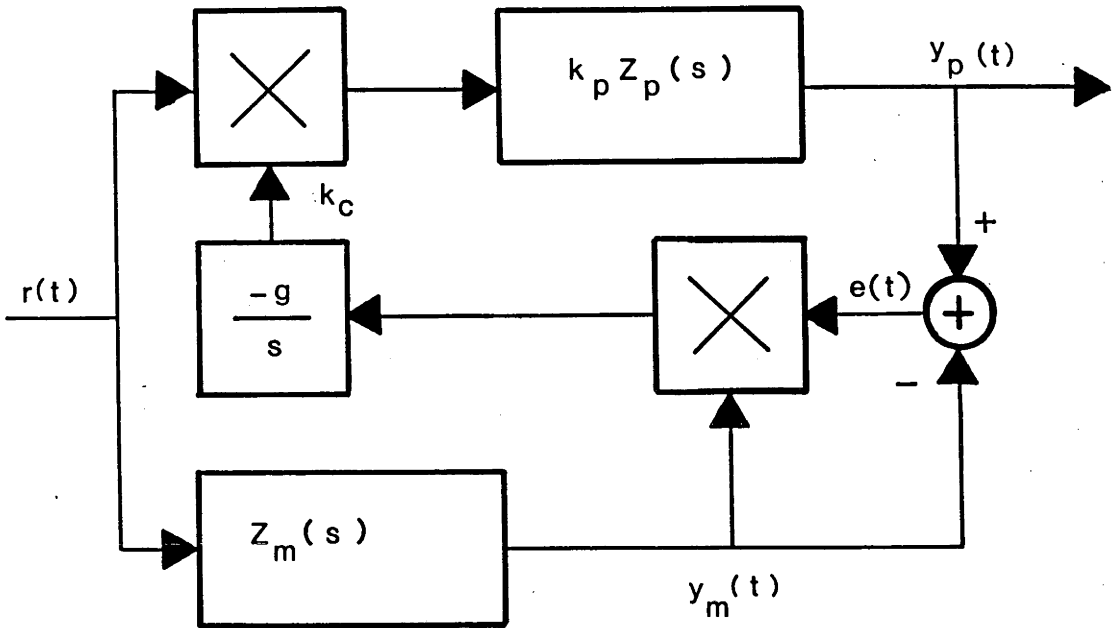
$$u_p(t) = k_c r(t) \quad (k_c = 1/k_p)$$

- (3) The identification mechanism is

$$\dot{k}_c = -g y_m(t) (y_p(t) - y_m(t)).$$

$\square$

Figure 2.2 The M.I.T. rule



(R.2.3) In the sequel we demonstrate that conditions which ensure the approximation (2.2.6) are desirable for good adaptive control performance. Notice that the essence of (2.2.6) is to make the control objective well posed.  $\square$

(R.2.4) The right hand side of (2.2.7) is affine in  $k_C$  and nonlinear in  $r$ . This explains why it is easy to characterize the global dynamics in state space ( $k_C$ ), only stable or unstable behaviour has to be considered, and why it is difficult to characterize the stability/instability boundary in parameter space ( $r$ ).  $\square$

It might appear rather strange to try to minimise the integral squared error (2.2.1) using an instantaneous gradient scheme (2.2.3). However, keeping in mind that  $k_p$  may really be slowly time varying and that it is therefore necessary to track it in order to achieve the control objective, might justify this approach. This observation motivates us to consider the two following alternatives for the M.I.T. rule:

$$\dot{k}_{C,W} \propto - \frac{\partial}{\partial k_C} \frac{1}{T} \int_{t-T}^t e^2(\tau, k_C) d\tau \quad (2.2.8)$$

$$\dot{k}_{C,E} \propto - \frac{\partial}{\partial k_C} \int_0^t e^{-\sigma(t-\tau)} e^2(\tau, k_C) d\tau, \sigma > 0 \quad (2.2.9)$$

These update laws take some of the history of the output error into account to update  $k_C$ . They are called respectively integral gradient with finite window and integral gradient with exponential forgetting. Following the same heuristic argument used to derive the M.I.T. rule leads to the following implementations of respectively (2.2.8) and (2.2.9).

$$\dot{k}_{C,W} = - \frac{g}{T} \int_{t-T}^t (y_p(\tau) - y_m(\tau)) y_m(\tau) d\tau \quad (2.2.10)$$

$$\dot{k}_{C,E} = - g \int_0^t e^{-\sigma(t-\tau)} y_m(\tau) (y_p(\tau) - y_m(\tau)) d\tau \quad (2.2.11)$$

Having introduced the M.I.T. rule via the usual heuristic arguments the pertinent questions are "What are the stability properties?", and "Does it come anywhere near optimizing the integral squared error (2.2.1)?".

Remarks:

(R.2.5) For future comparison with the M.I.T. rule, we notice that the model reference adaptive control alternative uses as update law for the feedforward gain [4]:

$$\dot{k}_{C,M} = - gr(t)(y_p(t) - y_m(t)) \quad (2.2.12)$$

or, in full

$$\begin{aligned} \dot{k}_{C,M} = & - gr(t)[Z_p(s)(rk_p k_{C,M})](t) \\ & + gr(t)[Z_m(s)(r)](t) \end{aligned} \quad (2.2.13)$$

Notice in particular that (2.2.13) is affine in  $k_{C,M}$  and nonlinear in  $r$ , but of a different type from (2.2.7).  $\square$

(R.2.6) We collect here the standing assumptions we make about the M.I.T. rule and its operating conditions:

H1:  $r(t)$  is a bounded, piecewise continuous function on  $R^+$

H2:  $k_p(t)$  the plant gain is a strictly positive, bounded, piecewise

continuous function on  $\mathbb{R}^+$

$$0 < \underline{k}_p \leq k_p(t) \leq \bar{k}_p < \infty \quad t \in \mathbb{R}^+$$

H3:  $Z_p(s)$  and  $Z_m(s)$  are causal, strictly stable transfer functions.

H4:  $Z_p(0) = Z_m(0) = 1$

Assumptions H1 and H2 are quite innocent and typical in the context of tracking a slowly time varying parameter. Hypothesis H3 is essential in order to make the feedforward control strategy meaningful. Assumption H4 excludes the model and the plant having d.c. gain of opposite sign. In view of the fact that  $Z_m(s)$  is supposedly a good model for  $Z_p(s)$ , this is a very realistic assumption. With minor modifications, which become clear in the sequel, we can deal with the situation where this assumption does not hold. However, pursuing this does not add anything substantial to our understanding of the M.I.T. rule. Notice in particular that any nonzero, positive d.c. gain different from 1 can be absorbed in  $g$  or  $k_p$  respectively for the model and the plant.  $\square$

### 2.3 Instability Mechanisms

In this section we illustrate three types of instability mechanisms: high gain, resonance phenomena and modelling errors, i.e. when the plant does not belong to the model set.

#### 2.3.1 High Gain Instability

The possibility of high gain instability can be most easily demonstrated with constant reference input  $r(t) \equiv R$ . This corresponds to setpoint regulation, a very common situation in the control of industrial processes. For  $r(t) \equiv R$  and  $k_p$  constant, the M.I.T. rule (and its alternatives) becomes asymptotically (or modulo an exponentially decaying initial condition effect) a linear, time invariant system, which can be analyzed using the Nyquist criterion, or root locus method.

**Lemma 2.1:** Under the hypotheses H.3 and H.4 (of R.2.6) the M.I.T. rule with  $r(t)$  and  $k_p(t)$  constant ( $r(t) \equiv R$ ,  $k_p(t) \equiv k_p$ ) has infinite gain margin (i.e. for all positive  $g$  and  $R$  the adaptive law is stable, independent of the gain  $k_p$ ) iff:

$$-\frac{\pi}{2} < \arg Z_p(j\omega) < \frac{3\pi}{2} \quad \omega \in \mathbb{R} \quad (2.3.1) \square$$



Proof: Neglecting the effect of initial conditions (R.2.1), the equation (2.2.7) with  $r(t) \equiv R$ ,  $k_p(t) \equiv k_p$  becomes in Laplace domain:

$$s k_c(s) = -gR^2 k_p Z_p(s) k_c(s) + gR^2 \frac{1}{s}$$

It follows that the M.I.T. rule is stable iff the zeros of

$$1 + gR^2 k_p Z_p(s)/s$$

are in the left half plane. □

Remarks:

(R.2.7) The M.I.T. rule becomes unstable, i.e. the output and the adaptive gain  $k_c(t)$  are unbounded, for sufficiently large  $g$  or  $r$  whenever  $Z_p(s)$  contains non-minimum phase zeros and/or has at least two poles more than zeros (e.g.  $Z_p(s) = 1/(s+1)^2$  yields unstable response whenever  $gR^2 > 2/k_p$ , even if  $Z_p(s) \equiv Z_m(s)$ ). This is a typical instance of high gain instability. To avoid this kind of instability requires knowledge of the low pass characteristics of  $Z_p(s)$  and an upperbound for the plant's gain  $k_p$  as well as information about the relative degree. □

(R.2.8) Condition (2.3.1) is satisfied by all strictly positive real transfer functions  $Z_p(s)$ . □

(R.2.9) For the model reference alternative (2.2.13), the same result holds. □

(R.2.10) The integral gradient alternatives (2.2.11) and (2.2.12) have even lesser gain margins. The condition (2.3.1), guaranteeing infinite gain margin for the M.I.T. rule, becomes:

$$-\frac{\pi}{2} < \arg \frac{Z_p(j\omega)}{j\omega + \sigma} < \frac{3\pi}{2} \quad (2.3.2)$$

for the integral gradient algorithm with exponential forgetting ( $\sigma$ ) and

$$-\frac{\pi}{2} + \frac{\omega T \bmod 2\pi}{2} < \arg Z(j\omega) < \frac{3\pi}{2} + \frac{\omega T \bmod 2\pi}{2}; \quad \omega \geq 0 \quad (2.3.3)$$

for the integral gradient with finite window ( $T$ ). Even strictly positive real transfer functions do not satisfy (2.3.2) or (2.3.3), hence the gain margin is always finite! □

(R.2.11) Whenever the M.I.T. rule (or any of the alternatives) is exponentially stable for  $r(t) \equiv R$  and  $k_p$  constant, the adaptive gain  $k_c$  becomes asymptotically optimal. Indeed,  $k_c$  converges exponentially to  $1/k_p$  which also minimizes the integral squared error  $I$  (2.2.1) which in this situation is given by

$$I = (k_C k_P R - R)^2. \quad \square$$

### 2.3.2 Resonance Phenomena

The instability due to resonance phenomena is illustrated using an example. With  $k_P = 1$  and  $Z_P(s) = Z_M(s) = a/(s+a)$  ( $a>0$ ) we avoid high gain instability and modelling error issues. Using as reference input  $r(t) = \cos \omega t$ , the M.I.T. rule now becomes:

$$\dot{k}_C = -g \left[ \frac{a}{s+a} (r(k_C - 1)) \right](t) - \left[ \frac{a}{s+a} (r) \right](t) \quad (2.3.5)$$

which can be equivalently represented in state space form as:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -a & a \cos \omega t \\ -g y_m(t) & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad (2.3.6)$$

where

$$x_2 = k_C - 1 (= k_C - k_P) \quad (2.3.7)$$

$$y_m(t) = \frac{a}{\sqrt{\omega^2 + a^2}} \cos(\omega t - \arctan \frac{\omega}{a})$$

and  $x_1$  is the state of the plant  $Z_P(s)$ . Initial conditions have been disregarded.

Defining  $\tau = t/a$ ,  $x_1(t) = x_1(\tau/a) = z_1(\tau)$ ,  $g/a = g'$  and  $\omega/a = \omega'$ , (2.3.6) can be re-written in the normalised form:

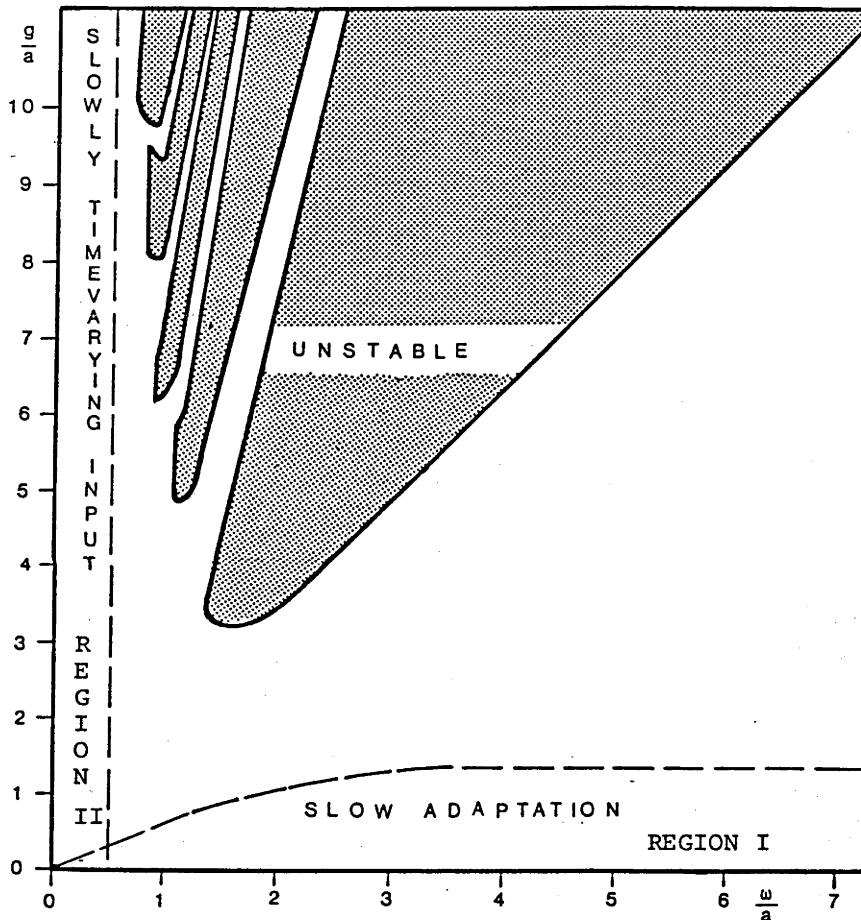
$$\begin{bmatrix} \dot{z}_1(\tau) \\ \dot{z}_2(\tau) \end{bmatrix} = \begin{bmatrix} -1 & \cos \omega' \tau \\ -g' y_m'(\tau) & 0 \end{bmatrix} \begin{bmatrix} z_1(\tau) \\ z_2(\tau) \end{bmatrix} \quad (2.3.7)$$

$$y_m'(\tau) = \frac{1}{\sqrt{\omega'^2 + 1}} \cos(\omega' \tau - \arctan \omega')$$

(2.3.7) is a linear differential equation with periodic coefficients. Its stability properties are readily analysed using Floquet theory [3, Chapter 2] in conjunction with numerical integration. (In reference [5] an analogous example has been discussed, [5] also contains the equivalent of Figure 2.3.) The results are displayed in Figure 2.3, which shows the stability domain in the frequency ( $\omega/a$ ) - gain ( $g/a$ ) parameter plane. The stability-instability boundary is extremely complex, due to the strong interaction between the system's dynamics

and the input excitation - resonance instability. The gain margin which is infinite at  $\omega = 0$  is drastically reduced around the cut off frequency ( $\omega/a = 1$ ) of the plant.

Figure 2.3 Stability boundary in parameter space for the M.I.T rule



Remarks:

(R.2.11) For parameters  $(\omega', g')$  on the stability/instability boundary the system (2.3.7) has bounded, nontrivial solutions (i.e.  $\neq 0$ ) defined on  $\mathbb{R}^+$ . (The characteristic multipliers are  $\exp(-2\pi/\omega')$  and 1 in modulus.) On the boundary the dynamics are structurally unstable, as the slightest change in the parameters can change the stability properties (Chapter 1, Section 1.3).  $\square$

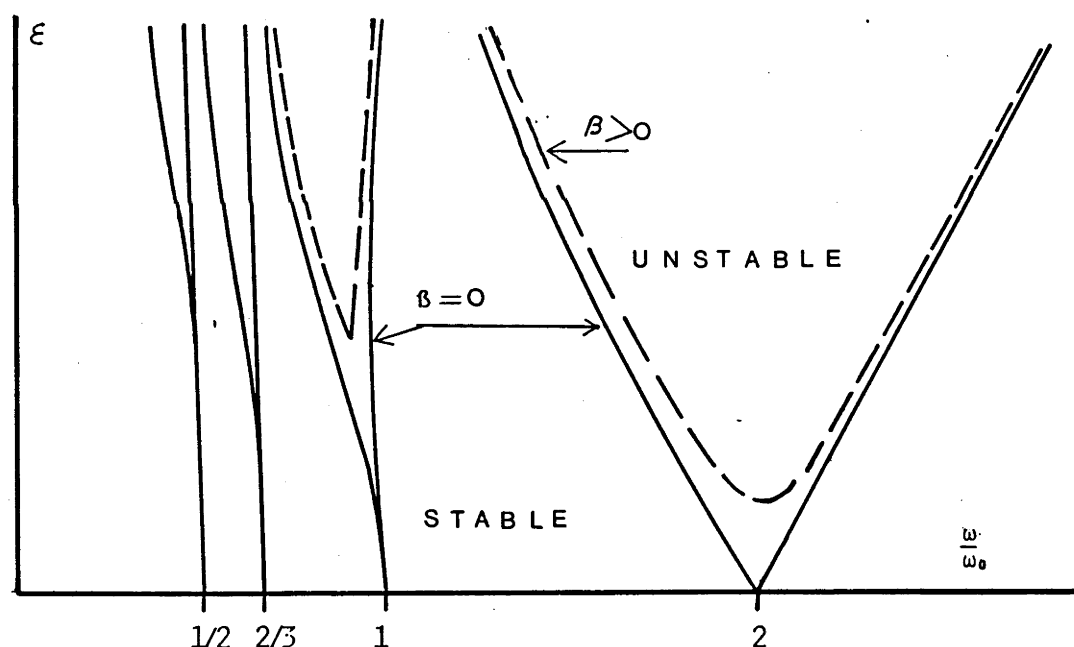
(R.2.12) Although the adaptive feedforward control problem is described by a

two dimensional set of first order, linear, periodically time varying, ordinary differential equations - how far do we want to simplify? - we are unable to give substantial analytic results without making extra assumptions. We have to resort to numerical methods. This should not surprise us too much. Compare Figure 2.3 with Figure 2.4, which displays the stability domain in parameter space for the Mathieu equation [6,7]:

$$\ddot{y}(t) + \beta \dot{y}(t) + \omega_0^2(1 + \epsilon \cos \omega t)y(t) = 0 \quad (2.3.8)$$

$$\beta \geq 0, \quad \epsilon > 0, \quad \omega > 0$$

Figure 2.4 Stability boundary in parameter space for the Mathieu equation



The relevant parameters, after time rescaling are  $\epsilon$  (parallels our gain) and  $\omega/\omega_0$  (parallels our  $\omega/a$ ). The solid curves delineate the stability/ instability boundary for  $\beta = 0$ , no damping, the dashed lines are for  $\beta > 0$  (the M.I.T. rule has damping ( $a > 0$ )). The instability regions accumulate at the the  $\omega/\omega_0 = 0$  line. The similarity is striking! Although the Mathieu equation has been around since 1868, and many analytic properties are known [6,7], basically the only way to obtain Figure 2.4 is numerical integration. This does not augur well for presenting a full analysis of (even) the linear M.I.T. feedforward adaptive control algorithm.  $\square$

(R.2.13) For the given plant, which is strictly positive real, the model reference adaptive control is uniformly asymptotically stable for all gains and frequencies. However, for plants which do not possess a strictly positive real transfer function, the model reference control and the M.I.T. rule have very similar properties. The fundamental difference between the M.I.T. rule and the model reference adaptive control algorithm for feedforward control is that the model  $Z_m(s)$  does not affect the stability properties of the model reference control law but is instrumental for the M.I.T. rule. This implies that under the same circumstances (same model and plant) the M.I.T. rule and the model reference algorithm can behave quite differently. However, both adaptive algorithms exhibit the same phenomena (high gain instability, resonance and instability due to model errors) but for different plant model combinations.  $\square$

### 2.3.3 Model Errors

Normally the plant's transfer function is not "exactly" known to the designer of the control loop, and it is therefore very likely that the plant does not belong to the model set. Modelling errors can drastically deteriorate the performance of the M.I.T. rule (cf. Lemma 1). If in the previous example  $Z_m(s) = ae^{-s}/(s+a)$ , overestimating the delay in the plant, the stability domain is drastically reduced, as is displayed in Figure 2.5. Notice in particular that the algorithm becomes unstable for  $\omega \approx \pi/2$ , and that the gain margin is even further reduced compared to the previous exact matching situation.

#### Remark:

(R.2.14) Also in this example, the model reference controller is stable, though due to the plant-model mismatch the performance is unacceptable. On the other hand if  $Z_p(s) = ae^{-s}/(s+a)$  also, then the M.I.T. rule is stable, and has

excellent performance, but the model reference controller becomes unstable for  $\omega > \pi/2$ . Figure 2.6 illustrates this point, comparing the response of the M.I.T. rule and the model reference control law, for  $r(t) = \cos 2t$ ,  $Z_p(s) = Z_m(s) = e^{-s}/(s + 1)$ .  $\square$

These examples illustrate three different instability mechanisms active in the M.I.T. rule. This does not augur well for the adequate performance of the M.I.T. rule in many situations. A rudimentary analysis of these examples (especially Figures 2.3 and 2.5) indicates that good performance is possible for a wide class of input signals for slow adaptation and indicates that higher gains are acceptable if the inputs are slow. This idea of separate time scales (respectively  $g \ll \omega$  and  $\omega \ll g$ ) will be pursued in greater detail (to rescue the M.I.T. rule).

Figure 2.5 Stability boundary in parameter space (model errors)

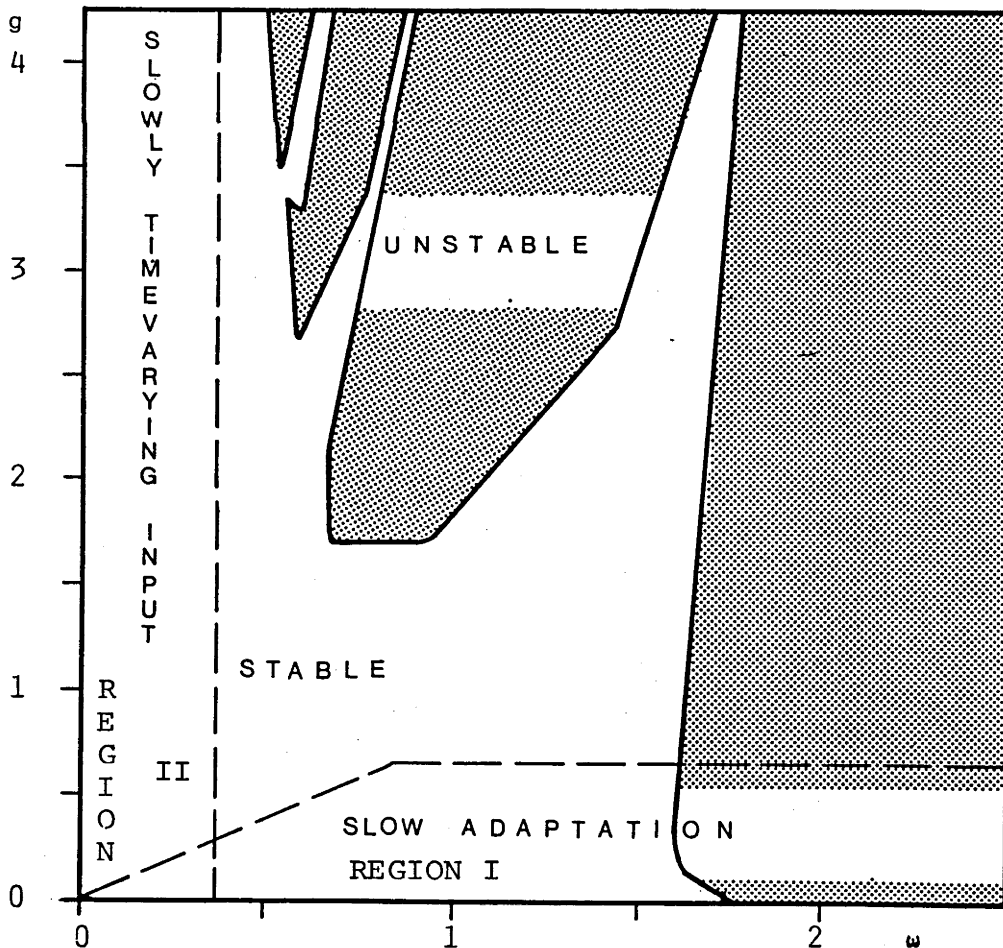
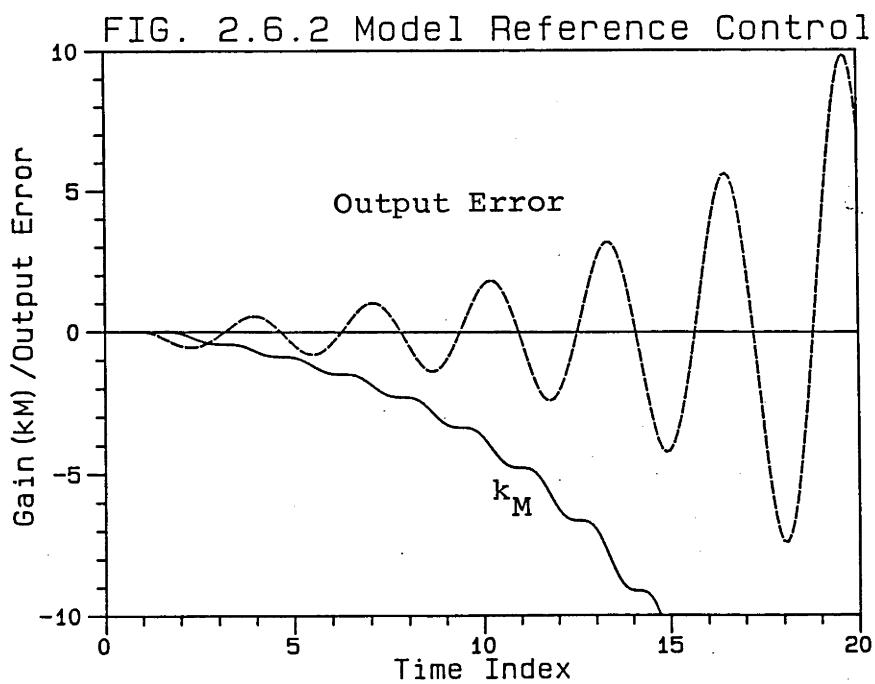
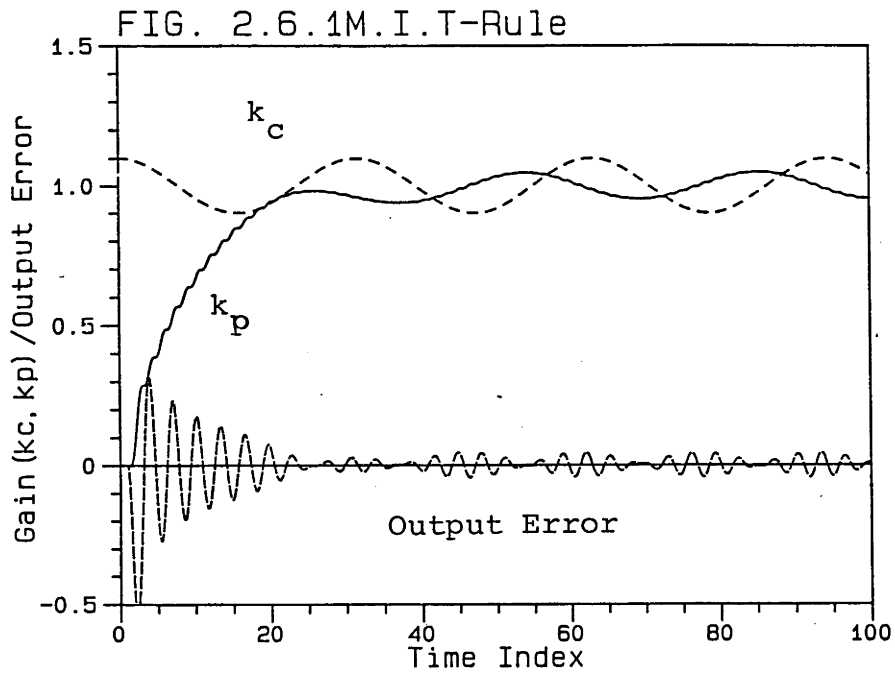


Figure 2.6 Comparing M.I.T rule with Model Reference Control



### 2.4 Stability Analysis via Averaging

Abiding by the warnings of the instability mechanisms of the previous section, we seek to consider the case of time-scale separation between the plant and the adaptation. This restriction allows us to use averaging and/or singular perturbation techniques to obtain some intuitively appealing sufficient requirements for good performance (necessary and sufficient conditions are too hard, cf. Mathieu equation and (R.2.12)). We consider two different types of timescale separation: (i) the adaptation is slow relative to the plant and reference signals; and (ii) the reference input is slow relative to the plant and the adaptation. Both stability and instability results are presented.

#### 2.4.1 Slow Adaptation

The algorithm is (cf.(2.2.7)):

$$\begin{aligned} \dot{k}_C = & -g[Z_m(s)r](t)[Z_p(s)(k_p k_C r)](t) \\ & +g[Z_m(s)r]^2(t) \end{aligned} \quad (2.4.1)$$

Assuming that  $g$  is small, i.e.  $k_C$  is slowly time varying, it is reasonable to approximate (2.4.1) by

$$\begin{aligned} \dot{k}_C^* = & -g[Z_m(s)r](t)[Z_p(s)(k_p r)](t)k_C^*(t) \\ & +g[Z_m(s)r]^2(t) \end{aligned} \quad (2.4.2)$$

which is obtained from (2.4.1) by formally treating  $k_C$  as a constant in the right hand side of (2.4.1). For sufficiently small  $g$  (2.4.1) and (2.4.2) have similar stability properties. In particular, exponential stability or instability of (2.4.2) implies the same for (2.4.1), provided  $g$  is sufficiently small. The following result is immediate.

#### Theorem 2.1:

Under the conditions:

- C1:  $r(t)$  is a bounded, piecewise continuous function on  $\mathbb{R}^+$   
 $k_p(t)$  is a bounded, piecewise continuous function on  $\mathbb{R}^+$ ;
- C2:  $Z_p(s)$ ,  $Z_m(s)$  are causal, strictly stable transfer functions allowing a finite dimensional state space representation;
- C3: the limits



$$\lim_{T \uparrow \infty} \frac{1}{T} \int_{\tau}^{\tau+T} [Z_m(s)r](t)[Z_p(s)k_p r](t)dt = \alpha \quad (2.4.3)$$

$$\lim_{T \uparrow \infty} \frac{1}{T} \int_{\tau}^{\tau+T} [Z_m(s)r]^2(t)dt = \beta \quad (2.4.4)$$

exist (uniformly in  $\tau \in \mathbb{R}^+$ );

there exists a positive constant  $g^*$ , such that for all  $g \in (0, g^*)$ :

- (i) if  $\alpha < 0$ , the M.I.T. rule (2.4.1), and its approximation (2.4.2) are unstable;
- (ii) if  $\alpha > 0$ , there exist positive monotonically increasing functions  $\delta_i(g)$  ( $\delta_i(0) = 0$ ,  $\delta_i(g) > 0$ ,  $g \in (0, g^*)$ )  $i=1,2$  such that the gain  $k_C$  as adapted by the M.I.T. rule (2.4.1) converges exponentially fast to:

$$k_C \xrightarrow{(\exp)} \frac{\beta}{\alpha} + \delta_1(g) \text{ as } t \uparrow \infty \quad (2.4.5)$$

and also

$$k_C^* \xrightarrow{(\exp)} \frac{\beta}{\alpha} + \delta_2(g) \text{ as } t \uparrow \infty \quad (2.4.6)$$

where  $k_C^*$  is the solution of (2.4.2). □

Proof: Provide a state space realization for (2.4.1) and apply Theorem A.18 of the appendix. □

Remarks:

(R.2.14) In the case that  $r(t)$  and  $k_p(t)$  are finite sums of periodic signals, the limits (2.4.3) and (2.4.4) exist and the function  $\delta(g) = 0(g)$ . (See appendix, Lemma A.2, Remark A.2.3.) □

(R.2.15) The constant  $g^*$  above, may be quantified in terms of  $\alpha$  in (2.4.3) and further characteristics of  $r$ . ( $g^*$  is proportional to  $\alpha$ ). Averaging theory permits us only to look with surety up to  $g^*$  and gives us no further information about the properties of (2.4.1) in terms of (2.4.2). The boundary may be conservative. (Higher order averaging, as opposed to first order averaging used in the above, can however be appealed to to obtain stability

information even when  $\alpha = 0$ , cf. [6,7,9].)  $\square$

(R.2.16) The time invariance of  $k_p$  has not been invoked. Given the usual rationale of adaptive systems of adjustment to slowly-varying parameter values  $k_p$ , one can allow for  $k_p$  time variations, provided at least that  $k_p$  does not change sign (cf R.2.6, Hypothesis H.2). Using the same averaging principles in allowing  $k_p$  to vary more slowly than the adaptation, it is possible to split the timescales into three different components and (2.4.3) can be replaced by:

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_{\tau}^{\tau+T} [Z_m(s)r](t)[Z_p(s)r](t)dt = \alpha' > 0 \quad (2.4.6)$$

$\square$

(R.2.17) Conclusion (ii) of Theorem 2.1 indicates how the M.I.T. rule performs in terms of optimizing the integral squared error  $I$  (2.2.1).

Assuming that the limit

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_{\tau}^{\tau+T} [Z_p(s)rk_p]^2(t)dt = \gamma \quad (2.4.7)$$

exists, the integral squared error  $I$  (under the conditions of Theorem 2.1) is optimized by:

$$k_c^{\text{opt}} = \alpha/\gamma \quad (2.4.8)$$

From the Cauchy-Schwartz inequality we obtain that, for all  $\tau$  and  $T \in \mathbb{R}^+$ :

$$\left[ \int_{\tau}^{\tau+T} [Z_p(s)rk_p](t)[Z_m(s)r](t)dt \right]^2 \leq \int_{\tau}^{\tau+T} [Z_p(s)rk_p]^2(t)dt \cdot \int_{\tau}^{\tau+T} [Z_m(s)r]^2(t)dt$$

hence

$$\alpha/\gamma \leq \beta/\alpha$$

which implies that the gain  $k_c$  as updated by the M.I.T. rule is biased in general (equality holds for  $Z_p^*(s) = Z_m(s)$  and  $k_p$  constant). We have:

$$k_C - k_C^{\text{opt}} \xrightarrow{(\text{exp})} \frac{\beta}{\alpha} - \frac{\alpha}{\gamma} + \delta(g) \quad t \uparrow \infty \quad (2.4.9)$$

for all  $g \in (0, g^*)$ . □

Before proceeding to interpret (2.4.6), we demonstrate that  $k_C^*$  (2.4.2) is a good approximation for  $k_C$  (2.4.1), without invoking the existence of the averages (2.4.3), (2.4.4), nor using the assumption that  $Z_p(s)$  is finite dimensional. The proof is specifically helpful in quantifying  $g^*$ , however only stability (cf. conclusion (ii) in Theorem 2.1) is addressed. Firstly rewrite (2.4.1) in the following equivalent form:

$$\begin{aligned} \dot{k}_C(t) = & -gy_m(t) \int_0^t h(t-\tau)k_p(\tau)r(\tau)k_C(\tau)d\tau \\ & + gy_m^2(t); \quad k_C(0) \end{aligned} \quad (2.4.10)$$

where  $y_m(t) = [Z_m(s)r](t)$  and  $h(t)$  is the impulse response (at time  $t$  for an impulse at 0) of the transfer function  $Z_p(s)$ . The effect of initial conditions of the plant is not taken into account (cf. (R.2.1)). Equation (2.4.10) is of the form:

$$\dot{x}(t) = g \int_0^t H(t,\tau)x(\tau)d\tau + gB(t); \quad x(0) = x_0 \quad (2.4.11)$$

(Comparing (2.4.11) with (2.4.10)  $H$  and  $B$  can be identified as respectively  $H(t,\tau) = -y_m(t)h(t-\tau)k_p(\tau)r(\tau)$  and  $B(t) = y_m^2(t)$ .)

We first demonstrate the following result.

**Lemma 2.2:** Assume that the following conditions hold:

- C1:  $B(t)$  is a continuous, bounded (matrix) function on  $R^+$ ;
- C2: the kernel  $H(t,\tau)$  has the properties:

$$\begin{aligned} \text{(i)} \quad & \int_0^\tau H(t,s)ds \text{ is piecewise continuous in } (t,\tau) \text{ on } R^+ \times R^+ \\ \text{(ii)} \quad & \left| \int_0^\tau H(t,s)ds \right| \leq K_2 e^{-b(t-\tau)} \quad \forall t \geq \tau \end{aligned} \quad (2.4.12)$$

for some positive constants  $K_2$  and  $b$ ;

C3:  $\Phi(t, \tau)$  the fundamental matrix of the linear, time varying equation

$$\dot{y}(t) = g \left( \int_0^t H(t, \tau) d\tau \right) y(t)$$

satisfies

$$\|\Phi(t, \tau)\| \leq K_1 e^{-ga(t-\tau)} \quad \forall t \geq \tau \quad (2.4.13)$$

for some positive constant  $a > 0$  and  $K_1 \geq 1$ .

Then there exists a positive constant  $g^*$  such that for all  $g \in (0, g^*)$  the solution  $x(t, 0, x_0)^{\dagger}$  of (2.4.11) satisfies:

$$\|x(t, 0, x_0) - y(t, 0, x_0)\| = o(g) \quad \forall t \in \mathbb{R}^+ \quad (2.4.14)$$

where  $y(t, 0, x_0)$  is the solution of:

$$\dot{y}(t) = g \left( \int_0^t H(t, \tau) d\tau \right) y(t) + gB(t) \quad (2.4.15)$$

passing through  $x_0$  at  $t = 0$ . □

#### Remarks:

(R.2.18) Notice that  $H(t, \tau)$  may contain an impulse function, i.e.  $Z_p(s)$  is allowed to have a direct throughput (cf. (2.4.10) and (2.4.11)). □

(R.2.19) Although for the moment we need this result only for scalar  $x$ ,  $B$  and  $H$  it is demonstrated for the general matrix case, anticipating forthcoming generalizations. □

Proof: Starting from (2.4.11), and integrating by parts we obtain:

$$\dot{x}(t) = g \left( \int_0^t H(t, \tau) d\tau \right) x(t) - g \int_0^t \left( \int_0^{\tau} H(t, s) ds \right) \dot{x}(\tau) d\tau + gB(t)$$

---

<sup>†</sup>By  $x(t, t_0, x_0)$  we denote this solution of an ordinary differential equation ( $\dot{x} = f(x, t)$ ) which passes through  $x_0$  at time  $t_0$ . (e.g.  $z(t, t_0, z_0) = \exp(a(t-t_0))z_0$ , for  $\dot{z} = az$  with  $z(t_0) = z_0$ ).

which can be rewritten in operator form as

$$[(I+F)\dot{x}](t) = g \int_0^t H(t, \tau) d\tau x(t) + gB(t) \quad (2.4.15)$$

where  $F$  is a linear operator defined as

$$(Fy)(t) = g \int_0^t \left( \int_0^\tau H(t, s) ds \right) y(s) ds. \quad (2.4.16)$$

$F$  is defined on the Banach space of continuous, bounded functions, equipped with the supremum norm, its induced operator norm is [3,10]

$$\|F\| = gK_2/b \quad (2.4.17)$$

which follows from (2.4.16) upon using condition (4.13). Choosing  $g < g_1 = b/K_2$ , the operator  $I+F$  has an inverse, which can be written as:

$$(I+F)^{-1} = I - F + F^2 - F^3 + \dots \quad (2.4.18)$$

denoting

$$G = -F + F^2 - F^3 + \dots \quad (2.4.19)$$

we have

$$\|G\| \leq (gK_2/b)/(1-gK_2/b) \quad (2.4.20)$$

$$\|(I+F)^{-1}\| \leq 1/(1-gK_2/b)$$

and

$$\begin{aligned} \dot{x}(t) &= g \int_0^t H(t, \tau) d\tau x(t) + gB(t) \\ &+ G \left( g \int_0^\cdot H(\cdot, s) ds x(\cdot) \right)(t) + G(gB(\cdot))(t). \end{aligned} \quad (2.4.22)$$

Furthermore we obtain

$$\begin{aligned} x(t) &= \Phi(t, 0)x(0) + \int_0^t \Phi(t, s)gB(s)ds \\ &+ \int_0^t \Phi(t, s)G \left( g \int_0^\cdot H(\cdot, \tau) d\tau x(\cdot) \right)(s)ds \end{aligned}$$

$$+ \int_0^t \Phi(t,s) G(gB(\cdot))(s) ds . \quad (2.4.23)$$

Define the operator  $N$ , on the Banach space of continuous bounded functions equipped with the supremum norm as

$$\begin{aligned} N(y)(t) &= \Phi(t,0)x(0) + \int_0^t \Phi(t,s) [(I+F)^{-1}gB(\cdot)](s) ds \\ &+ \int_0^t \Phi(t,s) G\left(g \int_0^s H(\cdot, \tau) d\tau y(\cdot)\right)(s) ds . \end{aligned}$$

We now show that  $N$  is a contraction operator for sufficiently small  $g$ , and therefore has a unique fixed point, solution of (2.4.23), and therefore of (2.4.11). Indeed  $N$  is well defined:

$$\begin{aligned} \|N(y)\| &\leq K_1 \|x(0)\| + \frac{K_1}{a} \| (I+F)^{-1} \| \cdot \|B\| \\ &+ \frac{K_1}{ag} \cdot \|G\| \cdot g \cdot K_2 \cdot \|y\| < \infty \end{aligned}$$

and  $N$  is a contraction for

$$g \leq g^* = \min \left[ \frac{b}{2K_2}, \frac{ab}{4K_1K_2^2} \right] \quad (2.4.24)$$

since

$$\begin{aligned} \|N(y_1) - N(y_2)\| &\leq \frac{K_1}{ag} \cdot \|G\| \cdot g \cdot K_2 \|y_1 - y_2\| \\ &\leq \frac{K_1K_2}{a} \cdot \frac{gK_2/b}{1-gK_2/b} \|y_1 - y_2\| \\ &\leq \left( \frac{2K_1K_2^2}{ab} \right) \cdot g \cdot \|y_1 - y_2\| \\ &\leq \frac{1}{2} \|y_1 - y_2\| \end{aligned} \quad (2.4.25)$$

The solution of (2.4.23) or (2.4.11) can then be approximated up to first order in  $g$  as:

$$\|x(t, 0, x(0)) - \Phi(t, 0)x(0) - \int_0^t g\Phi(t, s)B(s)ds\| = o(g).$$

It suffices to notice that:

$$\begin{aligned} & \|x(t, 0, x(0)) - N(0)\| \\ &= \lim_{k \uparrow \infty} \|W^k(0) - N(0)\| \\ &\leq \lim_{k \uparrow \infty} \sum_{l=1}^{k-1} \left[ \frac{2K_1 K_2^2}{ab} g \right]^l \|W(0)\| \\ &\leq \left[ \frac{2K_1 K_2^2}{ab} g \right] / \left[ 1 - \frac{2K_1 K_2^2}{ab} g \right] \|W(0)\| = o(g) \end{aligned}$$

and

$$\|W(0) - \Phi(t, 0)x(0) - \int_0^t g\Phi(t, s)B(s)ds\| = o(g)$$

because  $\|G\| = o(g)$  (see (2.4.20)).  $\square$

This lemma is readily applicable to the M.I.T. rule, or any of its discussed alternatives, it suffices to verify the conditions of the Lemma. The following Theorem is an immediate consequence of this Lemma, and complements Theorem 2.1.

**Theorem 2.2:**

Under the conditions that:

C1:  $r(t)$ ,  $k_p(t)$  are bounded, piecewise continuous functions on  $\mathbb{R}^+$ ;

C2:  $Z_p(s)$ ,  $Z_m(s)$  are causal, strictly stable transfer functions:

$$C3: \liminf_{T \uparrow \infty} \frac{1}{T} \int_0^T [Z_m(s)r](t)[Z_p(s)rk_p](t)dt > \alpha > 0 \quad (2.4.26)$$

there exists a positive constant  $g^*$ , such that for all  $g \in (0, g^*)$  the M.I.T. rule (2.4.1) has the properties:

- (i) there exists a unique solution to (2.4.1);  $k_C(t, 0, k_{C0})$  starting at  $t = 0$  in  $k_{C0}$ , bounded, well defined on  $\mathbb{R}^+$ , for any (finite)  $k_{C0}$ .
- (ii) the effect of the initial condition decays exponentially fast;
- (iii)  $|k_C(t, 0, k_{C0}) - k_C^*(t, 0, k_{C0})| = o(g) \quad \forall t \in \mathbb{R}^+ \quad (2.4.27)$   
 where  $k_C^*(t, 0, k_{C0})$  is the unique solution of (2.4.2) passing through  $k_{C0}$  at  $t = 0$  □

Proof: Condition C1 of Theorem 2.2 implies C1 of Lemma 2.2; condition C2 of Lemma 2.2 follows from C1 and C2 of Theorem 2.2 and condition C3 implies C3 of Lemma 2.2. □

Remarks:

(R.2.20) Theorem 2.2 imposes much less restrictive conditions than Theorem 2.1, yet yields basically the same conclusion - apart from an instability limit. At first glance the error estimation (2.4.27) might appear better than (2.4.24),  $0(g) \leftrightarrow \delta(g)$ . However (2.4.27) does not involve any averaging, the result compares  $k_C$  with  $k_C^*$ , not  $k_C$  or  $k_C^*$  with the response of the averaged equation; introducing averaging one obtains:

$$k_C, k_C^* \xrightarrow{(\exp)} \beta/\alpha + \delta(g)$$

as before. □

(R.2.21) From (2.4.24) and (2.4.26) it follows that  $g^*$  is roughly proportional to  $\alpha$ , inversely proportional to  $\|r\|^2$  and also inversely proportional to the gain of  $Z_p(s)$  and  $Z_m(s)$ . This implies in particular that  $g^*$  is proportional to the condition number of the excitation. It also follows that the "natural" adaptation gain of the algorithm is  $g\|r\|^2$ , cf. Lemma 2.1. □

(R.2.22) As pointed out, Lemma 2.2 is also applicable to the M.I.T. rule's alternatives, in particular, stability condition (2.4.26) becomes:

$$\liminf_{S \uparrow \infty} \frac{1}{S} \int_0^t \left[ \int_0^{\tau_1} e^{-\sigma(t-\tau_1)} \int_0^{\tau_2} [Z_m(s)r](\tau_2) [Z_p(s)rk_p](\tau_2) d\tau_2 d\tau_1 \right] dt \gg \alpha > 0 \quad (2.4.28)$$

for the integral gradient scheme with exponential forgetting, or



$$\liminf_{S \uparrow \infty} \frac{1}{S} \int_T^{S+T} \frac{1}{t} \int_{t-T}^t [Z_m(s)r](\tau)[Z_p(s)rk_p](\tau) d\tau dt \geq \alpha > 0 \quad (2.4.29)$$

for the integral gradient with finite window. Clearly (2.4.28) and (2.4.29) are harder to meet than condition (2.4.26), confirming our earlier observation that these alternatives have a smaller stability margin than the M.I.T. rule. For the model reference control algorithm we have

$$\liminf_{t \uparrow \infty} \frac{1}{T} \int_0^T r(t)[Z_p(s)k_{pr}](t) dt \geq \alpha > 0 \quad (2.4.30)$$

Replacing (2.4.26) in Theorem 2.2 by (2.4.28), (2.4.29) or (2.4.30) yields equivalent results for respectively the integral gradient and model reference control algorithms.  $\square$

(R.2.23) Loosely speaking, but in intuitively appealing terms, Theorem 2.1 and Theorem 2.2 assert that the M.I.T. rule, as far as the update mechanism is concerned, reacts as a first order system with cut off frequency of  $\alpha g$  (for  $g$  sufficiently small).  $\square$

Finally, we interpret condition (2.4.26) or (2.4.3) in terms of the spectral properties of almost periodic inputs  $r(t)$  of the form:

$$r(t) = \sum_{i=-\infty}^{+\infty} a_i e^{j\omega_i t}; \quad a_i = a_{-i}^*; \quad \omega_i = -\omega_{-i}; \quad \sum_{i=-\infty}^{+\infty} |a_i|^2 < \infty \quad (2.4.31)$$

We assume  $k_p$  constant and  $Z_p(0) = Z_m(0) = 1$ . (2.4.26) or (2.4.3) then become

$$|a_0|^2 + 2 \sum_{i=1}^{+\infty} |a_i|^2 \operatorname{Re}(Z_m(-j\omega_i)Z_p(j\omega_i)) \geq \alpha > 0 \quad (2.4.32)$$

(2.4.32) indicates that the M.I.T. rule is stable (provided  $g$  is sufficiently small) if  $Z_m$  is close to  $Z_p$  (at least in phase) on the spectral lines ( $\omega_i$ ) of the input  $r(t)$ , (cf.[8]). In particular (2.4.32) is always satisfied if:

$$|\arg Z_m(j\omega) - \arg Z_p(j\omega)| < \frac{\pi}{2}$$

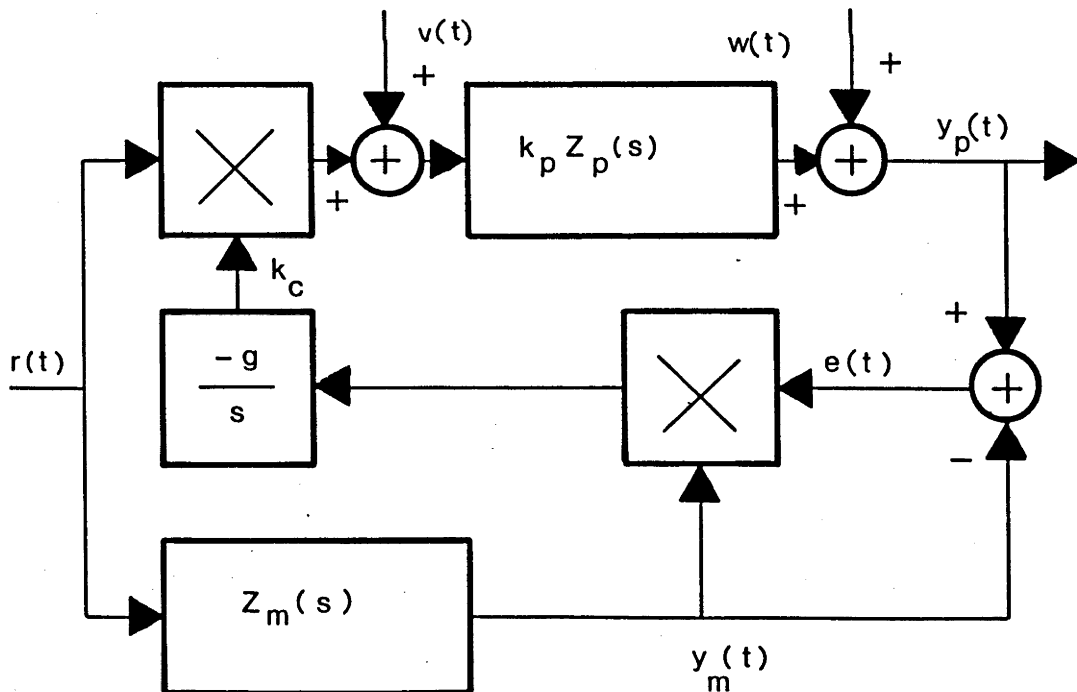
which is the equivalent of the strictly positive real condition of the model

reference approach for the M.I.T. rule. However, here this condition is not as powerful as the provision "g small" is necessary, whilst it is not for the model reference control law. If the left hand side of (2.4.32) is negative, then the M.I.T. rule is unstable, for g sufficiently small. Notice that this instability can only be due to model errors, as the small gain assumption avoids both resonance and high gain instability.

The stability properties asserted in Theorem 2.1 and Theorem 2.2 are robust - because of the exponential stability - with respect to other nonmodelled effects not taken into account. Small nonlinearities in the plant will not destroy stability [3], whilst bounded disturbances on plant input and/or output do not effect the stability, only add to the output error. Referring to Figure 2.7, in the situation of bounded disturbances, the M.I.T. rule can be described as:

$$\begin{aligned} \dot{k}_c = & -g[Z_m(s)r](t)[Z_p(s)k_p k_c r](t) \\ & + g y_m^2(t) - g y_m(t)w(t) - g y_m(t)[Z_p(s)k_p v](t). \end{aligned}$$

Figure 2.7 The M.I.T. rule with disturbances



Clearly, the input noise  $v$  and the output noise  $w$  do not affect stability and as the M.I.T. rule has effectively a cut off frequency of the order of  $g$  (for small  $g$ , cf. (R.2.23)) the effect of the noise on  $k_C$  is only felt if the noise has substantial power in an  $O(g)$  neighbourhood of the spectral lines of the input  $r(t) - (y_m w \text{ and } y_m v!)$ .

The analysis so far has concentrated on time scale separation involving slow adaptation compared with plant dynamics and input signals. These results predict in particular the simulation outcome displayed in Figure 2.3, Region I (Slow Adaptation), indicating stable behaviour for small  $g$  and  $g \ll \omega$  and predict instability for  $\pi/2 \leq \omega \leq 3\pi/2$  and small gain as displayed in Figure 2.5. Notice that in this case  $|\arg Z_p(j\omega) - \arg Z_m(j\omega)| = \omega!$

#### 2.4.2 Slowly Time Varying Inputs

The M.I.T. rule for adaptive feedforward control is described by

$$\dot{k}_C = -g[Z_m(s)r](t)[Z_p(s)rk_p k_C](t) + g[Z_m(s)r]^2(t) \quad (2.4.1)$$

assuming that  $r(t)$  and  $r(t)k_p(t)$  are slowly time varying, we can attempt to approximate  $k_C$  by  $\bar{K}_C$  solution of (see (R.2.6))

$$\dot{\bar{K}}_C = -gr^2(t)k_p(t)[Z_p(s)\bar{K}_C](t) + gr^2(t) \quad (2.4.33)$$

The stability properties of (2.4.33) may be derived simply as an extension of the root locus method, used in Section 2.3.1. From equation (2.4.33), provided we can demonstrate stability, it appears that after an exponentially decaying transient  $\bar{K}_C(t)$  should track  $1/k_p(t)$ . This heuristic approach can be justified. We have the result:

##### Theorem 2.3:

Under the hypotheses:

- C0:  $Z_p(s)$ ,  $Z_m(s)$  are strictly stable, causal transfer functions, which have a finite dimensional state space realization ( $Z_m(0) = Z_p(0) = 1$ ).
- C1: the zeros of  $s + g\gamma Z_p(s)$  have real part less than  $-\sigma$ ,  $\sigma > 0$ , for all  $\gamma \in (\gamma_0, \gamma_1)$ ,  $0 < \gamma_0 < \gamma_1 < \infty$ .
- C2:  $r(t)$ ,  $k_p(t)$  are bounded, continuous functions on  $\mathbb{R}^+$  satisfying
  - (i)  $r^2(t)k_p(t) \in (\gamma_0, \gamma_1) \quad \forall t \in \mathbb{R}^+$

$$(ii) |r(t_1) - r(t_2)| \leq \delta |t_1 - t_2| \quad \forall t_1, t_2 \in \mathbb{R}^+$$

$$|k_p(t_1)r(t_1) - k_p(t_2)r(t_2)| \leq \delta |t_1 - t_2| \quad \forall t_1, t_2 \in \mathbb{R}^+$$

Then for every  $\mu \in (0, \sigma)$  there exists a  $\delta^*$  such that for all  $\delta \in (0, \delta^*)$  the solution  $k_C(t, 0, k_{C0})$  of (2.4.1) starting in  $k_{C0}$  at  $t=0$  satisfies:

$$|k_C(t, 0, k_{C0}) - k_p^{-1}(t)| \leq K e^{-(\sigma - \mu)t} C_0 + O(\delta) \quad (2.4.34)$$

where  $K$  depends on  $\mu$  and  $C_0$  is the norm of the initial conditions of the plant and  $k_{C0}$ .  $\square$

Proof: Provide a state space realization for (2.4.1) and apply Proposition 5 in [11, Chapter 1, p.6] to obtain that the homogeneous equation is exponentially stable. Using the variation of constants formula and C2 yields the desired result.  $\square$

#### Remarks:

(R.2.24) With minor modifications, piecewise continuous functions can be treated. The same techniques handle the multivariable situation as well.  $\square$

(R.2.25) An alternative proof can be obtained from [10, pp125-127].  $\square$

(R.2.26) Theorem 2.3 explains the stability properties of the M.I.T. rule for slowly time varying inputs. In particular it predicts the stability in regions II in Figures 2.3 and 2.5. Notice that Theorem 2.3 is only concerned with high gain type of instability, as the assumption of slow inputs avoid the resonance phenomena, and model errors are irrelevant because  $Z_m(0) = Z_p(0) = 1$ , indicating good model-plant match at low frequencies. This theorem extends Lemma 2.1, which can be obtained as a special case from Theorem 2.3 with  $\delta=0$  and  $\mu=0$ . In this sense, the results of this section and the previous one complement each other, as they discuss different working conditions in the parameter space.  $\square$

(R.2.27) Under the conditions of Theorem 2.3, the M.I.T. rule is close to optimal behaviour, in the sense of minimizing the integral squared error, since  $1/k_p$  is indeed optimal.  $\square$

### 2.5 Generalizations

Although the object of our discussion was a scalar parameter update law, all theorems and main lemmas (cf. Appendix and Lemma 2.2) have been

presented for the multivariable situation. Instances where these results are directly applicable are adaptive feedforward control schemes for multivariable plants, where  $k_c$  is a matrix and adaptive equation error identification with regressor and/or error filtering. In both cases, the overall system can be represented as a linear in the state, multivariable time varying system. In general adaptive control leads to nonlinear systems in which situation, these results can be used to describe their local properties (cf. [1] for more along these lines).

### 2.6 Historical Overview of the M.I.T. Rule

The M.I.T. rule for adaptive control, here presented in its simplest form, was formulated in the late fifties and early sixties as a model reference adaptive control law for linear systems modelled as a cascade of a known stable plant and a single unknown gain (see Figure 2.1). The names generally associated with the formulation are Whitaker, Osburn and Kezer [2]. The initial intended application was to optimize the performance of aircraft, where the single unknown gain was related to dynamic pressure.

In the history of adaptive control, or at least its folklore, the M.I.T. rule represents a watershed. The method was simply formulated, easily appreciated and was directly applicable. Consequently, this approach to self-optimizing systems was taken up by theorists and practitioners alike as a potential route to enhanced performance. In application trials with aircraft dynamics, however, the M.I.T. rule adaptive controller led to unpredicted instability with a considerable associated loss of face and confidence in ad hoc adaptive control.

Simulation studies provided some idea [5] of the rule's stability properties and indicated the likely complexity of any analysis. Donalson and Leondes (1963) [12] indicated engineering guidelines, rules of thumb "guaranteeing" good performance, which are disturbingly close to currently emerging modern notions of suitable operating conditions for adaptive controllers. We quote Donalson and Leondes, transliterating into our framework:

"Assumption 1:  $k_p$  varies slowly compared to the basic time constants of the physical process  $Z_p(s)$  and the reference model  $Z_m(s)$ ";  
and

"Assumption 2:  $k_p$  varies slowly compared to the rate at which the adjusting mechanism, to be designed, updates  $k_c$ ";

and

"Assumption 3: the adjusting mechanism updates  $k_c$  fast compared to the effect caused by the reference input on the output error"

(This corresponds to the slow input case.) Our contribution (see also [13,14,15]) provides a rigorous basis for the design guidelines derived earlier (cf. the above quotes from [12] and Theorem 2.3). Also we extend these guidelines in discussing a larger portion of the design-variables space, including slow adaptation compared to the effect of the input on the output error (opposite to Assumption 3 in [12]). Moreover, we explain why and how the M.I.T. rule can fail. The guiding principles for good performance emerging from our analysis are:

- (a) the natural timescales present in the adaptive system should be well separated (plant, model/input/adaptive gain/time varying gain  $k_p$ );
- (b) model and plant need to be well matched over the frequency range where the input has its dominant power;
- (c) input's power should be concentrated outside the frequency range of significant noise power.

These conditions have also been recently espoused as good engineering sense in [1] with regard to a broad class of adaptive algorithms.

Concisely, we have revisited the M.I.T. rule demonstrating the reasons for the loss of confidence in this rule and have suggested potential remedies necessary to resuscitate it. We also indicated that it can out-perform the more fashionable model reference adaptive control designed via Lyapunov techniques [12], see Figure 2.5, as opposed to the M.I.T.'s criterion minimisation approach. (A discussion of more general adaptive control algorithms based on direct criterion minimisation can be found in [16].)

In any event, the M.I.T. rule served our primary purpose of highlighting the dependence of the adaptive control performance on the design parameters. Especially in the (likely) situation of model-plant-mismatch a careful selection of the adaptation gain and the spectral properties of the reference input is essential for guaranteeing good adaptive performance. Notice also the complexity of describing the asymptotic dynamics (stability/instability) in parameter space (Figures 2.3 and 2.5) ... a complete analytical description is beyond our capabilities, (as we have argued, hopefully convincingly!).

2.7 APPENDIX: Averaging Theorems for linear timevarying systems.

Consider the piecewise continuous and bounded matrix function  $A(t)$  defined on  $\mathbb{R}^+$ .

A.1 Definition [14,17]: The matrix function  $A(t)$  possesses a uniform average  $\bar{A}$  if there exists a bounded decreasing, positive function  $c$  such that

$$\sup_{s \in \mathbb{R}^+} \left\| \bar{A} - \frac{1}{T} \int_s^{s+T} A(t) dt \right\| \leq c(T), \quad T \in \mathbb{R}^+ \quad (2.7.1)$$

where  $c$  decreases monotonically to zero as  $T$  increases:

$$c(T) \downarrow 0 \text{ as } T \uparrow \infty \quad (c(T_1) \leq c(T_2) \quad \forall T_1 > T_2).$$

$c$  is called the convergence function. □

The first Lemma links the notions of integral smallness and uniform average:

A.2 Lemma: If the matrix  $A(t)$  possesses the uniform average  $\bar{A}$ , then for any positive constant  $h$  there exists a monotonically nondecreasing, positive function  $\delta_h(\mu)$  such that

$$\delta_h : \mathbb{R}^+ \rightarrow \mathbb{R}^+ \quad \mu \mapsto \delta_h(\mu), \quad \delta_h(0) = 0$$

$$\delta_h(\mu_1) \leq \delta_h(\mu_2) \quad \forall \mu_1 \leq \mu_2$$

and

$$\left\| \int_{t_1}^{t_2} \left( A\left(\frac{t}{\mu}\right) - \bar{A} \right) dt \right\| \leq \delta_h(\mu) \quad (2.7.2)$$

for all  $|t_1 - t_2| < h$ . □

Proof: By assumption  $A(t) - \bar{A}$  is bounded, say  $\|A(t) - \bar{A}\| < M$  for all  $t$ , hence for all  $|t_2 - t_1| \leq \delta(\mu)/M$  we have

$$\left\| \int_{t_1}^{t_2} \left( A\left(\frac{t}{\mu}\right) - \bar{A} \right) dt \right\| \leq \delta(\mu) \quad \forall \mu \in \mathbb{R}^+.$$

On the other hand, for  $\delta(\mu)/M \leq |t_2 - t_1| \leq h$ , we have

$$\begin{aligned} \left\| \int_{t_1}^{t_2} \left( A\left(\frac{t}{\mu}\right) - \bar{A} \right) dt \right\| &= \left\| \mu \int_{t_1/\mu}^{t_2/\mu} (A(t) - \bar{A}) dt \right\| \\ &\leq h \cdot c \left[ \left| \frac{t_1 - t_2}{\mu} \right| \right] \\ &\leq h \cdot c(\delta(\mu)/\mu M) \end{aligned}$$

Define  $\delta_h(\mu)$  as either the solution of

$$\delta - h c(\delta/\mu M) = 0$$

if  $\delta \leq h/M$ , where  $c$  is the convergence function of  $A(t)$ , or  $\delta_h(\mu) = hM$ .  $\delta_h$  is uniquely defined, because  $c(T) > 0$  and monotonically decreasing. This yields the desired result.  $\square$

**A.3 Remark:** The estimate for the function  $\delta_h(\mu)$  derived in the Lemma is conservative. Indeed for  $A(t)$  periodic or a finite sum of periodic matrices,  $\delta_h(\mu)$  is of the order of  $\mu$  ( $\delta(\mu) \leq hM$ ) and the Lemma yields  $\delta_h(\mu)$  of the order of  $\sqrt{\mu}$  because  $c(T)$  is of the order of  $1/T$ .  $\square$

**A.4 Remark:** If condition (2.7.2) of the Lemma A.2 alone holds, then  $A(t)$  will possess a uniform average, which is  $\bar{A}$ .  $\square$

In the sequel we provide a proof using the principle of integral smallness (condition 2 in Lemma A.2) [11] and "L-decomposition" technique (e.g. [1, Chapter 3]) for an extended version of the general averaging theorem for linear time varying differential equations. The extension consists of the fact that we can treat matrices  $\bar{A}$  which have eigenvalues with zero real part.

We need the following Lemma [11, Chapter 1]:

**A.5 Lemma:** Let  $F(t)$ ,  $B(t)$  be bounded and piecewise continuous matrix functions on  $\mathbb{R}^+$ , such that

$$\|F\|_{\infty} \leq M, \quad \|B\|_{\infty} \leq M$$

suppose that the fundamental matrix  $X(t)$  of

$$\dot{x}(t) = F(t)x(t) \tag{2.7.3}$$

satisfies

$$\|X(t)X^{-1}(s)\| \leq Ke^{-a(t-s)} \quad \forall t \geq s \tag{2.7.4}$$

for some  $a \in \mathbb{R}$  not  $\mathbb{R}^+$  and  $K \geq 1$ . If



$$\| \int_{t_1}^{t_2} B(t) dt \| \leq \delta \quad \forall |t_1 - t_2| < h \quad (2.7.5)$$

for some positive constants  $\delta$  and  $h$ , then the fundamental matrix  $Y(t)$  of the perturbed equation

$$\dot{y}(t) = (F(t) + B(t))y(t) \quad (2.7.6)$$

satisfies the inequality

$$\|Y(t)Y^{-1}(s)\| \leq (1+\delta)K e^{-b(t-s)} \quad \forall t \geq s$$

where

$$b = +a - 3MK\delta - (\log[(1+\delta)K])/h \quad (2.7.7) \square$$

**A.6 Remark:** It follows from this Lemma that stability of (3) (i.e.  $a > 0$ ) is preserved in (6) with  $b \geq a/2 > 0$  if  $h \geq h^* = \max(4(\log K)/a, 1)$  and  $3MK\delta + \log(1+\delta)/h^* \leq a/4$ , which always can be satisfied for  $\delta$  sufficiently small.  $\square$

**A.7 Corollary:** Let  $A(t)$  be a bounded and piecewise continuous matrix function with uniform average  $\bar{A}$ . Assume that

$$\|\bar{A}\| \leq M \text{ and } \|\bar{A} - A(t)\| \leq M$$

and that  $\bar{A}$  is a stability matrix (i.e.  $\operatorname{Re} \lambda_i(\bar{A}) < 0, \forall i$ ) such that

$$\|e^{\bar{A}(t-s)}\| \leq K e^{-a(t-s)} \quad \forall t \geq s$$

for some positive constant  $a$  and  $K \geq 1$ . Then there exists  $\mu^*$  positive such that for all  $\mu \in (0, \mu^*)$  the fundamental matrix  $X(t)$  of

$$\dot{x}(t) = \mu A(t)x(t) \quad (2.7.8)$$

satisfies

$$\|X(t)X^{-1}(s)\| \leq (1+\delta(\mu))K e^{-b(\mu)(t-s)}$$

where

$$a \geq b(\mu) \geq a/2$$

and  $\delta(\mu)$  is a monotonically increasing function of  $\mu$  ( $\delta(0) = 0, \delta(\mu) > 0$ ).  $\square$

Proof: Rewrite (2.7.8) as  $\dot{y}(\tau) = [\bar{A} + (A(\frac{\tau}{\mu}) - \bar{A})]y(\tau)$ , with  $y(\tau) = x(\frac{\tau}{\mu}) = x(t)$ .

Identify this equation as an equation of the form (2.7.6), with

$$B(\tau) = A(\frac{\tau}{\mu}) - \bar{A}$$

and

$$F(t) = \bar{A}.$$

By assumption we have:  $\exists K \geq 1$  and  $a > 0$ :

$$\|\bar{A}(t-\tau)\| \leq K e^{-a(t-\tau)} \quad \forall t \geq \tau.$$

Select then  $h = \max(4 \log K/a, 1)$  construct  $\delta_h(\mu) \equiv \delta(\mu)$  as in Lemma A.2 (for the matrix  $A(\tau/\mu)$ ).  $\mu^*$  is then defined as:

$$\sup \{ \mu > 0, \quad 3MK\delta(\mu) + \log(1+\delta(\mu))/h \leq a/4 \}$$

as suggested by Remark A.6. □

Integral smallness as expressed in Condition 2 in Lemma A.2 is preserved under multiplication with a signal having a "band limited" spectrum:

**A.8 Lemma:** Provided  $v(t)$  and  $\dot{v}(t)$  are piecewise continuous and bounded functions

$$\|v\|_{\infty} = V_0 \quad \text{and} \quad \|\dot{v}\|_{\infty} = V_1$$

and that  $B(t)$  is integral small:

$$\exists h > 0, \delta > 0: \quad \forall |t_1 - t_2| \leq h, \quad \left\| \int_{t_1}^{t_2} B(t) dt \right\| \leq \delta \quad (2.7.9)$$

$B(t)v(t)$  is integral small:

$$\forall |t_1 - t_2| \leq h, \quad \left\| \int_{t_1}^{t_2} B(t)v(t) dt \right\| \leq (V_0 + hV_1)\delta \quad \square$$

Proof: Integrating by parts we have:

$$\begin{aligned} \left\| \int_{t_1}^{t_2} B(t)v(t)dt \right\| &= \left\| \int_{t_1}^{t_2} B(t)dt \cdot v(t_2) - \int_{t_1}^{t_2} \left( \int_{t_1}^t B(\tau)d\tau \right) \cdot \dot{v}(t)dt \right\| \\ &\leq \delta V_0 + \delta V_1 h \end{aligned} \quad \square$$

Integral small signals are filtered out by "low" pass systems:

A.9 Lemma Provided  $A(t), B(t)$  be bounded and piecewise continuous functions on  $R^+$  with  $\|A\|_\infty \leq M$  and  $\|B\|_\infty \leq M$ . Assume that the fundamental matrix  $X(t)$  of (2.7.3) satisfies (2.7.4) with a positive and that  $B(t)$  is integral small, satisfying (2.7.9) then the solution of

$$\dot{x}(t) = A(t)x(t) + B(t), \quad x(t_0), \quad t \geq t_0 \quad (2.7.10)$$

satisfies

$$\|x(t)\| \leq K e^{-a(t-t_0)} \|x(t_0)\| + C\delta \quad (2.7.11)$$

with

$$C = K(1+M/a)/(1-e^{-ah}) \quad \square$$

Proof: Using the variation of constants formula, the solution of (2.7.10) can be written as:

$$x(t, t_0, x_0) = X(t)X^{-1}(t_0)x_0 + \int_{t_0}^t X(t)X^{-1}(s)B(s)ds \quad (2.7.12)$$

Define

$$C(\tau) = \int_{t_1}^{\tau} B(s)ds, \quad t_1 \text{ arbitrary}$$

Integrating by parts, we have that

$$\int_{t_1}^{t_2} X(t)X^{-1}(s)B(s)ds = X(t)X^{-1}(s)C(s) \Big|_{t_1}^{t_2} - \int_{t_1}^{t_2} A(s)X(t)X^{-1}(s)C(s)ds \quad (2.7.13)$$

Using the integral smallness of  $B$ , we have

$$\left| \int_{t_1}^{t_2} X(t)X(s)^{-1}B(s)ds \right| \leq \delta K \left(1 + \frac{M}{a}\right) e^{-a(t-t_2)} \quad (2.7.14)$$

$$\forall |t_1 - t_2| \leq h, \quad t_1 \leq t_2$$

Hence, for the last term in (2.7.12), we find using (2.7.13)

$$\begin{aligned} \left| \int_{t_0}^t X(t)X^{-1}(s)B(s)ds \right| &\leq \sum_{k=0}^{n-1} \left| \int_{t_0+kh}^{t_0+(k+1)h} X(t)X^{-1}(s)B(s)ds \right| \\ &\quad + \left| \int_{t_0+nh}^t X(t)X^{-1}(s)B(s)ds \right| \end{aligned} \quad (2.7.15)$$

where  $n$  is such that  $t_0+nh < t \leq t_0+(n+1)h$ . Using the estimate (2.7.14) in (2.7.15) yields the desired result.  $\square$

The first main averaging theorem is:

#### A.10 General Averaging Theorem I

Consider

$$\dot{x}(t) = \mu A(t)x(t) + B(t); \quad x(t_0) = x_0, \quad t \in \mathbb{R}^+ \quad (2.7.16)$$

where  $A(t)$  and  $B(t)$  are bounded, piecewise continuous matrix functions of  $t$  on  $\mathbb{R}^+$  with uniform averages  $\bar{A}$  and  $\bar{B}$  respectively. Let

$$M = \max(\|\bar{A}\|, \|\bar{B}\|, \sup_{t \in \mathbb{R}^+} \{\|A(t) - \bar{A}\|, \|B(t) - \bar{B} + \bar{A}^{-1}B(A(t) - \bar{A})\|\}).$$

Assume that  $\bar{A}$  is a stability matrix:

$$\|e^{-\bar{A}(t-s)}\| \leq Ke^{-a(t-s)} \quad \forall t \geq s$$

for some positive  $a$  and  $K \geq 1$ , then there exist a positive constant  $\mu^*$  and a monotonically increasing function  $\delta(\mu)$  ( $0 \leq \delta(\mu) \leq \delta(\mu^*)$ ,  $\delta(0)=0$ ) such that for all  $\mu \in [0, \mu^*)$

- (i) the homogenous part of (2.7.16) is exponentially stable;
- (ii) the solution of (2.7.16),  $x(t, t_0, x_0)$  satisfies:

$$\|x(t, t_0, x_0) - \bar{A}^{-1}\bar{B}\| \leq K(1+\delta(\mu))e^{-\mu b(\mu)(t-t_0)}\|x(t_0)\| + K(1-\delta(\mu))(1+M/b(\mu))/(1-e^{-b(\mu)h})\cdot\delta(\mu) \quad (2.7.17)$$

with

$$h = \max(4(\log K)/a, 1), \quad a/2 \leq b(\mu) \leq a \quad \square$$

Proof: Write (2.7.16) as:

$$\dot{y}(\tau) = [\bar{A} + (A(\frac{\tau}{\mu}) - \bar{A})]y(\tau) + D(\frac{\tau}{\mu}) \quad (2.7.18)$$

where

$$y(\tau) = x(\frac{\tau}{\mu}) - \bar{A}^{-1}\bar{B} = x(t) - \bar{A}^{-1}\bar{B}$$

$$D(\frac{\tau}{\mu}) = (B(\frac{\tau}{\mu}) - \bar{B}) + \bar{A}^{-1}\bar{B}(A(\frac{\tau}{\mu}) - \bar{A})$$

and apply Corollary A.7 to obtain the stability conclusion and then Lemma A.9 to obtain (ii).  $\square$

A.11 Remark: The term "general" refers to the fact that  $A(t)$  does not need to be periodic. In the case that  $A(t)$  and  $B(t)$  are periodic or a finite sum of periodic matrices then  $\delta(\mu)$  is of the order of  $\mu$ , and (2.7.17) reads as

$$\limsup_{t \uparrow \infty} \|x(t, t_0, x_0) - \bar{A}^{-1}\bar{B}\| = O(\mu) \quad \square$$

In order to handle unstable  $\bar{A}$  we need one more auxiliary result:

A.12 Lemma: Provided that  $F(t)$ ,  $B_1(t)$ ,  $B_2(t)$  are bounded and piecewise continuous matrix functions on  $\mathbb{R}^+$  with  $\|F\|_\infty \leq M$ ,  $\|B_1\|_\infty \leq M$  and  $\|B_2\|_\infty \leq M$ . Assume that the fundamental matrix of (2.7.3) satisfies (2.7.4) with a positive and that  $B_1(t)$ ,  $B_2(t)$  are integral small satisfying condition (2.7.9) then there exists a  $\delta^* > 0$  such that for all  $\delta \in [0, \delta^*)$  there exists a unique solution on  $\mathbb{R}^+$  to:

$$\dot{L}(t) = F(t)L(t) + L(t)B_1(t)L(t) + B_2(t), \quad L(0) = 0 \quad (2.7.19)$$

satisfying

$$\|L\|_\infty \leq M \quad \|L\|_\infty \leq M + M^2 + M^3 \quad (2.7.20)$$

with  $L \equiv 0$  if  $\delta = 0$ .  $\square$

Proof: Consider in the Banach space of continuous matrix functions  $L$  defined

on  $R^+$ , equipped with the supremum norm ( $\|L\|_\infty$ ) the following class of functions:

$$B = \{L \mid \|L\|_\infty \leq M\}.$$

Define the operator:

$$\begin{aligned} (Ty)(t) = & \sum_{k=1}^{n-1} X(t)X^{-1}(s)y(s)C_{k-1}(s)y(s) \Big|_{(k-1)h}^{kh} + X(t)X^{-1}(s)y(s)C_n(s)y(s) \Big|_{nh}^t \\ & - \sum_{k=1}^{n-1} \left\{ \int_{(k-1)h}^{kh} X(t)X^{-1}(s)y(s)C_{k-1}(s)y(s)ds \right. \\ & + \int_{(k-1)h}^{kh} X(t)X^{-1}(s)y(s)C_{k-1}(s)[F(s)y(s)+y(s)B_1(s)y(s)+B_2(s)]ds \\ & + \int_{(k-1)h}^{kh} X(t)X^{-1}(s)[F(s)y(s)+y(s)B_1(s)y(s)+B_2(s)]C_{k-1}(s)y(s)ds \\ & \left. + \int_0^t X(t)X^{-1}(s)B_2(s)ds \right\} \end{aligned}$$

where  $n$  is such that  $nh \leq t < (n+1)h$  and  $C_k$  is defined by :

$$C_k(s) = \int_{kh}^s B_1(t) dt \quad s \in [kh, (k+1)h)$$

Using the integral small property of  $B_1$ , it is not hard to verify that  $T$  is well defined on the class  $B$  and that it is a contraction on  $B$  for sufficiently small  $\delta$ . It follows that  $T$  has a unique fixed point. This fixed point is the solution of (2.7.19) satisfying (2.7.20) by construction.  $\square$

### A.13 General Averaging Theorem II

Consider

$$\dot{x}(t) = \mu A(t)x; \quad t \in R^+ \quad (2.7.20)$$

assume that  $A(t)$  is a piecewise continuous and bounded matrix function on  $\mathbb{R}^+$  with uniform average  $\bar{A}$ , satisfying

$$\|\bar{A}\| \leq M, \quad \|\bar{A} - A\|_{\infty} \leq M.$$

Then there exists a positive constant  $\mu^*$  such that for all  $0 < \mu < \mu^*$ :

- (i) (2.7.20) is strictly stable if  $\bar{A}$  is a stability matrix (i.e.  $\text{Re} \lambda_i(\bar{A}) < 0$ );
- (ii) (2.7.20) is unstable if at least one eigenvalue of  $\bar{A}$  has positive real part.  $\square$

**A.14 Remark:** This result extends a classical averaging result [9] because it does not require  $\bar{A}$  to be hyperbolic, i.e. to have no eigenvalue with zero real part. In particular the instability result holds even when  $\bar{A}$  has eigenvalues with zero real part. An extension of averaging treating general (bounded and regulated) matrix functions  $A(t)$  can be found in [18].  $\square$

#### Proof of the General Averaging Theorem II:

##### Step 1: Time Scaling $t = \tau$

Define  $y(\tau) = x(\frac{\tau}{\mu}) = x(t)$ , to write (2.7.20) as

$$\dot{y}(\tau) = [\bar{A} + (A(\frac{\tau}{\mu}) - \bar{A})]y(\tau) \quad (2.7.21)$$

##### Step 2: Co-ordinate Transformation

Let  $T$  be a nonsingular matrix which satisfies:

$$TAT^{-1} = \text{diag}(A_S, A_U)$$

where

$$\text{Re} \lambda(A_S) < 0 \text{ and } \text{Re} \lambda(A_U) > 0.$$

Define  $z(\tau) = Ty(\tau)$ , and also

$$T(A(\frac{\tau}{\mu}) - \bar{A})T^{-1} = B(\frac{\tau}{\mu})$$

We have then

$$\dot{z}(\tau) = (\text{diag}(A_S, A_U) + B(\frac{\tau}{\mu}))z(\tau)$$

##### Step 3: Decouple Stable and Unstable

Introduce the transformation:

$$\begin{bmatrix} z_S \\ z_U \end{bmatrix} = \begin{bmatrix} I & -L(\tau) \\ 0 & I \end{bmatrix} \begin{bmatrix} w_S \\ w_U \end{bmatrix} \quad (2.7.22)$$

where  $L$  is the unique, bounded solution of:

$$\dot{L} = A_S L - L A_U + L B_{US} \left(\frac{\tau}{\mu}\right) L + B_{SU} \left(\frac{\tau}{\mu}\right), \quad L(0) = 0$$

where  $z_S$ ,  $z_U$  and  $B_{SS}$ ,  $B_{SU}$ ,  $B_{US}$ ,  $B_{UU}$  are partitionings of  $z$  and  $B$  respectively, conforming to  $\text{diag}(A_S, A_U)$ . Lemma A.12 guarantees that  $L$  exists and is bounded. Because  $\text{Re}\lambda(A_S) < 0$  and  $\text{Re}\lambda(A_U) < 0$ , we have that the linear, homogeneous equation:

$$\dot{M} = A_S M - M A_U$$

is exponentially stable. Using this transformation, we obtain:

$$\begin{bmatrix} \dot{w}_S \\ \dot{w}_U \end{bmatrix} = \begin{bmatrix} A_S + L(\tau) B_{US} \left(\frac{\tau}{\mu}\right) & 0 \\ B_{US} \left(\frac{\tau}{\mu}\right) & A_U - B_{UU} \left(\frac{\tau}{\mu}\right) L(\tau) \end{bmatrix} \begin{bmatrix} w_S \\ w_U \end{bmatrix} \quad (2.7.23)$$

Notice in particular that (2.7.23) has the same stability properties as (2.7.21) ((2.7.22) is a Lyapunov transformation [19]). Because  $\|d/d\tau L(\tau)\|$  is bounded, Lemma A.8 assures that  $L(\tau) B_{US}(\tau/\mu)$  and  $B_{UU}(\tau/\mu) L(\tau)$  are integral small for sufficiently small  $\mu$ . From Lemma A.5 it follows then that the equation

$$\dot{w}_U = (A_U - B_{UU} \left(\frac{\tau}{\mu}\right) L \left(\frac{\tau}{\mu}\right)) w_U$$

is exponentially unstable. This proves the result.  $\square$

**A.15 Remark:** It is possible to go one step further in the decomposition to obtain:

$$\dot{w} = \begin{bmatrix} A_S + \dots & 0 & 0 \\ x & A_C + \dots & 0 \\ x & x & A_U + \dots \end{bmatrix} w$$

where  $\text{Re}\lambda(A_S) < 0$ ,  $\text{Re}\lambda(A_C) = 0$ ,  $\text{Re}(A_U) > 0$ . The result then demonstrates that part of the state decays exponentially, part explodes exponentially and the states corresponding to  $A_C$  can be either stable or unstable, depending on higher order ( $\mu^2$ ) effects.  $\square$

**A.16 Remark:** We believe that this proof is novel and relatively straightforward, and self contained, in that it only uses very basic principles of functional



analysis. We were unable to locate a theorem like A.13, except for [1, chapter 3], which however is proven along totally different lines and discusses only the periodic case or complete instability. For alternative approaches to averaging, see [3,9]. [3] uses a co-ordinate transformation approach presenting both stability and instability results, however, without details about estimating the errors. Reference [9] only discusses stability, and places more emphasis on finite time results. The Theorem 4.2.1 in [9, Chapter 4] discusses stability and is inadequate for our purpose, because of the (suspicious) lack of "uniform average". Both [3] and [9] discuss averaging for nonlinear systems.

A minor extension includes matrices of the form  $A(t, \mu)$ , bounded and piecewise continuous with respect to  $t$  and Lipschitz continuous with respect to  $\mu$  in a neighbourhood of the origin. If  $\bar{A}$  is the uniform average of  $A(t, 0)$  all the above results still apply, because the key Lemma A.2 is still valid.

**A.17 Lemma:** Let the matrix  $A(t, \mu)$  be bounded, piecewise continuous in  $t$ , Lipschitz continuous in  $\mu$ , with Lipschitz constant  $L$ :

$$|A(t, \mu) - A(t, 0)| < L\mu \quad \forall \mu \in [0, \mu^*] \quad \forall t$$

Assume that  $A(t, 0)$  has a uniform average  $\bar{A}$ . Under these conditions for any positive constant  $h$ , there exists a monotonically decreasing function  $\delta_h(\mu)$ ,  $\mu \in [0, \mu^*]$  such that:

$$\left| \int_{t_1}^{t_2} (A(t, \mu) - \bar{A}) dt \right| < \delta_h(\mu) \quad \forall |t_1 - t_2| < h \quad \square$$

**Proof:** Using Lemma A.2, we obtain:

$$\begin{aligned} & \left| \int_{t_1}^{t_2} (A(t, \mu) - \bar{A}) dt \right| \\ & \leq \left| \int_{t_1}^{t_2} (A(t, 0) - \bar{A}) dt \right| + \left| \int_{t_1}^{t_2} (A(t, \mu) - A(t, 0)) dt \right| \\ & \leq \delta_h(\mu) + Lh\mu = \delta_h(\mu), \quad \mu \in [0, \mu^*] \end{aligned}$$

where  $\delta_h(\mu)$  is constructed in Lemma A.2. □

In an adaptive control context we encounter systems of the following form:

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} A_{11}(t) & A_{12}(t) \\ \mu A_{21}(t) & \mu A_{22}(t) \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} B_1(t) \\ \mu B_2(t) \end{bmatrix} \quad (2.7.24)$$

having multiple time scales. Averaging is not directly applicable, but under the conditions that  $A_{11}(t)$  defines an exponentially stable system, one can reduce (2.7.24) to the standard form (2.7.16). We omit details, which can be found in [1, Chapter 3]. First, one separates slow and fast variables, using a Lyapunov transformation:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} I & -L(t, \mu) \\ 0 & I \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \quad (2.7.25)$$

where  $L(t, \mu)$  is the unique bounded solution of

$$L = A_{11}(t)L - A_{12}(t) - \mu LA_{22}(t) + \mu LA_{21}(t)L \quad (2.7.26)$$

which satisfies the boundary condition:

$$L(t, 0) = L_0(t)$$

$$L_0(t) = A_{11}(t)L_0(t) - A_{12}(t) \quad (2.7.27)$$

$L_0$  being the unique bounded solution of (2.7.27) defined on  $\mathbb{R}$  (analogous to a steady state solution). The transformation (2.7.25) brings (2.7.24) into the following block triangular form:

$$\begin{bmatrix} \dot{w}_1(t) \\ \dot{w}_2(t) \end{bmatrix} = \begin{bmatrix} A_{11}(t) + \mu L(t, \mu)A_{21}(t) & 0 \\ \mu A_{21}(t) & \mu(A_{22}(t) - A_{21}(t)L(t, \mu)) \end{bmatrix} \begin{bmatrix} w_1(t) \\ w_2(t) \end{bmatrix} + \begin{bmatrix} B_1(t) + \mu L(t, \mu)B_2(t) \\ \mu B_2(t) \end{bmatrix} \quad (2.7.28)$$

Stability of (2.7.24) can therefore be analysed, considering only a set of two decoupled equations:

$$\dot{Z}_1(t) = (A_{11}(t) + \mu L(t, \mu)A_{21}(t))Z_1(t) \quad (2.7.29)$$

$$\dot{Z}_2(t) = \mu(A_{22}(t) - A_{21}(t)L(t, \mu))Z_2(t) \quad (2.7.30)$$

The system (2.7.29) can be ensured to be stable for sufficiently small  $\mu$ , as  $A_{11}(t)$  defines a stable linear system, and (2.7.30) is in the standard form for averaging analysis. Using the General Averaging Theorem I, A.10, the following result emerges.

#### A.18 Theorem: Averaging in Multiple Timescale Case

Let the matrices  $A_{ij}(t), B_i(t)$   $i, j = 1, 2$  be bounded, piecewise continuous functions on  $\mathbb{R}^+$ , and let  $A_{11}(t)$  define a stable linear system. Assume that  $B_2(t)$  and the matrix  $A_{22}(t) - A_{21}(t)L(t, 0)$ , ( $L(t, 0)$  defined in (2.7.27)) have uniform averages  $\bar{B}$  and  $\bar{A}$  respectively. Then, there exists a positive constant  $\mu^*$  such that for all  $\mu \in (0, \mu^*)$ :

- (i) the system (2.7.24) is unstable if  $\bar{A}$  has at least one eigenvalue with positive real part;
- (ii) The system (2.7.24) is exponentially stable if  $\bar{A}$  is a stability matrix, and further more  $x_2(t, t_0, x_2(t_0))$  is close to  $\bar{A}^{-1}\bar{B}$ ; in the sense that  $\exists K \gg 1$ ,  $a(\mu) > 0$  and a monotonically increasing positive function  $\delta(\mu)$  ( $\delta(0)=0$ ) such that

$$\|x_2(t, t_0, x_2(t_0)) - \bar{A}^{-1}\bar{B}\| \leq K e^{-a(\mu)(t-t_0)} \|x_1(t_0) \times x_2(t_0)\| + \delta(\mu)$$

In the case where the matrices  $B_i(t), A_{ij}(t)$ ,  $i, j=1, 2$ , are periodic or a finite sum of periodic signals then  $\delta(\mu)=0(\mu)$ . □

## 2.8 REFERENCES

- [1] B.D.O. Anderson, R.R. Bitmead, C.R. Johnson Jr, P.V. Kokotovic, R.L. Kosut, I.M.Y. Mareels, L. Praly, B.D. Riedle, Stability of Adaptive Systems: Passivity and Averaging Analysis, M.I.T. Press, Boston, 1986.
- [2] P.V. Osburn, H.P. Whitaker, A. Kezer, 'New developments in the design of model reference adaptive control systems', IAS Paper, No.61-39, Feb. 1961.
- [3] J.K. Hale, Ordinary Differential Equations, Interscience Series on Pure Applied Mathematics, Vol.21, Wiley and Sons, New York 1969.
- [4] P.C. Parks, 'Lyapunov redesign of model reference adaptive control systems', IEEE Trans. Auto. Control, vol.AC-11, 1966, pp362-367.
- [5] D.J.G. James, 'Stability of model reference control systems', AIAA Journal, Vol.9, No.5, pp.950-952.
- [6] N.W. McLachlan, Theory and Application of Mathieu Functions, Clarendon Press, Oxford, 1947.
- [7] P. Hagedorn, Nonlinear Oscillations, Clarendon Press, Oxford, 1981.
- [8] B.D. Riedle, P.V. Kokotovic, 'A stability-instability boundary for disturbance free adaptation and unmodelled dynamics', IEEE Trans. on Auto. Control, vol.AC-30, pp1027-1030.
- [9] J.A. Sanders, F. Verhulst, Averaging in nonlinear systems, Applied Mathematics Vol.59, Springer Verlag, 1985.
- [10] C.A. Desoer, M. Vidyasagar, Feedback systems: Input-Output Properties, Academic Press, New York 1975.
- [11] W.A. Coppel, Dichotomies in Stability Theory, Springer Verlag, 1978, Lecture Notes on Mathematics Vol.629.
- [12] D.D. Donalson, C.T. Leondes, 'A model reference parameter tracking technique for adaptive control systems', IEEE Trans. on Application and Industry, Sept. 1963, pp241-261.

- [13] I.M.Y. Mareels, B.D.O. Anderson, R.R. Bitmead, M. Bodson, S.S. Sastry, 'Revisiting the M.I.T. rule for adaptive control', Proc. IFAC Workshop on Adaptive Systems in Control and Signal Processing, June 1986, pp.67-72.
- [14] M. Bodson, S.S. Sastry, B.D.O. Anderson, I.M.Y. Mareels, R.R. Bitmead, 'Nonlinear averaging theorems and the determination of parameter convergence rates in adaptive control', Systems and Control Letters, Vol 7, No3, May 1986, pp 145-159.
- [15] B.D.O. Anderson, I.M.Y. Mareels, 'Marrying frequency domain intuition and adaptive control', Proc. 3rd Conf. on Contr. Engineering, Sydney, 1986, pp66-72.
- [16] E. Trulsson, Adaptive control based on explicit criterion minimization, Ph. D. thesis, Linkoping (Sweden), 1983.
- [17] L. Fu, M. Bodson, S.S. Sastry, 'New stability theorems for averaging and their application to the convergence analysis of adaptive identification and control schemes', Proc. 24th IEEE CDC, Fort Lauderdale, 1985, pp473-477.
- [18] R.L. Kosut, B.D.O. Anderson, I.M.Y. Mareels, 'Stability theory for adaptive systems : methods of averaging and persistency of excitation', IEEE Trans. Auto. Contr., to appear, 1986.
- [19] F.R. Gantmacher, The Theory of Matrices, VolII, Chelsea Publishing Company, New York, N.Y., 1974.

### 3. BIFURCATIONS IN ADAPTIVE MODEL REFERENCE CONTROL

#### 3.1 Introduction

The adaptive feedforward control problem, discussed in the previous chapter, could be represented as a linear, time varying system whose dynamics depended in a very nonlinear way on the design and plant parameters. However, the linearity simplified the analysis by restricting the possible structurally stable dynamics to being either asymptotically stable or unstable. Using averaging techniques, we characterized sets of parameters describing the system, yielding either good adaptive performance (asymptotic stability) or unstable behaviour. In general however an adaptively controlled system is nonlinear and the above analysis can only be invoked to obtain information about the local dynamics of the system - in the neighbourhood of equilibria [1].

The second case study is chosen to emphasise the effect undermodelling has on the (global) dynamics of a nonlinear adaptive scheme. We choose to study a simple adaptive model reference scheme exhibiting a wealth of nonlinear phenomena. The plant to be controlled is a second order, but not necessarily stable, linear system. The controller is designed on the assumption that the system is first order, and the control task is that the plant output should track the output of a first order model. This leads to a three dimensional system of nonlinear differential equations, whose dynamics we study as a function of the relevant parameters: the plant parameters; the model parameters; the adaptation gain, and the parameters characterizing the reference input. This is the simplest example possible to exhibit a wide variety of nonlinear effects, which is still amenable to analysis such that we are able to recognise the fundamental mechanisms which cause these nonlinear, complex dynamics. Using perturbation techniques we then argue that the observed phenomena (Hopf bifurcation, homoclinic explosions) are not a product of this particular problem, but are rather generically present in adaptive control.

A complete analysis of the dynamical behaviour of a particular nonabstract nonlinear system as a function of the parameters describing it is a task far beyond purely analytical mathematical tools. Numerical techniques are indispensable tools in gaining a good understanding of the dynamics of a particular class of nonlinear systems of differential equations. Mathematicians of the 20th century have not only recognised but have also come (fatalistically)

to live with this fact. However, it is not our goal to go into the complete nitty-gritty details of the dynamics of the particular adaptive problem we are about the present. Such analysis is possible, but would lead us too far afield, and probably would generate too great a volume of results, not all interesting for the general adaptive control problem. (As a case in point, consider the study of the Lorenz equations presented in [2], which are in many respects much simpler than the particular adaptive control problem of the sequel, but required a 270 page book without fully exhausting their study!)

In Section 2 the problem is set up. Its basic and known features are concisely discussed in Section 3. The bulk of the analysis then follows in Sections 5 and 6, after we have introduced some terminology from global analysis (Section 4). The analytical results of sections 5 and 6 are supported by numerical experiments (Section 7), and using a blend of analytical results and numerical simulation evidence we extend our understanding of the global dynamics as a function of the relevant plant and design variables. The attention is focussed on structurally stable dynamics and on the mechanisms that cause changes in these dynamics. In Section 7 we discuss our results from an adaptive control point of view. An historical overview pinpointing the major results in model reference control and situating our results ends the chapter.

### 3.2 Problem Description

We describe a simple adaptive control problem of the model output tracking type in the presence of undermodelling. The plant to be controlled is assumed to be a causal, linear and time invariant, continuous time system with a transfer function belonging to the class  $P'$  (cf. Section 1.2, Chapter 1):

$$P' = \{Z_p(s) \mid Z_p(s) = \frac{p_1}{s^2 + p_1 s + p_1 p_2}, p_1, p_2 \in \mathbb{R} \ p_1 \neq 0\} \quad (3.2.1)$$

The plant can be represented in state space form as

$$\begin{aligned} \begin{bmatrix} \dot{x}_{1p} \\ \dot{x}_{2p} \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ -p_2 p_1 & -p_1 \end{bmatrix} \begin{bmatrix} x_{1p} \\ x_{2p} \end{bmatrix} + \begin{bmatrix} 0 \\ p_1 \end{bmatrix} u_p \\ y_p &= (1 \quad 0) \begin{bmatrix} x_{1p} \\ x_{2p} \end{bmatrix} \end{aligned} \quad (3.2.2)$$

The control objective is to track the output (state) of the first order system:

$$\dot{x}_m = -ax_m + r(t) \quad (3.2.3)$$

where  $a$  is a strictly positive constant and  $r(t)$  is a bounded piecewise continuous function of  $t$  - the reference input. The control designer assumes (mistakenly) that the plant can be adequately represented by a first order system with transfer function belonging to the model set:

$$P = \{Z_m(s) \mid Z_m(s) = \frac{1}{s+a_p}, \quad a_p \in R\} \quad (3.2.4)$$

and therefore implements the control law:

$$u_p(t) = -\theta(t)y_p(t) + r(t) \quad (3.2.5)$$

where  $\theta$  is the parameter estimate generated by a normalized gradient identification algorithm:

$$\dot{\theta}(t) = \frac{\epsilon y_p(t)(y_p(t) - x_m(t))}{(1 + \mu(y_p(t) - x_m(t))^2)} \quad (3.2.6)$$

$\epsilon$ ,  $\mu$  are positive constants;  $\epsilon$  is the adaptation gain and  $\mu$  is the normalization constant. Some modified algorithms which have been recently introduced [3,4,5] are of the form:

$$\begin{aligned} \dot{\theta} = & \epsilon y_p(t)(y_p(t) - x_m(t)) / (1 + \mu(y_p(t) - x_m(t))^2) \\ & - g(y_p(t), x_m(t), \theta) \end{aligned}$$

where  $g$  is a "fix-it" function. Some typical examples of fixes are:

$$g = -\gamma(\theta - \theta_0) \quad (3.2.7)$$

algorithm with exponential forgetting, biased towards  $\theta_0$  and

$$g = \gamma\sigma_1(\theta)\sigma_2(y_p(t) - x_m(t)) \quad (3.2.8.1)$$

with

$$\begin{aligned} \sigma_1(\theta) = 0 & \quad \text{if } |\theta - \theta_0| < \theta_1 \\ = \theta - \theta_0 - \theta_1 & \quad \text{if } \theta - \theta_0 > \theta_1 \\ = \theta - \theta_0 + \theta_1 & \quad \text{if } \theta - \theta_0 < -\theta_1 \end{aligned} \quad (3.2.8.2)$$

and



$$\sigma_2(e) = e^2 / (1 + \mu_2 e^2), \quad \mu_2 \geq 0 \quad (3.2.8.3)$$

which gives an error scaled exponential forgetting with dead zone of width  $2\theta_1$  centred at  $\theta_0$ . Most of our analysis deals with the situation  $g \equiv 0$ , the original model reference control law, but we indicate (some of) the changes in our results due to the presence of a fix function.

Remarks:

(R.3.1) Notice that the modelset  $P$  is contained in the class  $P'$  of allowable plants in a singular way, i.e. by allowing  $p_1 = \infty$ . In other words, for  $p_1$  large,  $1/s + p_2$  is a good low frequency approximation for  $p_1/(s^2 + p_1 s + p_1 p_2)$ .  $\square$

(R.3.2) The modifications (3.2.7) and (3.2.8) seem rather artificial from the control objective point of view, exponential forgetting (3.2.7) makes it even impossible to achieve zero tracking error (unless  $\theta_0$  happens to be the correct parameter value for zero tracking).  $\square$

(R.3.3) Notice the slightly nonstandard normalization used in (3.2.9). Normalizing by  $1 + \mu(y_p - x_m(t))^2 \geq 1$  serves however the same purposes as the more classical normalization of the regressor (here the output  $y_p$ ) by  $1 + \mu y_p^2$ . Because  $x_m(t)$  is by assumption a bounded signal, we have that

$$1 + \mu y_p^2 \leq K_1 (\|x_m(t)\|) (1 + \mu(y_p - x_m)^2)$$

and

$$1 + \mu(y_p - x_m)^2 \leq K_2 (\|x_m(t)\|) (1 + \mu y_p^2)$$

Hence both forms of normalization are equivalent. Any form of normalization allows us to demonstrate that no finite escape time can exist.  $\square$

The adaptive closed loop system can be described by a three dimensional set of ordinary first order, time varying, coupled nonlinear differential equations:

$$\dot{x}_1 = x_2 \quad (3.2.9.1)$$

$$\dot{x}_2 = -p_1 x_2 - p_2 p_1 x_1 - p_1 x_3 x_1 + p_1 r(t) \quad (3.2.9.2)$$

$$\dot{x}_3 = \epsilon x_1 (x_1 - x_m(t)) / (1 + \mu(x_1 - x_m(t))^2) - g(x_1, x_m(t), x_3) \quad (3.2.9.3)$$

where we identified

$$x_1 = x_{1p}, \quad x_2 = x_{2p}, \quad x_3 = \theta$$

and  $g$  is given by (3.2.7) or (3.2.8) and  $x_m(t)$  is defined by (3.2.3). For any given initial condition and any starting time there exists a unique solution well defined in the future: there is no finite escape time.

Lemma 3.1:

Assume that the normalization constant  $\mu$  is strictly positive. For any initial condition  $(x_{10}, x_{20}, x_{30})$  and  $t_0 \in \mathbb{R}^+$ , and bounded piecewise continuous function  $r(t)$ , defined on  $(t_0, +\infty)$  there exists a unique solution  $(x_1(t, t_0, x_{10}), x_2(t, t_0, x_{20}), x_3(t, t_0, x_{30}))$  starting at  $(x_{10}, x_{20}, x_{30})$  at  $t_0$  to the adaptive systems equation (3.2.9) defined on  $(t_0, \infty)$ .  $\square$

Proof: Using the definition of  $x_m(t)$  (3.2.3) and of  $g$  (3.2.7) or (3.2.8) it follows that:

$$|\dot{x}_3| \leq C_1 |x_3| + C_2$$

for some  $C_1, C_2 > 0$ , if  $\mu > 0$ . Hence, using (3.2.9)

$$|\dot{x}_1| \leq |x_2|$$

$$|\dot{x}_2| \leq C_3(t) |x_1| + C_4 |x_2| + C_5$$

for some exponentially overbounded function  $C_3(t)$ , and some  $C_4, C_5 > 0$ .  $\square$

In the subsequent sections we analyse the dynamics of the adaptive control problem (3.2.9) as a function of its relevant parameters:  $p_1, p_2$  the plant parameters, where  $p_1$  characterizes the undermodelling effect and  $p_2$  determines the d.c. gain of the plant;  $\epsilon$  the adaptation gain,  $a$  the model transfer function's pole and the different parameters necessary to characterize the reference input  $r(t)$ . The other parameters (the normalization constant  $\mu$ , and the various parameters describing the fixes) play only a secondary role.

### 3.3 Basic Properties of the Adaptive Response

In this section we concisely review the classical analytic results available in the literature for this particular adaptive control problem. The techniques involved are basically Lyapunov arguments and differ considerably from the methods used in the sequel (and so do the results).

The first result concerns the asymptotic zero tracking error ( $x_1(t) - x_m(t) \equiv 0$ ) property of the adaptively controlled system in the absence of modelling errors, using the unmodified algorithm ( $g \equiv 0$  cf. (3.2.9)). The closed loop can be represented as:

$$\dot{x}_1 = -p_2 x_1 - x_3 x_1 + r(t) \quad (3.3.1.1)$$

$$\dot{x}_3 = \epsilon(x_1 - x_m(t))x_1 / (1 + \mu(x_1 - x_m(t))^2) \quad (3.3.1.2)$$

which can be obtained from (3.2.9) by formally setting  $d/dt(x_2) \equiv 0$  and solving for  $x_2$ . The system (3.3.1) has the nontrivial solution:  $x_1(t) \equiv x_m(t)$ ,  $x_3(t) \equiv -p_2 + a$ .

**Theorem 3.1:**

Assume that  $r(t)$  is a piecewise continuous bounded function defined on  $\mathbb{R}^+$ , and that the model (3.2.3) is strictly stable ( $a > 0$ ). For any initial time  $t_0 \in \mathbb{R}^+$ , and any initial condition  $(x_{10}, x_{30})$  the adaptive system (3.3.1) has a unique solution  $(x_1(t, t_0, x_{10}), x_3(t, t_0, x_{30}))$  uniformly bounded on  $(t_0, \infty)$  which satisfies the asymptotic property:

$$x_1(t, t_0, x_{10}) \rightarrow x_m(t) \text{ as } t \uparrow \infty \quad (3.3.2)$$

If in addition  $r(t)$  is persistently exciting in the sense that there exists a positive constant  $r_0$ :

$$\begin{aligned} \forall t_1, t_2 \in \mathbb{R}^+, \quad t_1 < t_2 - 1, \quad \exists t_3, t_4, \quad t_1 < t_3 < t_4 < t_2 \\ \left| \int_{t_3}^{t_4} r(t) dt \right| > r_0 \end{aligned} \quad (3.3.3)$$

then  $x_1$  converges uniformly to  $x_m(t)$  and  $x_3$  converges uniformly to  $-p_2 + a$  as  $t$  increases.  $\square$

Proof: Define  $V(x_1, x_3, t)$  as:

$$V = \begin{cases} \frac{1}{2} \{ (x_1 - x_m(t))^2 + \frac{1}{\epsilon} (x_3 + p_2 - a)^2 \} & \text{if } \mu = 0 \\ \frac{1}{2} \left\{ \frac{1}{\mu} \ln(1 + \mu(x_1 - x_m(t))^2) + \frac{1}{\epsilon} (x_3 + p_2 - a)^2 \right\} & \text{if } \mu > 0 \end{cases}$$

Along the solutions of (3.3.1) the total derivative of  $V$  with respect to time is given by:

$$\dot{V}(3.3.1) = - \frac{e^2(t)}{1 + \mu e^2(t)} < 0 ; e = x_1(t) - x_m(t) \quad (3.3.4)$$

from which it follows that:

- (i)  $x_1(t, t_0, x_{10}), x_3(t, t_0, x_{30})$  are uniformly bounded, and well defined on  $(t_0, \infty)$
- (ii)  $e(t)$  and  $d/dt(e(t))$  are bounded and  $e(t)$  is square integrable, hence  $e(t)$  converges to zero.

The last claim of the Theorem follows at once from (3.3.4) if  $r(t)$  and  $x_m(t)$  are periodic. For the general case it follows from the uniform identifiability of the feedback parameter  $-p_2+a$  due to the persistently exciting  $x_m(t)$  - details can be found in [6,7].  $\square$

As it stands the stability property expressed by Theorem 3.1 with the assumption of persistently exciting input signals is robust with respect to small nonlinearities or small modelling errors, e.g. singular perturbations, i.e.  $p_1$  large compared with  $\epsilon$ ,  $a_1$ ,  $|p_2|$  and  $|d/dt(r(t))|$ . This robustness follows from a total stability argument [8].

Specifically, using the exponential forgetting modification (3.2.7) it is possible to show that bounded disturbances and singular perturbations of the plant do not destroy (local) stability, provided  $|d/dt(r(t))|$  is bounded [3]. More importantly, one can quantify a large region in state space for which the response is bounded and converges to a small residual set, of irremovable errors. Notice that with the exponential forgetting (3.2.7) it is impossible to match exactly the model output ( $e(t) \equiv 0$ ) (unless  $\theta_0 = -p_2+a$ ) even when no modelling errors are present, as the exponential forgetting pulls the parameter estimate towards  $\theta_0$ . The modification (3.2.8) which scales the exponential forgetting factor according to the error  $e(t) = x_1(t) - x_m(t)$  retains the property of exact matching in the absence of modelling errors and can achieve exact matching starting from any initial condition provided the excitation is sufficiently large as compared to the exponential forgetting. This scheme is discussed in detail in [5]. Using the techniques of [3] it is not hard to show that this algorithm is also robust with respect to singular perturbations ( $p_1$  large) in the sense that for a large set of initial conditions the adaptive response is bounded and the error (deviation from desired response) ( $e(t) = x_1(t) - x_m(t), x_2(t), x_3(t) - p_2+a$ ) converges to a small residual set of irremovable errors. It is possible to roughly quantify the set of initial conditions as well as the residual set in terms of the undermodelling parameter ( $p_1$ ). For  $p_1 \uparrow \infty$  the former tends to the whole state space, the latter collapses

to zero. (The analysis, we are about to present, justifies most of the claims made in the above.)

Observe that these results discuss the robustness of model reference control with respect to undermodelling of singular perturbation type, i.e.  $p_1$  large compared to  $\epsilon$  the gain of the adaptive scheme, the model pole and  $|p_2|$  which characterizes the low frequency behaviour of the plant and  $|d/dt(r(t))|$ . In this situation the plant has two real poles, one close to  $-p_2$ ,  $-p_2 + O(1/p_1)$  which can be adequately modelled, the other stable and large  $O(p_1)$  which is neglected. The requirement that  $|d/dt(r(t))|$  be small compared with  $p_1$  ensures that the natural time scale separation of the plant is preserved under adaptive feedback, hence the robustness result. In the subsequent analysis the assumption  $p_1$  is large is not made. The analysis focuses on the behaviour of the system dynamics as a function of its parameters, and a priori no parts of the parameter space are excluded.

### 3.4 Some Notions from Global Analysis [9,10]

Consider the nonlinear system of differential equations:

$$\dot{x} = f(x, p), \quad x(0) = x_0 \in \mathbb{R}^n, \quad p \in \mathbb{R}^m \quad (3.4.1)$$

$f$  is called a vector field, with parameter  $p$ . Assume that  $f$  is differentiable with respect to  $x$  and  $p$ .

The flow of (3.4.1) is a mapping from  $\mathbb{R}^n$  to  $\mathbb{R}^n$ , parametrized by time as  $\varphi_t(x_0) = x(t, x_0)$ , where  $x(t, x_0)$  is the solution of (3.4.1) starting in  $x_0$  (because the system is time invariant the initial time is irrelevant).

A fixed point  $x^*(p)$  is a solution of  $f(x, p) = 0$ . A fixed point  $x^*(p)$  is called hyperbolic if the Jacobian

$$Df(x, p) = \left[ \frac{\partial f_i}{\partial x_j}(x, p) \right] = (Df(x, p))_{(i, j)}$$

evaluated at  $x^*(0)$  has no eigenvalues with zero real part. If a fixed point is hyperbolic then the flow  $\varphi_t(x_0)$  in a neighbourhood of  $x^*(p)$  is well approximated by  $\exp(tDf(x^*(p), p))$  (Hartman-Grobman Theorem, see [9, p.13]). The fixed point  $x^*$  is a saddle if it is hyperbolic and if the Jacobian  $Df(x^*, p)$  has both eigenvalues with positive and negative real part. A stable (unstable) node is a fixed point  $x^*$  for which the Jacobian  $Df(x^*, p)$  has all eigenvalues with negative (positive) real part.

In the sequel we need a slightly stricter definition of structural stability than the one we have given in Chapter 1, Section 1.3. However, apart from its use in the forthcoming bifurcation analysis, we use the definition in Section 1.3 as the stricter definition cannot handle singular perturbations. Firstly we need the notion of equivalent vector fields. Two vector fields  $f$  and  $g$  are topologically equivalent if there exists a homeomorphism  $h$  which maps the flow of  $f$  ( $\varphi_t^f$ ) into the flow of  $g$  ( $\varphi_t^g$ ), not necessarily preserving parametrization by time, i.e. for any  $x$ , and  $t_1$ , there is a  $t_2$  such that:

$$h(\varphi_{t_1}^f(x)) = \varphi_{t_2}^g(h(x))$$

An  $\epsilon$ -perturbation of the vector field  $f$  is a (differentiable) vector field  $f^-(x,p)$ ,  $x \in \mathbb{R}^n$ ,  $p \in \mathbb{R}^m$  which satisfies:

$$\sup_{x \in K} (|f(\cdot, p) - f^-(\cdot, p)|) < \epsilon \quad \forall p$$

and

$$\sup_{x \in K} (|\frac{\partial}{\partial x} f(\cdot, p) - \frac{\partial}{\partial x} f^-(\cdot, p)|) < \epsilon \quad \forall p$$

for some compact set  $K \subset \mathbb{R}^n$ , and which equals  $f$  outside  $K$ .

Finally, a vector field is called structurally stable in the strict sense if all  $\epsilon$ -perturbations of  $f$  are topologically equivalent to  $f$ . (Notice that this definition has all the ingredients of the definition in Section 1.3, only is less versatile.)

A value  $p^* \in \mathbb{R}^m$  for which the flow of (3.4:1) is not strictly structurally stable is a bifurcation value of  $p$ .  $p$  is the bifurcation parameter. For example,  $p^*$  is a bifurcation value if  $Df(x^*(p), p)$  becomes singular for  $p = p^*$ , or has an eigenvalue with zero real part for  $p = p^*$ , and does not have this property for  $p \neq p^*$  in a neighbourhood of  $p^*$ . A Hopf Bifurcation occurs when for a parameter value  $p_0$  and equilibrium point  $x^*(p_0)$  the Jacobian  $Df(x^*(p_0), p_0)$  has a simple pair of pure imaginary eigenvalues and no other eigenvalues with zero real part. A bifurcation set is a locus in parameter space for which a particular bifurcation phenomenon (e.g. Hopf Bifurcation) occurs in the flow.

A bifurcation diagram is a locus in the combined state  $\times$  parameter space indicating the change of the asymptotic dynamics, e.g. a locus of the family of

limit cycles generated by the Hopf Bifurcation.

An hyperbolic fixed point  $x^*(p)$  has associated with it a stable manifold, the set of all points converging to  $x^*(p)$  in forward time, locally defined as

$$W_{loc}^S(x^*) = \{x \in U \subset \mathbb{R}^n \mid \varphi_t(x) \rightarrow x^* \text{ as } t \uparrow \infty, \\ \varphi_t(x) \in U, \forall t \geq 0\} \quad (3.4.2)$$

and globally extended as

$$W^S(x^*) = \bigcup_{t \leq 0} \varphi_t(W_{loc}^S(x^*))$$

and an unstable manifold, the set of points converging to  $x^*(p)$  in reverse time, locally defined by

$$W_{loc}^U(x^*) = \{x \in U \subset \mathbb{R}^n \mid \varphi_t(x) \rightarrow x^* \text{ as } t \downarrow -\infty, \\ \varphi_t(x) \in U \quad \forall t \leq 0\} \quad (3.4.3)$$

and globally extended as

$$W^U(x^*) = \bigcup_{t \geq 0} \varphi_t(W_{loc}^U(x^*)).$$

For an hyperbolic fixed point  $x^*$  the local stable manifold has the same dimension as (and is tangent to) the eigenspace of the linearized system  $\{d/dt(\xi) = Df(x^*, p)\xi\}$  corresponding to the eigenvalues with strictly negative real part of  $Df(x^*, p)$  mutatis mutandis for the unstable local manifold.

A trajectory connecting different fixed points is a heteroclinic orbit, a trajectory connecting a fixed point to itself (in a loop) is a homoclinic orbit. Heteroclinic and homoclinic orbits are structurally unstable phenomena in the flow of a nonlinear system which have an enormous impact on the flow. The bifurcation associated with a homoclinic orbit is called homoclinic explosion.

For the nonautonomous nonlinear differential equation

$$\dot{x} = f(t, x, p) \quad x(t_0) = x_0 \quad t, t_0 \in \mathbb{R} \quad x \in \mathbb{R}^n \quad p \in \mathbb{R}^m \quad (3.4.4)$$

a surface  $S_p$  in  $(t, x)$  space is an integral manifold of (3.4.4) if for any point  $B$  in  $S_p$  the solution  $x(t)$  of (3.4.4) passing through  $B$  is such that  $(t, x(t))$  remains in  $S_p$  for all  $t$  in the domain of definition of the solution  $x(t)$  [8].

The ordinary differential equation is called globally stable or Lagrange stable if for any bounded initial condition  $x_0$  and initial time  $t_0$  the solution

$x(t, t_0, x_0)$  is well defined and bounded in the future:

$$\forall |x_0| < C \quad \forall t_0 \quad \exists K(C, t_0) : |x(t, t_0, x_0)| < K \quad \forall t \geq t_0 \quad (3.4.5)$$

and uniformly globally stable if  $K$  is independent of  $t_0$ .

The idea behind bifurcation theory is to classify in topological terms the underlying mechanism that causes a certain change in the structure of the flow of a vector field. For instance, if the change of a parameter alters the flow of a vector field such that a stable node becomes unstable and instead a unique stable limit cycle appears (a Hopf-Bifurcation), then bifurcation theory says that the observed phenomenon is essentially equivalent to the bifurcation at the origin for  $\mu = 0$  of the planar vector field described by:

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} d\mu & \omega + c\mu \\ -\omega - c\mu & d\mu \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} (x^2 + y^2)(ax + by) \\ (x^2 + y^2)(bx + ay) \end{bmatrix} \quad \omega > 0$$

It is this concept that makes bifurcation theory an excellent tool to understand the complex dynamics of adaptive control.

For the reader unfamiliar with these concepts we include a simple example illustrating the different ideas [9]. Consider the scalar differential equation:

$$\dot{x} = \mu x - x^2 = (\mu - x)x.$$

For  $\mu < 0$ , the origin is a stable node and  $x = \mu$  is an unstable node. At  $\mu = 0$ , an exchange in stability properties occurs, the so called transcritical bifurcation. For  $\mu > 0$ , the origin is unstable and the fixed point  $x = \mu$  is a stable node. The bifurcation set is  $\{\mu = 0\}$ , the bifurcation diagram is depicted in Figure 3.1.

Finally we discuss briefly the notions of Poincaré map and integral manifold associated with a periodic orbit of the differential equation (3.4.1). A periodic orbit is a solution  $x(t, x_0)$  of (3.4.1) such that for some  $T > 0$   $x(t+T, x_0) = x(t, x_0)$ . The smallest such  $T$  is the period of the periodic orbit. An integral manifold associated with this periodic orbit is a cylinder in  $\mathbb{R} \times \mathbb{R}^n$ :

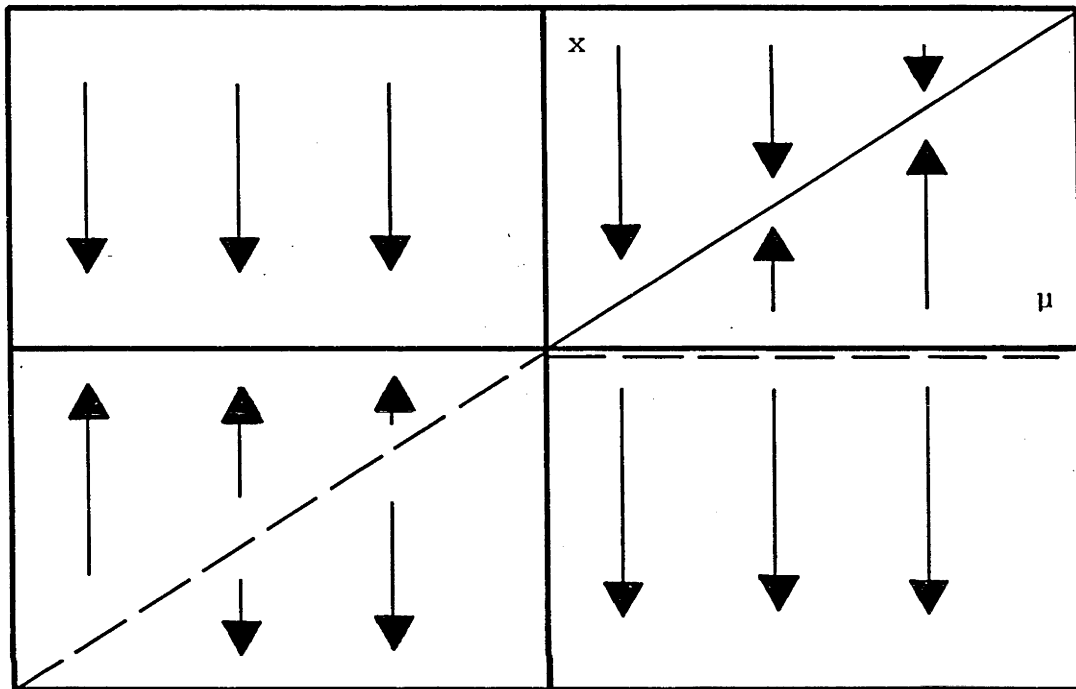
$$S_p = \mathbb{R} \times C_p = \mathbb{R} \times \{x \in \mathbb{R}^n \mid x = u_p(\varphi) \quad \varphi \in (0, T_p)\}.$$

Where we introduced the subscript  $p$  to emphasize that the periodic orbit and its period  $T_p$  depend on the particular value of the parameter  $p$ .  $C_p$  describes the locus in state space of the points on the periodic orbit  $u_p(\varphi)$ ,  $\varphi \in (0, T_p)$  is a parametrization for it. Consider a local cross section  $\Sigma$  (the return plane), i.e.



a subset of a hyperplane of dimension  $n-1$ , transversally intersected by  $C$  in a unique point  $x^*$  and such that locally the flow of (3.4.1) intersects  $\Sigma$  transversally. (cf. Figure 3.2.)

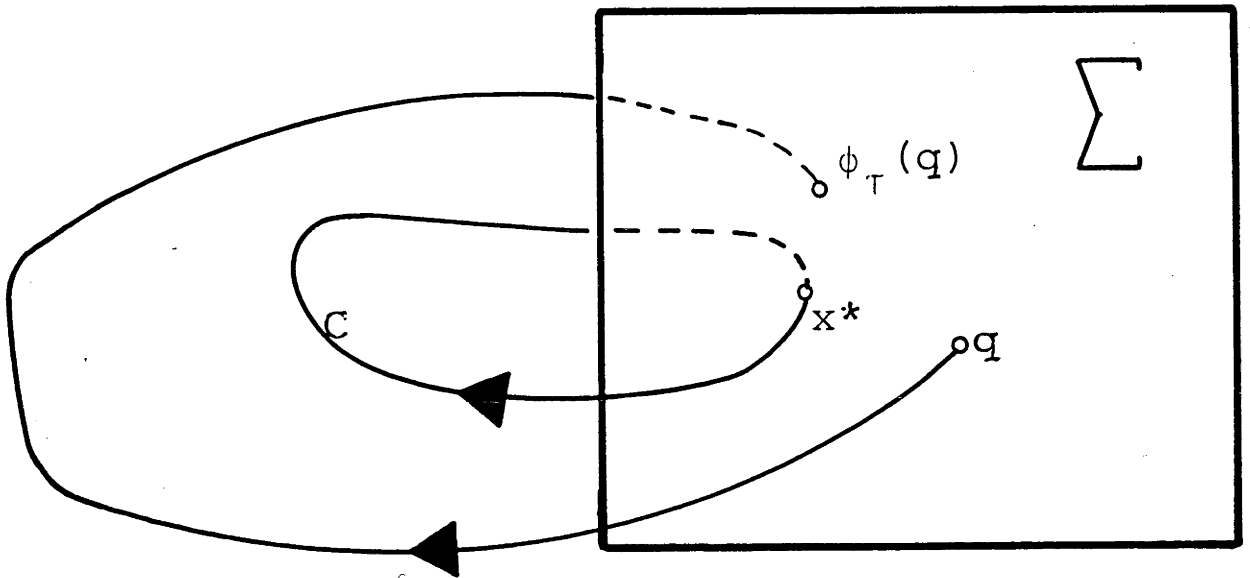
**Figure 3.1 Bifurcation Diagram for a Transcritical Bifurcation**



—— stable - - - - unstable ———> direction of attraction

The Poincaré map, or the first return map associates with each point  $q$  of  $\Sigma$  the point  $\varphi_\tau(q)$  where  $\tau = \tau(q)$  is the (shortest) time taken by the flow to return to  $\Sigma$  when started in  $q$ .

The point  $x^*$  is obviously a fixed point for the Poincaré map. The eigenvalues of the Jacobian of the Poincaré map evaluated at  $x^*$  correspond to  $n-1$  eigenvalues of  $\exp\{Df(x^*, p)T_p\}$  after deleting one eigenvalue which equals 1. (The existence of such an eigenvalue follows from the identity  $x^* = \exp\{Df(x^*, p)T_p\}x^*$ .) The Poincaré map reduces the study of a periodic orbit to the study of a fixed point (for a discrete map). In this way one can define hyperbolic periodic orbits, manifolds... by using the analogous concept for the fixed point of the Poincaré map.

Figure 3.2 First Return Map or Poincaré Map3.5 Adaptive System Dynamics for Constant Reference Input

Restricting the input  $r(t)$  to be constant simplifies the analysis considerably - the adaptive system (3.2.9) becomes an autonomous system - yet allows us to gain significant insight into the dynamics of the adaptive control problem, valuable even when the inputs are not constant. General reference inputs are dealt with in the next section.

For  $r(t) \equiv r$ , the adaptive control system (3.2.9) reduces to:

$$\dot{x}_1 = x_2 \quad (3.5.1.1)$$

$$\dot{x}_2 = -p_1 x_2 - p_1 p_2 x_1 + p_1 r - p_1 x_3 x_1 \quad (3.5.1.2)$$

$$\dot{x}_3 = \epsilon x_1 (x_1 - r/a) / (1 + \mu (x_1 - r/a)^2) - g(x_1, r/a, x_3) \quad (3.5.1.3)$$

where we have set  $x_m(t) \equiv r/a$ . For all nonzero  $r$ ,  $a$  and  $\epsilon$  the system (3.5.1) has a unique fixed point  $(r/a, 0, -p_2/a) = x^*$ , (if  $g \equiv 0$ ), its local stability

properties are discussed in the following Lemma.

Lemma 3.2:

The fixed point  $x^* = (r/a, 0, a-p_2)$  of the adaptive system (3.5.1) with  $g \equiv 0$  is locally uniformly asymptotically stable iff:

$$(i) \quad a > 0, \quad p_1 > 0, \quad \epsilon > 0 \quad (3.5.2)$$

$$(ii) \quad \epsilon r^2 < p_1 a^3 \quad (3.5.3)$$

□

Proof: The result follows directly from linearizing the system (3.5.1) around the fixed point. The Jacobian, evaluated at  $x^*$ , is given by:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ -p_1 a & -p_1 & -p_1 \frac{r}{a} \\ +\epsilon \frac{r}{a} & 0 & 0 \end{bmatrix} \quad (3.5.4)$$

The characteristic equation is given by:

$$\det(sI - A) = s^3 + p_1(s^2 + as + \epsilon \frac{r^2}{a^2}) \quad (3.5.5)$$

The Routh-Hurwitz test for stability of (3.5.5) - verifying that the polynomial (3.5.5) has only zeros with negative real part - yields the desired result.

Remarks:

(R.3.4) The fixed point is only unique when  $r \neq 0$ . This condition is a persistency of excitation condition for the model reference control scheme. The equilibrium  $x^* = (r/a, 0, a-p_2)$  corresponds to the desired response; the control objective is achieved because the plant output ( $x_1 = r/a$ ) tracks the model output ( $x_m(t) = r/a$ ). The parameter ( $x_3 = a-p_2$ ) is such that the d.c. gain of the controlled plant matches the d.c. gain of the model ( $1/a$ ). □

(R.3.5) Conditions (3.5.2) and (3.5.3) of Lemma 3.2 are independent of  $p_2$ , the d.c. gain of the plant. Not only are the local dynamics independent of  $p_2$ , but actually the complete adaptive behaviour is independent of  $p_2$ . This can be seen from the following description of the system (3.5.1) with  $g \equiv 0$ , in function of  $\tilde{x} = x - x^* = (x_1 - x_1^*, x_2, x_3 - x_3^*)$ :

$$\dot{\tilde{x}}_1 = \tilde{x}_2 \quad (3.5.6.1)$$

$$\dot{\tilde{x}}_2 = -p_1 a \tilde{x}_1 - p_1 \tilde{x}_2 - p_1 r/a \tilde{x}_3 - p_1 \tilde{x}_1 \tilde{x}_3 \quad (3.5.6.2)$$

$$\dot{\tilde{x}}_3 = \epsilon(\tilde{x}_1 + r/a)\tilde{x}_1/(1 + \mu\tilde{x}_1^2) \quad (3.5.6.3)$$

Of course, although the dynamics are independent of  $p_2$ , the actual adaptive signals do depend on  $p_2$  (cf.  $x^*$ ).  $\square$

(R.3.6) For the adaptive law using the estimation scheme with exponential forgetting  $g = \gamma(x_3 - \theta_0)$ , the fixed points are given by

$$x^e = (x_1^e, 0, -p_2 + \frac{r}{x_1^e}); \quad r \neq 0 \quad (3.5.7)$$

where  $x_1^e$  is a solution of

$$x_1^2(x_1 - \frac{r}{a}) + \frac{\gamma}{\epsilon}(-r + (p_2 + \theta_0)x_1)(1 + \mu(x_1 - \frac{r}{a})^2) = 0 \quad (3.5.8)$$

and the characteristic equation of the associated Jacobian is:

$$\lambda^3 + (p_1 + \gamma)\lambda^2 + p_1(\gamma + \frac{r}{x_1^e})\lambda + p_1\{x_1^e(2x_1^e - \frac{r}{a}) + \frac{\gamma r}{x_1^e}\} = 0 \quad (3.5.9)$$

A root locus argument shows that for  $(p_2 + \theta_0) > 0$  there is only one equilibrium point  $x^e$  with  $x_1^e$  belonging to the interval with endpoints  $r/a$  and  $r/(p_2 + \theta_0)$ . This point is a stable node provided:

$$p_1 > 0, \quad a < 0, \quad \epsilon > 0, \quad \gamma > 0, \quad \mu \geq 0 \quad (3.5.10.1)$$

$$\epsilon x_1^e(2x_1^e - r/a) \leq p_1^2 \frac{r}{x_1^e} + p_1\gamma^2 + \gamma^2 \quad (3.5.10.2)$$

Clearly for  $\gamma$  decreasing to zero, everything reduces to the situation discussed in the above. For  $\gamma > \epsilon$ ,  $x_1^e$  is heavily biased towards  $p_2 + \theta_0$  (cf. 3.5.8), which indicates large tracking errors. Worse still, for  $p_2 + \theta_0 < 0$ , there are three different equilibria for small  $\gamma$  ( $0 < \gamma < \gamma^*$ ) which can be stable or unstable, depending on the roots of the characteristic equation (3.5.9). The advantage of this scheme is that it can handle  $r = 0$ , or insufficient excitation. In this situation the fixed point is  $x^e = (0, 0, \theta_0)$ , which is stable provided  $p_2 + \theta_0 > 0$ . This advantage is therefore rather dubious, as one has to know a stabilizing control parameter in order to implement the adaptive control law in the absence of sufficient excitation! In the event of persistently exciting input  $r \neq 0$ , the local behaviour, which characterizes the performance of the adaptive scheme, is inferior to that of the unmodified algorithm. It can only be acceptable for  $\gamma \ll \epsilon$  and  $p_2 + \theta_0 > 0$ !  $\square$

(R.3.7) The adaptive control using the error scaled exponential forgetting

modification (3.2.3) preserves the fixed point  $x^*$  of the unmodified scheme, but may introduce two other equilibria. For  $g = \gamma(x_3 - \theta_0)(x_1 - r/a)^2 / (1 + \mu(x_1 - r/a)^2)$ , (no dead zone), the equilibria are given by

$$x^e = (x_1^e, 0, -p_2 + r/x_1^e)$$

where  $x_1^e$  is a solution of:

$$(x_1 - \frac{r}{a}) \left[ x_1^2 + \frac{\gamma}{\epsilon} (x_1 - \frac{r}{a}) (x_1 - \frac{r}{p_2 + \theta_0}) (p_2 + \theta_0) \right] = 0$$

Again for  $p_2 + \theta_0 > 0$  ( $\gamma, \epsilon, a > 0$ ) the only equilibrium is  $x_1 = r/a$ , provided  $\gamma$  is small, for  $p_2 + \theta_0 < 0$  and  $\gamma$  small there are two extra equilibria, besides  $x^*$ , one is stable, one is unstable. The fixed point  $x^*$  is always present, and has the same local stability properties as for the unmodified scheme. This scheme has the same (dubious) advantage of being able to cope with  $r = 0$  as the exponential forgetting modified scheme, but is able to perform as well as the unmodified algorithm when  $r(t)$  is persistently exciting.  $\square$

The conditions (3.5.2) in Lemma 3.2 are not only necessary for local stability, they are also necessary for global stability (cf. (3.4.5)) as we now will argue. It is clear that  $a$  needs to be positive otherwise the model would be unstable (this would also invalidate the replacement of  $x_m(t)$  by  $r/a$ ). The plant parameter  $p_1$  needs to be strictly positive, else the global response is unstable.

Lemma 3.3:

Assume that  $p_1$  is negative. The adaptive system (3.5.6) (or (3.2.9) with  $g \equiv 0$ ) is globally unstable in the sense that under the flow of the adaptive system any (nontrivial) volume of initial conditions is strictly expanding.  $\square$

Proof: It suffices to notice that the trace of the Jacobian of the vector field defined by (3.5.6) (or (3.2.9) with  $g \equiv 0$ ) is  $-p_1$ , independent of the state of the system.  $\square$

Remarks:

(R.3.8) The condition  $p_1 > 0$  implies that a plant belonging to  $P'(3.2.1)$  (with  $p_1 > 0$ ) can be stabilized by constant output feedback. This implies that for the particular situation of undermodelling described in Section (3.2), the unmodified model reference control law can stabilize any plant in  $P'$  which could be stabilized with constant output feedback, were the parameters known. In other words, the adaptive control problem is well posed.  $\square$

(R.3.9) Lemma 3.3 does not hold for the modified schemes, because the trace

of the Jacobian becomes  $-p_1 - (\partial g / \partial x_3)$ . This indicates however that these modified algorithms cannot stabilize a plant for which  $p_1 < -\gamma$  for the exponential forgetting scheme and  $p_1 < -\gamma/\mu$  for the error scaled exponential forgetting modification.  $\square$

Also  $\epsilon$ , the adaptation gain, should be positive, otherwise the parameter estimate  $x_3$  may remain negative, which leads to instability. More precisely, the following result holds.

**Lemma 3.4:**

Suppose that  $\epsilon < 0$ ,  $\mu > 0$ ,  $p_1 > 0$ ,  $a > 0$ . For parameter values satisfying these conditions, the adaptive system (3.5.6) is not Lagrange stable, in the sense that for all initial conditions  $x_1(0) > 0$ ,  $x_2(0) > 0$ , and  $x_3(0) < -a$  the solutions of (3.5.6) are unbounded.  $\square$

**Proof:** The proof is given for the case  $r > 0$ ,  $r < 0$  can be dealt with along the same lines.

We demonstrate that trajectories starting from  $x_1(0) = x_{10} > 0$ ,  $x_2(0) = x_{20} > 0$  and  $x_3(0) = x_{30} < -a$  satisfy  $x_1(t) > x_{10}/2$  and  $x_2(t) > 0$  and  $x_3(t) < -a$ ,  $\forall t \geq 0$ , which enables us to conclude that  $x_3(t)$  tends to  $-\infty$  and that  $x_1(t)x_2(t)$  tends to  $+\infty$ . The proof is by contradiction.

Assume that  $t_1 = \inf\{t \geq 0, x_1(t) \leq x_{10}/2\} < \infty$ . Notice that by continuity  $t_1 > 0$ . Because  $x_1(t) > x_{10}/2 \forall t \in [0, t_1)$  we have from equation (3.5.6.3) that:

$$\dot{x}_3(t) \leq -c_1 \text{ some } c_1 > 0 \quad \forall t \in [0, t_1),$$

hence

$$x_3(t) \leq -c_1 t + x_{30} < -a \quad \forall t \in [0, t_1).$$

Computing the derivative of  $x_1(t)x_2(t)$  along the trajectories of (3.5.6) we find:

$$\begin{aligned} \frac{d}{dt} x_1(t)x_2(t) &= x_2^2(t) - p_1(a+x_3(t))x_1^2(t) \\ &\quad - p_1x_1(t)x_2(t) - p_1\frac{r}{a}x_3(t)x_1(t) \end{aligned}$$

Because

$$x_2^2(t) - p_1(a+x_3(t))x_1^2(t) - p_1\frac{r}{a}x_3(t)x_1(t) \geq c_2 + c_3t > 0$$

for some  $c_2 > 0$ ,  $c_3 > 0$  and all  $t \in [0, t_1)$  it follows that

$$x_1(t)x_2(t) \geq c_4 + c_5t > 0 \quad \forall t \in [0, t_1)$$

for some  $c_4$  and  $c_5 > 0$ . As all solutions are defined on  $[0, +\infty)$ , it follows that  $\bar{x}_2(t) > 0 \quad \forall t \in [0, t_1)$  and hence, using (3.5.6.1)  $\bar{x}_1(t) > x_{10} \quad \forall t \in [0, t_1)$  contradicting the definition of  $t_1$ . Consequently  $\bar{x}_1(t) > x_{10}/2$ ,  $\bar{x}_3(t) < -c_1 t + x_{30}$  and  $\bar{x}_1(t)\bar{x}_2(t) > c_4 + c_5 t \quad \forall t > 0$ .  $\square$

Having convinced ourselves that condition (3.5.2) is necessary for both local and global stability, we now consider the situation where the local stability condition (3.5.3) fails; assuming that (3.5.2) holds. In this case a soft loss of stability occurs, i.e. the equilibrium ceases to be stable and instead an asymptotically stable limit cycle appears.

### Theorem 3.2: Hopf Bifurcation

Assume that  $a > 0$ ,  $\epsilon > 0$ ,  $p_1 > 0$  and  $\mu > 0$ . Consider  $b = (\epsilon, r, p_1, a)$  as bifurcation parameter. A Hopf bifurcation occurs for any bifurcation parameter  $b$  on the bifurcation locus  $H$  described by:

$$H = \{b : \epsilon r^2 = p_1 a^3\} \quad (3.5.11.1)$$

There exists a neighbourhood  $U$  of  $H$ :

$$U \subset \{b : \epsilon r^2 > p_1 a^3\} \quad (3.5.11.2)$$

such that to any bifurcation parameter  $\bar{b}$  in  $U$  there corresponds a unique limit cycle for the unmodified adaptive system (3.5.1) which is locally uniformly asymptotically stable. The limit cycle is at a distance  $O(d(\bar{b}, H)^{1/2})$  of  $x^*$  and has a period  $2\pi/\sqrt{ap_1} + O(d(\bar{b}, H)^{1/2})$ , where  $d(\bar{b}, H)$  denotes the distance between  $\bar{b}$  and  $H$ , i.e.  $d(\bar{b}, H) = \inf\{\|b - \bar{b}\|; b \in H\}$ .  $\square$

### Remark:

(R.3.10) In terms of the adaptive control problem this means that when the parameters of the system are such that condition (3.5.1)  $\epsilon r^2 = p_1 a^3$  is (slightly) violated ( $>$ ), the control objective cannot be realised. (Recall that the input  $r(t)$  is constant  $r(t) \equiv r$ .) Instead an asymptotically periodic non-zero tracking error remains, whilst also the adaptive control parameter is oscillating. The period of the oscillation is (at the onset of limit cycle behaviour) completely determined by the model pole and  $p_1$ , the amplitude depends on how much (3.5.11) is violated.  $\square$

(R.3.11) Except for the asymptotic stability conclusion in Theorem 3.2, the

Theorem is also valid for the scheme modified with error scaled exponential forgetting, as this scheme has the same fixed point with the same Jacobian as the unmodified scheme. (For the adaptive algorithm with exponential forgetting, in the situation that  $p_2 + \theta_0 > 0$ , a Hopf Bifurcation occurs when (3.5.10.2) is violated, cf. Remark 3.6.) In principle we could verify the asymptotic stability in these situations also, but the algebraic manipulations involved are too tedious to be worth the effort. Simulations confirm that the limit cycle generated by the Hopf bifurcation in the latter case is uniformly asymptotically stable.  $\square$

### Proof of Theorem 3.2:

#### Step 1: Continuity Properties

The application of the Hopf-Bifurcation Theorem [9,11,12,13] requires that the vector field defined by (3.5.1) is  $C^4$  in both parameter and state. Here the vector field is  $C^\infty$  in both parameter and state.

#### Step 2: Eigenvalues at Bifurcation Value

The Jacobian of (3.5.1) ( $g \equiv 0$ ) evaluated at  $x^* = (r/a, 0, -p_2 + a)$  has two complex conjugate, purely imaginary eigenvalues and one real negative eigenvalue if  $\epsilon r^2 = p_1 a^3$ . Indeed, under this condition, one can factorize the characteristic equation (3.5.5) as  $(s^2 + p_1 a)(s + p_1)$ .

#### Step 3: Nondegenerate Hopf Bifurcation

The Hopf Bifurcation is nondegenerate when the complex conjugate eigenvalues cross the imaginary axis transversally at criticality. This step guarantees the existence of a unique limit cycle corresponding to each  $\bar{b}$  in  $U$ . Denote by  $\lambda(b)$  that eigenvalue of the Jacobian which equals  $i\sqrt{ap_1}$  for  $b \in H$ . We have that for all  $b \in H$ :

$$\left. \frac{\partial \lambda}{\partial \epsilon} \right|_b = \frac{r^2}{2a^3 - 2ia^2 \sqrt{ap_1}}; \quad \text{Re} \left. \frac{\partial \lambda}{\partial \epsilon} \right|_b > 0 \quad (3.5.12.1)$$

$$\left. \frac{\partial \lambda}{\partial r} \right|_b = \frac{\epsilon r}{a^3 - ia^2 \sqrt{ap_1}}; \quad r \cdot \text{Re} \left. \frac{\partial \lambda}{\partial r} \right|_b > 0 \quad (3.5.12.2)$$

$$\left. \frac{\partial \lambda}{\partial p_1} \right|_b = \frac{ia}{2\sqrt{ap_1} - 2p_1 i}; \quad \text{Re} \left. \frac{\partial \lambda}{\partial p_1} \right|_b < 0 \quad (3.5.12.3)$$

$$\left. \frac{\partial \lambda}{\partial a} \right|_b = \frac{-2p_1 + i\sqrt{ap_1}}{2a - 2i\sqrt{ap_1}}; \quad \text{Re} \left. \frac{\partial \lambda}{\partial a} \right|_b < 0 \quad (3.5.12.4)$$



## Step 4: Asymptotic Stability

Establishing the uniform asymptotic stability of the limit cycles generated by a Hopf Bifurcation requires the computation of the so called curvature coefficients. This is a rather tedious, computationally-involved task. The result is:

$$\text{Re}\Psi = -\frac{3}{4}\frac{a}{(a+p_1)^2} - \frac{a^3 p_1}{r^2(a+p_1)} \left( 7 + \frac{p_1 \sqrt{p_1 a + p_1 + 2a + 1}}{p_1 + 4a} \right) \quad (3.5.15)$$

with

$$\epsilon r^2 = p_1 a^3; \quad a > 0, \quad p_1 > 0, \quad \epsilon > 0$$

valid for all  $b \in S$ . In all circumstances,  $\text{Re}\Psi < 0$  which together with (3.5.12) implies asymptotic stability. ( $\Psi$  is calculated according to the formula given by Poore [12], see also [11], which we found in these circumstances the easiest to use. Details can be obtained from the author on serious request.  $\square$ )

Remarks:

(R.3.12) This Theorem 3.2 establishes existence, uniqueness and the local uniform stability for a family of limit cycles, for bifurcation parameters  $b$  in a neighbourhood of the bifurcation locus. Simulation experiments show that these limit cycles persist in a large region of the parameter space and further more are not only locally but appear to be globally uniformly asymptotically stable (see Section 3.7).  $\square$

(R.3.13) The condition  $\epsilon r^2 < p_1 a^3$  (3.5.3) can be interpreted as delineating the region of "slow" adaptation, it is only when this condition is violated that the typical nonlinear effects (here limit cycles) manifest themselves. The condition is best interpreted as follows:

$$\epsilon \left( \frac{r}{a} \right)^2 < p_1 a \quad (3.5.3)$$

The effective adaptive gain (product of the algorithm gain  $\epsilon$  and the excitation level  $(r/a)^2 \equiv x_m^2$ ) must be strictly less than the bandwidth of the model (which in ideal circumstances would be the bandwidth of the controlled plant as well) scaled by the degree of undermodelling. ( $p_1 = \infty$  corresponds to the ideal situation,  $p_1 \ll 0$  violates the necessary condition (3.5.2)). This is completely

analogous to the high gain instability mechanism of the M.I.T. rule (see R.2.7, section 2.3.1).  $\square$

### 3.6 Adaptive System Response for General Reference Input

In the previous section we analysed the adaptive control problem introduced in Section 3.2 for constant inputs and obtained results about the dynamics of the adaptive system as a function of the parameters describing it. Are our findings a miraculous product of the simplifications we introduced, or do the bifurcation phenomena persist, at least qualitatively, when the input is not a constant, or when small nonlinearities are present in the plant or if the plant is not second order but higher order? In other words, is this bifurcation effect structurally stable? Because of the local uniform asymptotic stability of the fixed point or the limit cycles this question can be answered in the affirmative.

We first focus our attention on inputs  $r(t) = \cos \omega t$  for which it is still possible to achieve the control objective. Then we exploit the structural stability of our previous result to demonstrate that, although we analysed a specific problem, our results are relevant for a large class of adaptive control problems.

#### 3.6.1 Reference input $r(t) = \cos \omega t$

Although only a scalar feedback law (3.2.5) is implemented, the model reference control algorithm (3.2.9) can achieve model output tracking even if  $r(t) = \cos \omega t$ . Indeed, it is not hard to verify that the trajectory  $(x_1^*(t), x_2^*(t), x_3^*(t))$  defined as

$$x_1^*(t) = x_m(t) = (\omega^2 + a^2)^{-\frac{1}{2}} \cos(\omega t - \tan^{-1}(\omega/a)) \quad (3.6.1.1)$$

$$x_2^*(t) = \dot{x}_m(t) = -\omega(\omega^2 + a^2)^{-\frac{1}{2}} \sin(\omega t - \tan^{-1}(\omega/a)) \quad (3.6.1.2)$$

$$x_3^*(t) = -p_2 + a + \omega^2/p_1 \quad (3.6.1.3)$$

is a solution of the adaptive system (3.2.9) ( $g \equiv 0$ ).

#### Remark:

(R.3.12) This is a quite remarkable feature of the model reference control algorithm which is a result of the particular structure of undermodelling discussed here. Notice that the constant feedback gain (3.6.1.3) is such that the second order plant reacts as a first order plant at the frequency  $\omega$ .

$(Z_p(s)/(Z_p(s)x_3^* + 1) = Z_m(s)$  for  $s = j\omega$ . Similar behaviour is possible in more complicated situations.  $\square$

Linearizing the adaptive system in a neighbourhood of the trajectory  $(x_1^*(t), x_2^*(t), x_3^*(t))$ , denoting the deviations as  $y_i(t) = x_i(t) - x_i^*(t)$ ,  $i=1,2,3$ , we obtain:

$$\begin{bmatrix} \dot{y}_1(t) \\ \dot{y}_2(t) \\ \dot{y}_3(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ -p_1(a + \frac{\omega^2}{p_1}) & -p_1 & -p_1 x_m(t) \\ \epsilon x_m(t) & 0 & 0 \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{bmatrix} \quad (3.6.2)$$

where  $x_m(t) = (\omega^2 + a^2)^{-1/2} \cos(\omega t - \tan^{-1}(\omega/a))$  (cf. (3.6.1.1)). The following local stability result is immediate.

**Lemma 3.5:** Assume  $p_1 > 0$ ,  $a > 0$ ,  $\mu > 0$ ,  $\epsilon > 0$ . The trajectory  $(x_1^*(t), x_2^*(t), x_3^*(t))$  defined by (3.6.1) is locally uniformly asymptotically stable for all  $\omega$  and  $\epsilon$  sufficiently small,  $\epsilon \in (0, \epsilon^*(\omega, a))$ .  $\square$

**Proof:** Using the averaging results of Chapter 2, Theorem 2.1, there exists an  $\epsilon^*(\omega, a) > 0$  such that (3.6.2) is uniformly asymptotically stable for all  $\epsilon \in (0, \epsilon^*)$  provided that

$$\alpha = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x_m(t) [Z(s)x_m](t) dt > 0 \quad (3.6.3)$$

where

$$Z(s) = p_1 / (s^2 + p_1 s + p_1 a + \omega^2)$$

After some algebra, we obtain

$$\alpha = a / [2(\omega^2 + a^2)^2] > 0 \quad \forall \omega \in \mathbb{R}, \quad \forall a \in \mathbb{R}_0^+$$

Hence the result.  $\square$

#### Remarks:

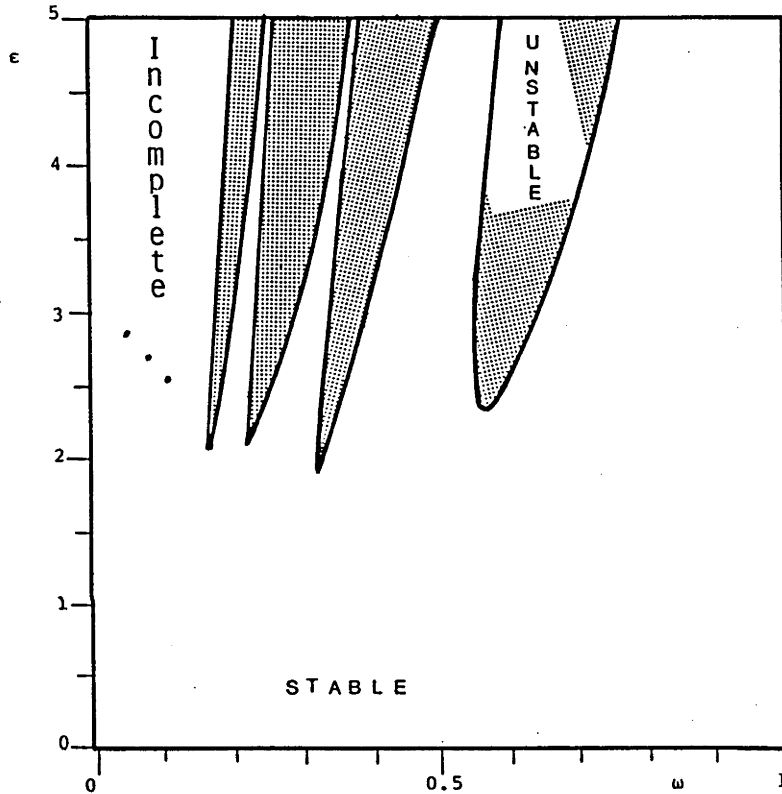
(R.3.13) From (R.2.21) we recall that  $\epsilon^*(\omega, a)$  is roughly proportional to  $\alpha$ , hence the stability gain margin ( $\epsilon^*$ ) reduces quickly with increasing frequency.  $\square$

(R.3.14) The complete stability analysis of (3.6.2) can only be achieved via numerical integration using Floquet Theory, (cf. Figures 2.3, 2.4 and 2.5 in Chapter 2). In this situation only the high gain instability - discussed in the

previous section - and the resonance phenomena are possible, because the adaptive controller eliminates the model error at the frequency  $\omega$  of the input. For non sinusoidal inputs instability due to model errors is possible.  $\square$

(R.3.15) In Figure 3.3 we display the instability/stability boundary in the parameter plane  $(\epsilon, \omega)$  for the linearized system described by equation (3.6.2).

### 3.3 Stability/Instability Boundary In Parameter $(\epsilon, \omega)$ Plane



The Figure 3.3 is obtained in the same way Figures 2.3, 2.4 and 2.5 in Chapter 2 were obtained. (The Figure is incomplete in the sense that it does not display the stability/instability boundary in the small frequency region  $\omega < 0.2$ , due to the extreme sensitivity of the stability properties with respect to small changes in the frequency in this region, (cf. Figure 2.4 in Chapter 2.) On part of the stability/instability boundary displayed in Figure 3.3 the dynamics undergo a Hopf type bifurcation. In this context this means that the origin becomes locally unstable in the sense that two of the characteristic multipliers (Floquet Theory [8, pp.117-121]) are complex conjugate and cross the unit circle, whilst the other characteristic multiplier is  $\pm e^{-p_1 T}$ ,  $T = 2\pi/\omega$ . This bifurcation phenomenon is extremely complex, and in some sense similar to the

Hopf Bifurcation encountered when  $r(t) \equiv r$ , only, the stability analysis including the nonlinear effects of the adaptive system in the neighbourhood of this bifurcation is prohibitive, and the new dynamics which emerge are difficult to describe (see [9, pp.162-163]). Notice in particular that one has earlier observed [1, Chapter 3, Figure 3.8, section 3.7.4] that in the situation where the averaging theory predicts local instability that the adaptive response was globally stable. The Hopf Bifurcation Theorem 3.2 (together with these observations) is a possible explanation for this phenomenon.  $\square$

(R.3.16) The same type of analysis can be carried out for the modified schemes, again the exponential forgetting modification (3.2.8) is more difficult to analyse, and appears not to perform as well as the error scaled exponential forgetting modification (3.2.9). For the latter the above analysis holds unchanged.  $\square$

### 3.6.2 General Reference Inputs

For general inputs (not simply sinusoidal) it becomes impossible to match the plant output with the model output. As it is difficult to identify, or establish the existence of nontrivial bounded solutions, we limit ourselves to input signals  $r(t)$  which in some sense can be regarded as a perturbation of a constant signal  $r$ , or of a purely sinusoidal signal.

We develop the results only in detail for  $r(t)$  close to a constant signal, indicating how it can be generalised to the sinusoidal case.

Firstly, we rewrite the adaptive system's equations (3.2.9) with  $g \equiv 0$  (no fixes) in the following equivalent form, more suitable for the subsequent analysis:

$$\begin{bmatrix} \dot{\tilde{x}}_1 \\ \dot{\tilde{x}}_2 \\ \dot{\tilde{x}}_3 \end{bmatrix} = \begin{bmatrix} \tilde{x}_2 \\ -p_1(a\tilde{x}_1 + \tilde{x}_2 + \frac{r}{a}\tilde{x}_3 + \tilde{x}_1\tilde{x}_3) \\ \epsilon \frac{\tilde{x}_1(\tilde{x}_1 + r/a)}{1 + \mu\tilde{x}_1^2} \end{bmatrix} + \begin{bmatrix} 0 \\ -p_1(r(t) - r) \\ \epsilon \Delta(\tilde{x}_1, x_m(t)) \end{bmatrix} \quad (3.6.4.1)$$

where  $\Delta$  is defined as:

$$\Delta(\tilde{x}_1, x_m(t)) = \frac{(\tilde{x}_1 + \frac{r}{a})(1 + \mu\tilde{x}_1(\tilde{x}_1 + \frac{r}{a} - x_m(t)))}{(1 + \mu\tilde{x}_1^2)(1 + \mu(\tilde{x}_1 + \frac{r}{a} - x_m(t))^2)} (x_m(t) - \frac{r}{a})$$

where  $(\bar{x}_1, \bar{x}_2, \bar{x}_3) = (x_1 - x_1^*, x_2 - x_2^*, x_3 - x_3^*)$  and  $x^* = (x_1^*, x_2^*, x_3^*) = (r/a, 0, a - p_2)$ , which is the fixed point for  $r(t) \equiv r$ . Obviously (3.6.4) is of the form:

$$\dot{\bar{x}} = f(\bar{x}) + F(t, \bar{x}), \quad (3.6.5)$$

which we consider as a perturbation of the system

$$\dot{\bar{x}} = f(\bar{x}) \quad (3.6.6)$$

corresponding to the constant input situation discussed in Section 3.5. The following two results are direct consequences of "classical" small parameter Theorems:

**Theorem 3.3:**

Assume that  $\epsilon > 0$ ,  $p_1 > 0$ ,  $a > 0$  and  $\epsilon r^2 < p_1 a^3$ , and that  $r(t)$  is a bounded, piecewise continuous function on  $R^+$ , which satisfies the integral small condition:

$$\left| \int_{t_1}^{t_2} (r(t) - r) dt \right| < \delta \quad \forall t_1, t_2 \in R^+ : |t_1 - t_2| < h \quad (3.6.7)$$

for some  $\delta$  and  $h$  positive. Under these conditions there exists a positive constant  $\delta_1$  such that for all  $\delta \in (0, \delta_1)$  the unmodified adaptive system (3.2.9) (or (3.6.4)) has a unique and locally uniformly asymptotically stable solution  $x^*(t, \delta)$  (or  $\bar{x}^*(t, \delta)$ ) in a neighbourhood of  $x^*$  (or 0). This solution depends continuously on  $\delta$  and satisfies  $x^*(t, 0) = x^*$  (or  $\bar{x}^*(t, 0) = 0$ ).  $\square$

**Proof:** Under the given assumptions the unperturbed system (3.6.5) has a locally uniformly asymptotically stable fixed point (0) by virtue of Lemma 3.2. The result then follows via total stability [8] upon noting that the perturbation  $F(t, \bar{x})$  satisfies:

$$\left| \int_{t_1}^{t_2} F(t, \bar{x}) dt \right| \leq C_1 \cdot \delta \quad \forall t_1, t_2 \in R^+, |t_1 - t_2| < h \quad (3.6.8)$$

uniformly in  $\bar{x} \in R^3$ , for some  $C_1 = C_1(\epsilon, \mu, a)$ .  $\square$

Remarks:

(R.3.17) This result is a statement about the dynamics of the adaptive system as a function of its parameters, local in state space, non local in parameter space. (Here local stands for in the neighbourhood of the nominal response.) The result is easily extended to incorporate other nonlinearities as long as condition (3.6.8) is satisfied.  $\square$

(R.3.18) This result establishes the existence of nontrivial solutions for the adaptive system in the situation that exact matching is impossible due to a nonsinusoidal reference input. It indicates therefore a constructive way for finding the so-called tuned solution [1] based on finding an input function close to the actual reference for which exact matching is possible (here a sinusoidal input) and obtaining the parameter setting for the controller from this input function. In the present case study this yields a unique parameter  $x_3$  or  $\theta$ , but in general for multidimensional parameter in the controller this is not the case. Also in the multidimensional situation it becomes a nontrivial issue to select the approximation function such that a stabilizing controller emerges.  $\square$

(R.3.19) Notice that the class of input functions (3.6.7) contains a subclass of fast time varying reference signals with mean  $r$  (cf. [14, Ch.1], and Lemma A.1 in the Appendix of Chapter 2).  $\square$

Completely similar results can be derived for the situations where the parameter values are such that the unperturbed system has a (locally) asymptotically stable limit cycle (Hopf Bifurcation Theorem 3.2). For clarity we recall the notation established in Theorem 3.2. The bifurcation parameter is  $b = (\epsilon, r, p_1, a)$ . The Hopf Bifurcation locus  $H$  is characterized by  $\epsilon r^2 = p_1 a^3$ , and the neighbourhood of  $H$  for which asymptotically stable periodic orbits appear in the flow of the unperturbed system (3.6.6) is denoted by  $U$  (cf. (3.5.6.2)). One cannot expect that the perturbed system (3.6.5) will possess a stable limit cycle (even for the smallest of non stationary perturbations). The best one can hope for is that it possesses a locally uniformly asymptotically stable integral manifold in the neighbourhood of the integral manifold (cylinder) of the unperturbed system. In control terms this means that although the control objective cannot be achieved due to both too large an adaptive gain and nonconstant inputs, the response remains bounded and close to the nominal periodic response. This fact is expressed in the next two theorems. We denote the integral manifold associated with the limit cycle

in the unperturbed system by (for all  $b \in U$ ):

$$S_{b,0} = R \times C_b = R \times \{x \in R^3 : x = u_b(\varphi) \varphi \in [0, T_b]\} \quad (3.6.9)$$

$C_b$  is the locus in state space of the limit cycle, which is parametrized by the function  $u_b$ ;  $T_b$  is the period of the periodic orbit. The subscript  $b$  emphasizes that the limit cycle (and its period) depend on the particular value of the bifurcation parameter  $b \in U$ . The following results hold:

**Theorem 3.4:**

Denote  $b = (\epsilon, r, p_1, a)$ , assume  $b \in U$  as in Theorem 3.2 (3.5.7.2). The unperturbed system (3.6.6) has a uniformly asymptotically stable limit cycle with corresponding integral manifold (3.6.9).

Assume that the reference input is a piecewise continuous function defined on  $R$  and satisfies

$$|r(t) - r| < \delta \quad \forall t \in R \quad (3.6.10)$$

Under these conditions there exist a positive constant  $\delta_1$  and a neighbourhood  $N_b$  of  $C_b$  such that the perturbed system (3.6.5) has for all  $\delta \in [0, \delta_1)$  an integral manifold  $S_{b,\delta}$  defined on  $R \times N_b$ , parametrized as:

$$S_{b,\delta} = \{(t, x) \in R \times N_b : x = u_b(\varphi) + v_b(t, \varphi, \delta) \\ (\varphi, t) \in [0, T_b] \times R\} \quad (3.6.11)$$

which reduces to  $S_{b,0}$  for  $\delta = 0$ , ( $v_b(t, \varphi, 0) = 0$ ).  $v(t, \varphi, \delta)$  is almost periodic ( $T$ -periodic) in  $t$  if  $r(t)$  is almost periodic ( $T$ -periodic) in  $t$ .  $S_{b,\delta}$  is locally uniformly asymptotically stable.  $\square$

**Proof:** Follows from the Hopf Bifurcation Theorem 3.2 and Theorem 7.1 in [8, pp. 244].  $\square$

**Theorem 3.5:**

Under the assumptions of Theorem 3.4, but with a continuous almost periodic reference input  $r(t)$  with mean  $r$ , i.e.

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T r(t) dt = r$$



there exists a positive constant  $\omega_1$  and a neighbourhood  $P_b$  of  $C_b$  such that the perturbed system (3.6.5) with input  $r(\omega t)$ ,  $\forall \omega \gg \omega_1$  has an integral manifold  $S_{b,\omega}$  defined on  $R \times P_b$ , parametrized as:

$$S_{b,\omega} = \{(t, x) \in R \times P_b: x = u(\varphi) + v_b(\omega t, \varphi, \frac{1}{\omega}) \\ (t, \varphi) \in R^+ \times [0, T_b]\} \quad (3.6.13)$$

which is uniformly asymptotically stable and reduces to  $S_{b,\infty} = R \times C_b$  as  $\omega \uparrow \infty$ . The function  $v_b(\omega t, \varphi, 1/\omega)$  has the properties that  $v_b(\omega t, \varphi, 1/\omega) \rightarrow 0$  as  $\omega \uparrow \infty$  and  $v_b$  is almost periodic in  $t$ .  $\square$

Proof: Follows from the Hopf Bifurcation Theorem 3.2 and Theorem 7.2 in [8, pp. 245].  $\square$

Remarks:

(R.3.20) Theorems analogous to Theorems 3.3 and 3.4 can be established for  $r(t)$  close to a sinusoidal signal, on the basis of Lemma 3.3. This extends our knowledge about the dynamics in parameter space considerably (small  $\epsilon$ , but extra parameter: frequency of sinusoidal reference signal).  $\square$

(R.3.21) Theorems 3.4 and 3.5 can be amended to include unmodelled nonlinear terms in the plant, as long as they are small in the neighbourhood of the desired response.  $\square$

(R.3.22) Theorems 3.3, 3.4 and 3.5 are readily adapted to see that the model reference control algorithm is robust with respect to small input noise and output noise. If input and output of the plant are corrupted by noise as:

$$u_p(t) = -x_3(t)x_1(t) + r(t) + w(t)$$

$$y_p(t) = x_1(t) + v(t)$$

It follows, after some algebraic manipulations, that for sufficiently small  $\delta$  the model reference control algorithm is stable if:

$$\left| \int_{t_1}^{t_1+1} w(t) dt \right| < \delta, \quad \left| \int_{t_1}^{t_1+1} v(t) dt \right| < \delta \quad \forall t \in R^+$$

$$\int_{t_1}^{t_1+1} v^2(t) dt < \delta, \quad \left| \int_{t_1}^{t_1+1} (x_m(t) - r/a) v(t) dt \right| < \delta \quad \forall t \in R^+ \quad \square$$

### 3.6.3 Robustness of the Bifurcation Phenomena

Using "classical" singular perturbation techniques one can establish that Hopf Bifurcations are persistent under singular perturbations [15,8]. We use this fact to establish that the results obtained previously are not a (miraculous) consequence of the particular control problem we have set up in Section 3.2.

Consider the following class of plants:

$$P^* = \{Z_p(s) \mid Z_p(s) = Z_1(s)Z_2(\mu s); \mu \in \mathbb{R}^+; Z_1(s) \in P^* \text{ and } Z_1(0) = 1, Z_2 \text{ is strictly stable and proper}\} \quad (3.6.14)$$

The plant  $Z_p(s)$  is basically second order, with some extra fast (if  $\mu$  is small) decaying modes. Using the methodology of Section 3.2 we arrive at a closed loop system which can be represented as:

$$\mu \dot{z} = Az + Br(t) \quad (3.6.15.1)$$

$$\dot{x}_1 = x_2 \quad (3.6.15.2)$$

$$\dot{x}_2 = -p_1 p_2 x_1 - p_1 x_2 + p_1 (C^T z + dr(t)) \quad (3.6.15.3)$$

$$\dot{x}_3 = \epsilon (x_1 - x_m(t)) x_1 / (1 + \mu (x_1 - x_m(t))^2) \quad (3.6.15.4)$$

where  $(A, b, c, d)$  is a minimal realization for  $Z_2$ :

$$Z_2(\mu s) = C^T (\mu s I - A)^{-1} B + d.$$

Thus  $A$  is a stability matrix and  $-C^T(A)^{-1}B + d = 1$  by assumption. The system (3.6.15) is in the standard form for the application of the main results in [15] about singularly perturbed Hopf Bifurcations. The following result holds for constant input  $r(t) \equiv r$  and  $x_m(t) \equiv r/a$ .

#### Theorem 3.6:

Assume that  $\epsilon > 0$ ,  $a > 0$ ,  $p_1 > 0$ ,  $\mu > 0$ ,  $r(t) \equiv r$ , and that  $A$  is a stability matrix. Denote  $\bar{b} = (\epsilon, r, p_1, a) \in H$  (cf. Theorem 3.2) if  $\epsilon r^2 = p_1 a^3$ . Denote  $x^* = (-A^{-1}Br, r/a, 0, -p_2 + a)$ ,  $x = (r, x_1, x_2, x_3)$ . Under these assumptions, there exists a  $\mu_0$  positive, such that for all  $\mu \in [0, \mu_0]$  the full system (3.6.15) undergoes a Hopf Bifurcation at an equilibrium  $x_\mu^*$  near  $x^*$  for a bifurcation value  $\bar{b}_\mu$  near  $\bar{b}$ . Moreover, there

exists a neighbourhood  $U_\mu$  of the bifurcation locus  $H_\mu$  such that for every  $\bar{b}_\mu \in U_\mu$  the limit cycle is locally uniformly asymptotically stable.  $\square$

Proof: Verify hypotheses H1-H4 of [15] and apply Theorems 3 and 4 of [15].  $\square$

Remarks:

(R.3.23) More complicated singular perturbations can be accommodated, but yield little extra insight. The above result describes just one class of adaptive problems where a Hopf Bifurcation occurs. We are not claiming that Hopf Bifurcations may be present in a generic adaptively controlled system, but that it is indicative of possible behaviour.  $\square$

(R.3.24) In a very similar way the integral manifold results of Theorems 3.4 and 3.5 can be amended to cope with singular perturbations as in (3.6.15). (Chapter VII in [8])  $\square$

### 3.7 Numerical Experiments

The analytical results concerning the dynamics of the adaptive control problem (3.2.9) discussed in the previous sections deal with:

- (1) The ideal situation, where the plant belongs to the model set ( $p_1 = \infty$ , formally). (See Section 3.3.)
- (2) The case where the plant can be well approximated by a first order system in the low frequency range; the fast mode is being neglected in the control design. (See Sections 3.3 and 3.6.)
- (3) The general undermodelling situation, the plant is second order and possibly cannot be adequately modelled by a first order system. (See Sections 3.5 and 3.6.)

Concisely, the information obtained about the dynamics is respectively:

- (1) The adaptive system is globally stable and the control is asymptotically optimal. The results indicate little information about the robustness of these properties.
- (2) For a large region of initial conditions the adaptive system's response is bounded with residual (small) errors. The results are valid for a region in parameter space characterized by  $p_1$  large compared to  $a$ ,  $\epsilon$ ,  $|p_2|$  and  $|d/dt(r(t))|$ . Little or no information is available about the performance (local dynamics, residual errors).
- (3) The local dynamics are precisely characterized over large regions of

the parameter space, providing detailed information about the performance and robustness of the local dynamics. However, little information is obtained about the global dynamics.

Clearly the above results complement each other, but there are some hiatuses. In this section, besides illustrating the above results, we fill in these gaps. Firstly, we argue using a blend of numerical and analytical results that the local dynamics discussed previously are the asymptotic dynamics for a large region of initial conditions (closing the gap between local and global results). Secondly, and more importantly, we explore further the possible dynamics over the parameter space discovering new bifurcation phenomena.

### 3.7.1 Local Dynamics?

Before presenting simulation evidence indicating that the local dynamics discussed in Sections 3.5 and 3.6 are indeed the asymptotic dynamics for a large set of initial conditions in the state space of the adaptive system (3.2.9), we make the following important observation about the global dynamics in general:

**Lemma 3.6:** Assuming that  $p_1 > 0$ , the flow of the adaptive system (3.2.9) is volume contracting.  $\square$

**Proof:** It suffices to observe that the trace of the Jacobian of the vector field defined in (3.2.9) is uniformly in state space less than or equal to  $-p_1 < 0$ . (This is the case for both the unmodified and modified control schemes.)  $\square$

One says that the global dynamics are uniformly hyperbolic. In particular Lemma 3.6 implies that any asymptotic dynamical behaviour is restricted to, at most, a two-dimensional subset of the state space.

Figure 3.4 illustrates that the local dynamics are the asymptotic dynamics for a large set of initial conditions. For these simulations the system parameters were set as follows:

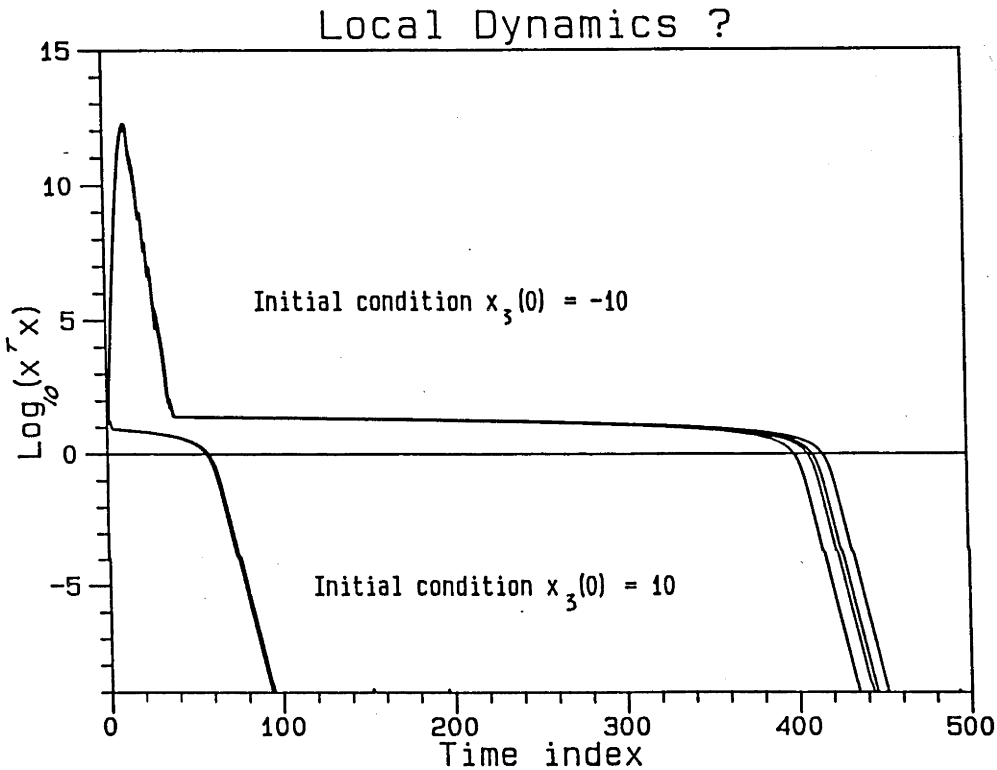
$$\epsilon = 1, \quad r = p_1 = a = 2, \quad p_2 = -\frac{3}{2}, \quad \mu = 1$$

which corresponds to the situation of a second order plant with poles at  $s = -1$  and  $s = 3$ , where the neglected pole has the same order of magnitude as the pole of the model ( $s = -a = -1$ ). The control objective was to track the constant model output  $r/a = 1$ . Notice that for the above parameter settings the desired response is locally uniformly asymptotically stable ( $\epsilon r^2 < p_1 a^3$ ). Figure 3.4 displays the decay of the logarithm of the norm  $(\bar{x}^T \bar{x})^{\frac{1}{2}}$  as a

function of time,  $\bar{x} = (x_1 - r/a, x_2, x_3 + p_2 - a)$  for nine initial conditions:  $\bar{x}_0 = (\delta_1, \delta_2, \delta_3)$ ,  $\delta_i = +10, -10$ ;  $i = 1, 2, 3$ .

Figure 3.4 demonstrates that the transient behaviour can be quite unacceptable! Completely analogous results are obtained for sinusoidal inputs.

Figure 3.4 Global Response (Transient response)



Remarks:

(R.3.25) Although it is easy to find initial conditions for which the simulations yield numerical overflow errors (on a Vax 11/780, working in quadruple precision), it is necessary in order to see this phenomenon to introduce errors in the initial conditions by selecting a destabilizing feedback parameter setting  $x_{30}$ . For large negative values of  $x_{30}$ , (such that  $a + \bar{x}_{30} = -p_2 + x_{30} < 0$ ) the initial response  $(\bar{x}_1, \bar{x}_2, \bar{x}_3)$  blows up very fast, which drives  $\bar{x}_3$  to positive values (possibly yielding overflow  $x_{30} < -100$ ). Then, once the feedback gain is "stabilizing" ( $\bar{x}_3 \geq a$ ),  $\bar{x}_1$  and  $\bar{x}_2$  the plant states decay quickly (in an oscillatory way), approximately converging to:

$$\dot{x}_1 = - \frac{\dot{x}_3 \frac{r}{a}}{\dot{x}_3 + a} \quad \dot{x}_2 \approx 0$$

which implies (cf. (3.2.9)) that  $\dot{x}_3$  decays. Once  $\dot{x}_3$  becomes small the local dynamics take over, for  $\epsilon r^2 < p_1 a^3$  this implies convergence to the desired settings. For  $\epsilon r^2 > p_1 a^3$  the above cycle repeats itself as the local dynamics are unstable yielding increasingly complex behaviour the more the parameters deviate from  $\epsilon r^2 = p_1 a^3$ .  $\square$

(R.3.26) Exhaustive simulations indicate that in all cases where the local dynamics are uniformly asymptotically stable, the adaptive controller gave good performance (including the transient behaviour) for initial errors up to 100% (relative to the desired response).  $\square$

### 3.7.2 Illustrating the Local Results

When the adaptive gain is too large (e.g.  $\epsilon r^2 > p_1 a^3$  for  $r(t) \equiv r$ ) and/or the input is not a purely sinusoidal signal the control objective cannot be achieved anymore, but the response remains bounded. The following figures illustrate this point, focussing our attention especially on the performance deterioration, in the asymptotic response.

Figures 3.5.1 and 3.5.2, display the asymptotic output tracking error ( $y_p(t) - x_m(t)$ ) as a function of time for  $r_1(t) = 1 + 0.1 \cos t$  and  $r_2(t) = 1 - 1.41 \cos 20t$  for the system described by the parameters  $p_1 = p_2 = 0.5$ ,  $a = 1$  and  $\epsilon = 0.25$ . In this situation exact tracking is achieved for  $r(t) \equiv 1$ . These figures illustrate Theorem 3.3. Notice that both inputs  $r_1(t)$  and  $r_2(t)$  satisfy the condition:

$$\left| \int_{t_1}^{t_2} (x_{m,i}(t) - 1) dt \right| < 0.15 \quad \forall t_1, t_2: |t_1 - t_2| < 1$$

(cf. (3.6.7)). Notice that the output error magnitude is about 10% of the desired response.

Figure 3.6.1 displays the output error, as a function of time for the inputs  $r_0(t) \equiv 1$  and  $r_1(t) = 1 + 0.1 \cos 0.1t$  for the same system as above, but with increased adaptive gain  $\epsilon = 1.0$ . In this case ( $\epsilon r^2 > p_1 a^3$ ) the control objective cannot be achieved even for  $r_0(t) \equiv 1$ , a periodic error remains. The extra

sinusoidal component in  $r_1(t)$  modulates this response. Figure 3.6.2 illustrates the same response in the state space, projected into the plane of output-error and feedback gain. Notice in particular how the integral manifold (for the  $r_1(t)$  response) wraps around the periodic orbit (response for  $r_0(t)$ ). These pictures illustrate the Hopf Bifurcation Theorem 3.2 and the integral manifold Theorem 3.4.

A typical response for the system described by the parameters  $p_1 = p_2 = a = \mu = \epsilon = 1$  and with sinusoidal input  $r(t) = \cos 0.5 t$  is displayed in Figure 3.7.1 and for  $\epsilon = 4.5$  in Figure 3.7.2. For  $\epsilon = 1$ , the control objective is achieved,  $\epsilon = 4.5$  proves to be too large, a tracking error remains, (cf. Figure 3.3). These figures illustrate Lemma 3.5 and Remark (R.3.14).

In Figures 3.6 and 3.7.2 one can easily identify the dominant frequencies in the tracking error, the fast oscillation is due to the Hopf Bifurcation, the slow modulation is due to the input excitation.

Remarks:

(R.3.27) These pictures illustrate the robustness of the model reference controller very well. Indeed the first order model  $Z_m(s) = 1/s+1$  is really a very bad model for the plant  $Z_p(s) = 0.5/s^2+0.5s+0.25$ , which has an oscillatory impulse response! This demonstrates that the original model reference scheme can have excellent robustness properties with respect to undermodelling provided a modest control task is imposed: the plant output should only be required to track a predominantly slowly time varying model output, as outlined in Theorems 3.3 and 3.4 and Lemma 3.5.  $\square$

(R.3.28) An intuitively appealing interpretation for the limit cycling phenomenon (as outlined in the Hopf Bifurcation Theorem) is that it reminds us of hunting. Due to the high adaptation gain, the adaptive controller wants to achieve too much too quickly and therefore keeps on hunting after the good setting, pumping energy into the plant, causing oscillations.  $\square$

(R.3.29) The same simulations have been repeated for the modified schemes, with very similar responses.  $\square$

(R.3.30) From our simulation studies and analytical results we conjecture that the model reference control algorithm as described by (3.2.9) driven by constant input  $r(t)$  has a bounded response for any initial condition and any parameter settings  $p_1 > 0$ ,  $\epsilon > 0$ ,  $a > 0$  and  $r$ .  $\square$

Figure 3.5 Output Tracking Error Caused by Nonsinusoidal Input

Figure 3.5.1

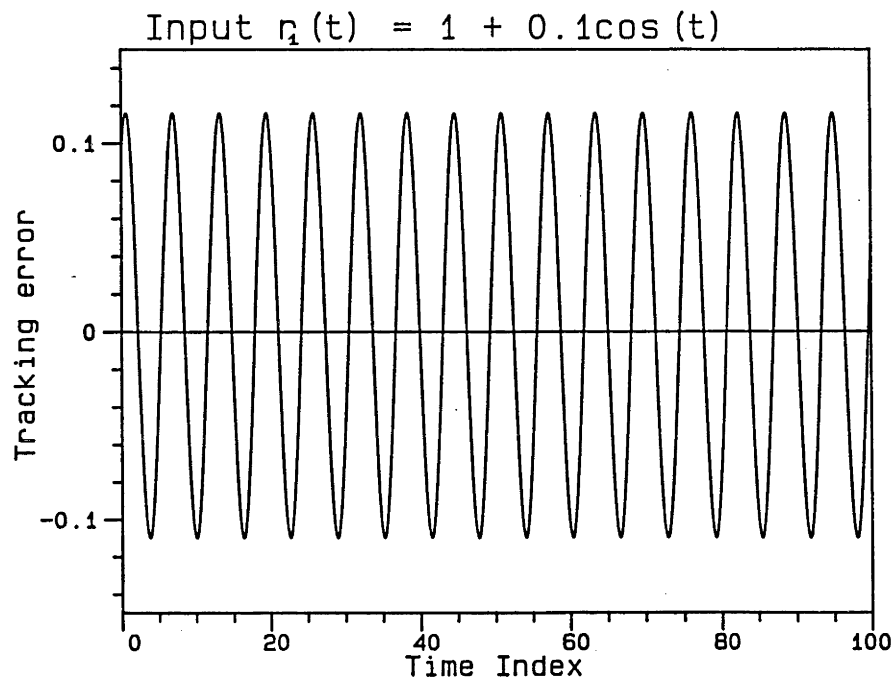


Figure 3.5.2

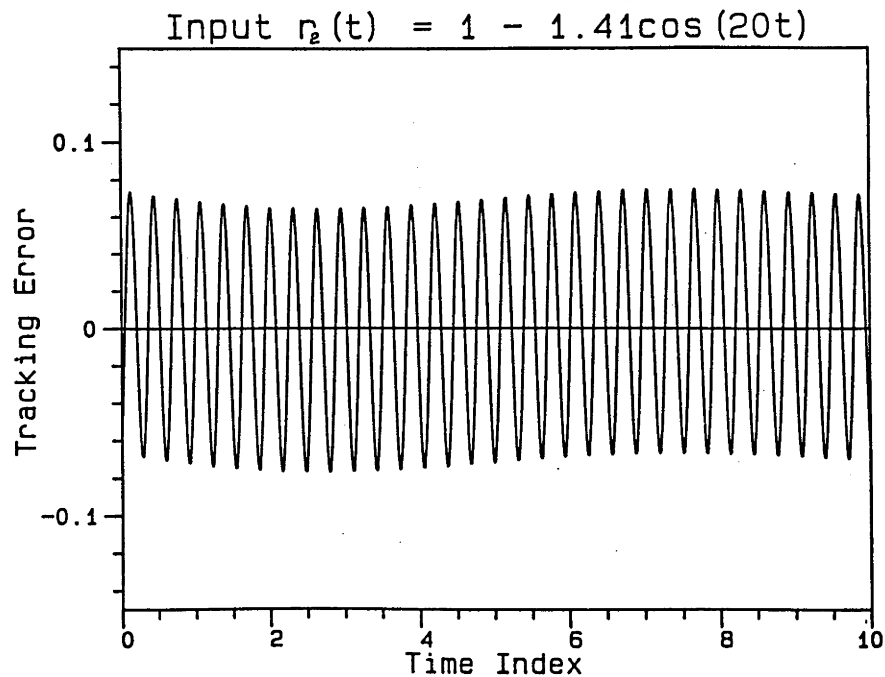




Figure 3.6 Output Tracking Error Caused by Large Adaptive Gain:

Figure 3.6.1

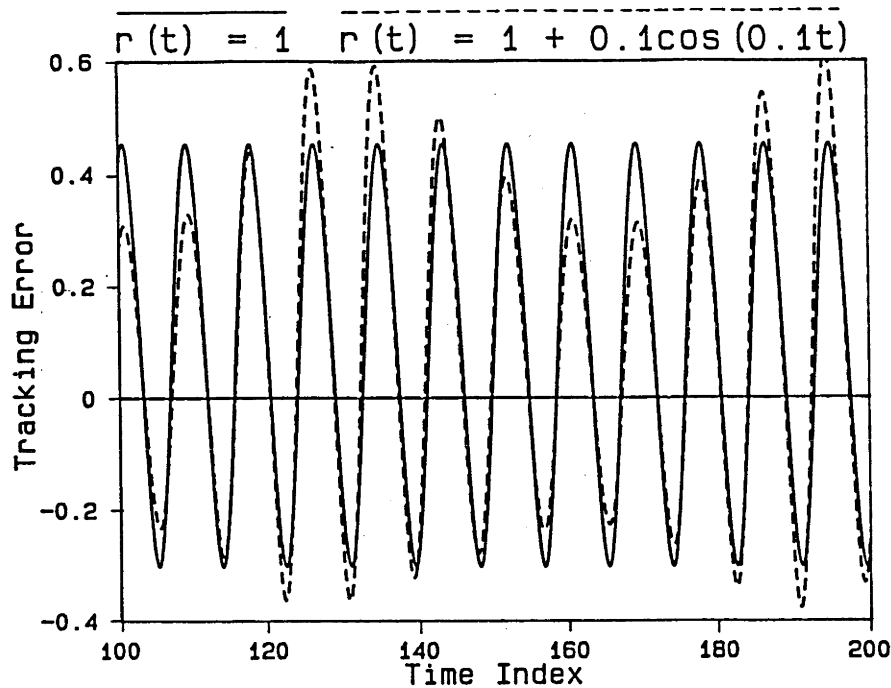


Figure 3.6.2 State Space Representation

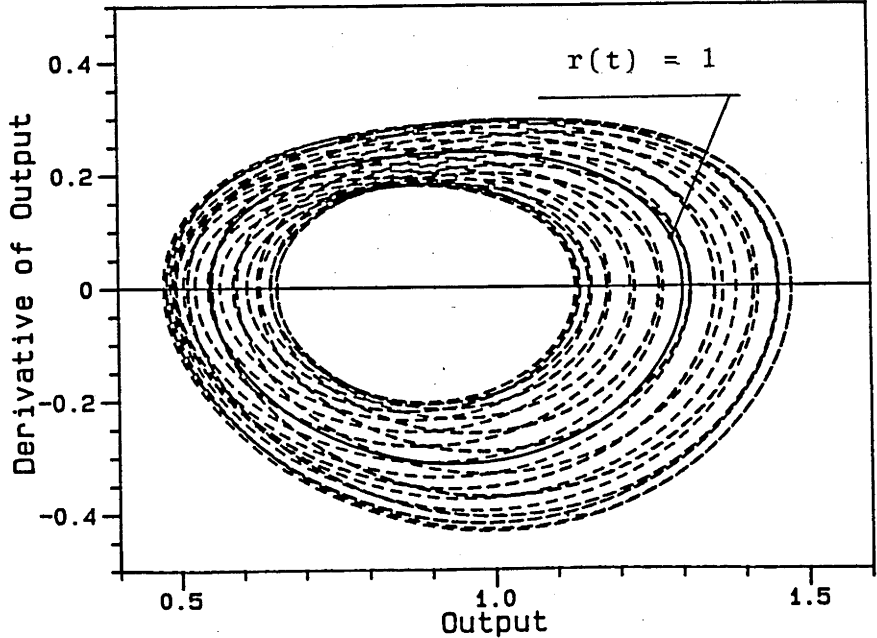


Figure 3.7 Tracking Error Caused by Large Adaptive Gain:

Figure 3.7.1

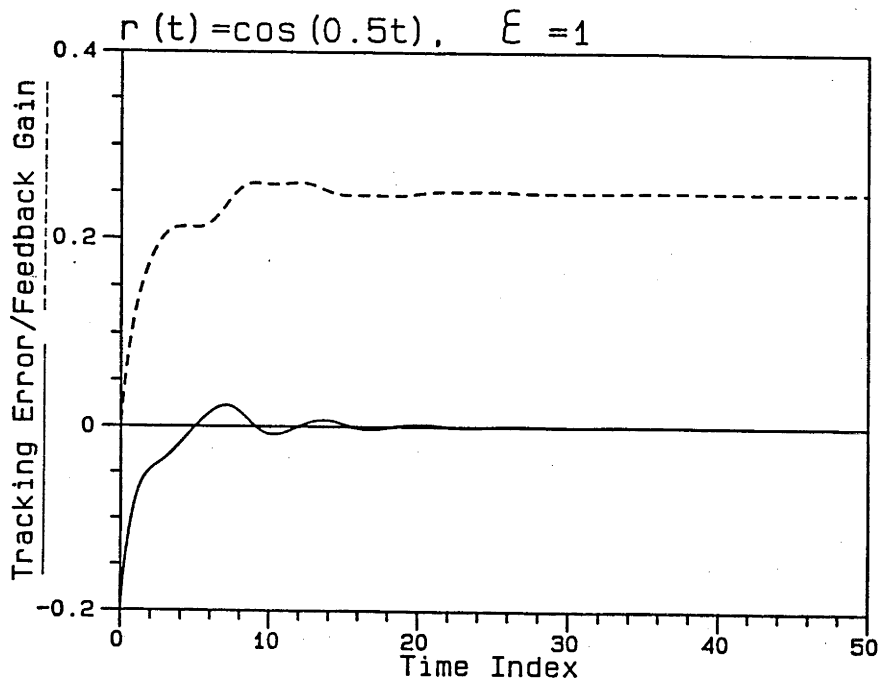
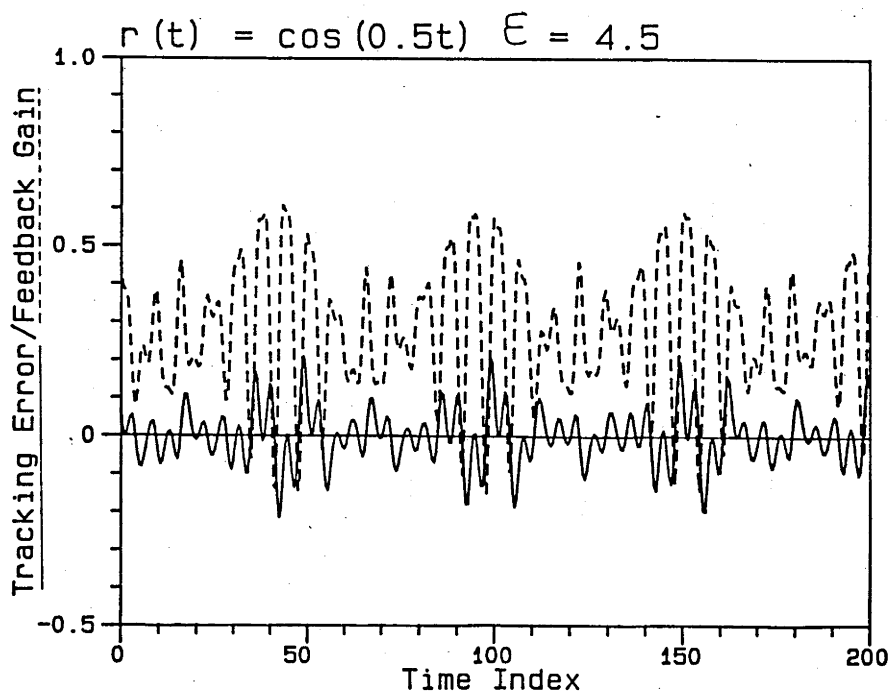


Figure 3.7.2



### 3.7.3 Exploring Further the Parameter Space

We analyse in greater detail the dynamics of the adaptive control problem set up in section 3.2 for constant inputs  $r(t) \equiv r$ . We focus our attention on the asymptotic dynamics.

Figure 3.8 displays in part the bifurcation diagram of the adaptive system (3.2.9) for the bifurcation parameter  $\epsilon$ , the adaptation gain (horizontal axis). The vertical axis represents a measure of the amplitude of the limit cycle or periodic orbit:

$$r = \sum_{i=1}^3 \max_{0 \leq t_1, t_2 \leq T(\epsilon)} |x_i(t_1) - x_i(t_2)|$$

where  $T(\epsilon)$  is the period of the limit cycle. The other parameters describing the adaptive system are set as: reference input  $r \equiv 1$ ; plant transfer function parameter  $p_1 = 1$ ; model pole  $a = 1$ ; and the normalization constant  $\mu = 1$ .

#### Remarks:

(R.3.31) The other plant parameter in the plant's transfer function,  $p_2$  which determines the d.c. gain and the stability of the plant, is irrelevant. Figure 3.8 is the bifurcation diagram for all  $p_2$ . The parameter  $p_2$  only affects the actual location in state space  $(x_1, x_2, x_3)$  of the limit cycle or asymptotic invariant set,  $(x_3^* = a - p_2 !)$  (see also remark (R.3.5)).  $\square$

The bifurcation diagram is obtained by locating the periodic orbits as fixed points of the Poincaré map (cf. Section 3.3) using a Newton Raphson procedure (cf. [2, Appendix E]). The local stability of the periodic orbit is then determined by integrating the first variational equations along the periodic orbit. (Stable periodic orbits are indicated by full lines, unstable periodic orbits are indicated by dashed lines in Figure 3.8.) Using Floquet Theory we obtain then the eigenvalues of the Jacobian of the Poincaré map evaluated at the fixed point corresponding to the periodic orbit. In this way new bifurcations can be detected, if they exist. The method can locate both stable and unstable periodic orbits, it suffices to have a good estimate of the period, and a reasonable estimate of its location. The Hopf Bifurcation Theorem 3.2 provides us with such estimates. Once we determine one periodic orbit, we can lock in

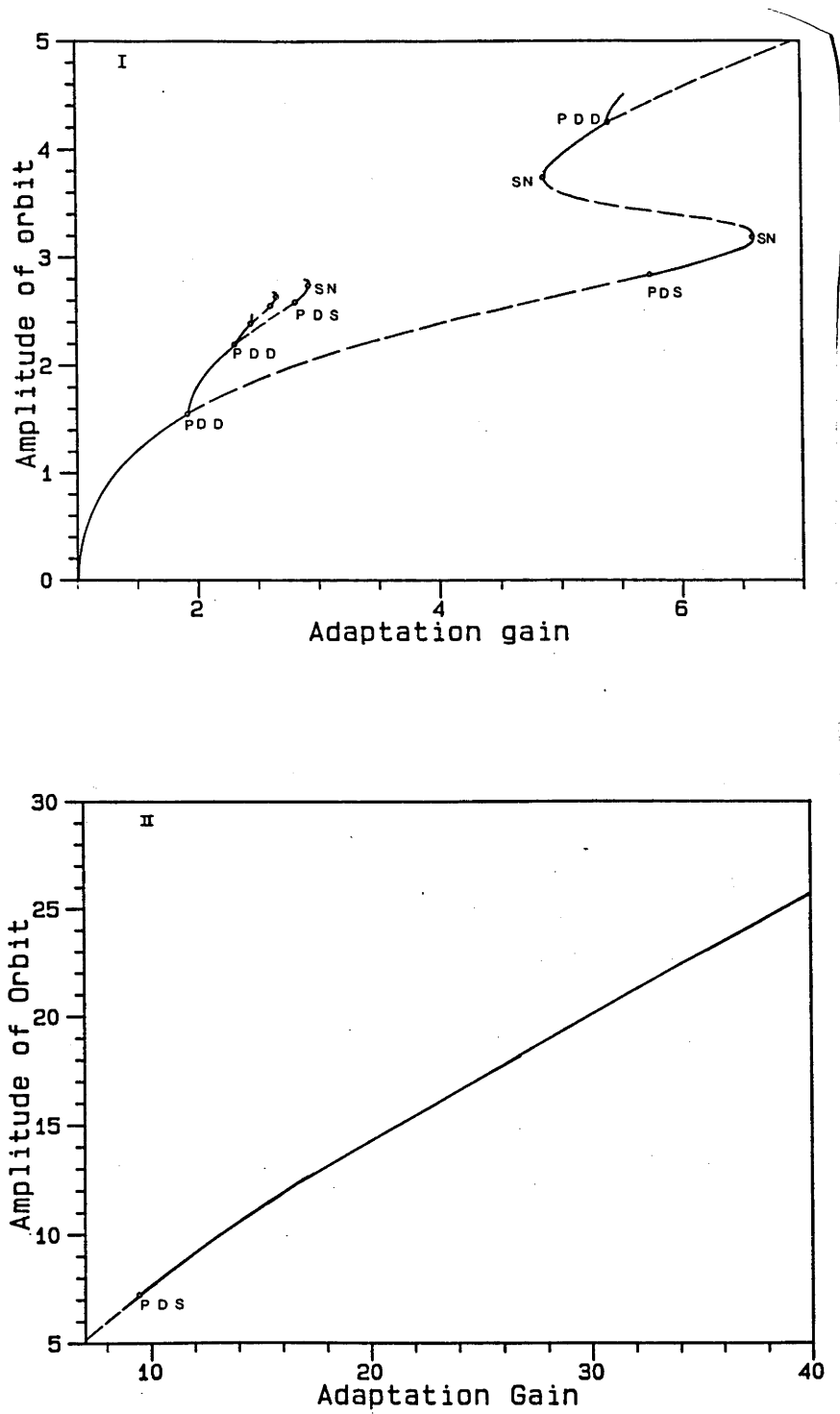
on it, and follow it through state space over the whole bifurcation parameter value range. In this way we can generate Figure 3.8, starting from the Hopf Bifurcation ( $\epsilon = 1$ ), and increasing  $\epsilon$  whilst tracking this orbit. If a new bifurcation phenomenon occurs, we identify its nature and if it generates a new periodic orbit, we follow that orbit as well. As return plane (cf. Figure 3.2,  $\Sigma$  is the return plane) we found  $\bar{x}_1 = -0.05$  adequate for the parameter range  $\epsilon \in (1, 10)$  which we studied most carefully, and  $\bar{x}_3 = 4$  for the parameter range  $\epsilon > 10$ .

Before explaining what Figure 3.8 implies we describe concisely the two bifurcations which were encountered besides the Hopf Bifurcation, in our numerical experiment. The Saddle-Node (SN) bifurcation is a phenomenon in which a stable and an unstable orbit annihilate themselves. (The stable orbit corresponds to a node in the Poincaré map, whilst, the unstable orbit (because the flow in our example is volume contracting) is a saddle in the Poincaré map, hence the name.) If the bifurcation value is  $\epsilon^*$ , then on one side of  $\epsilon^*$ , say  $\epsilon < \epsilon^*$ , the orbits both exist, and for  $\epsilon > \epsilon^*$  they do not exist. As  $\epsilon$  tends to  $\epsilon^*$  (from below) the periods and the locations of the orbits tend to a common limit. This bifurcation can in our example be recognised from the eigenvalues of the Jacobian of the Poincaré map, at criticality  $\epsilon = \epsilon^*$  these are  $+1$  and  $e^{-P_1 T}$ , where  $T$  is the (common) period. The Period-Doubling (PD) bifurcation is a phenomenon in which a stable (unstable) periodic orbit becomes unstable (stable) generating at the same time a stable (unstable) periodic orbit of about twice the period. At criticality both orbits coincide. This bifurcation (also called flip) can be easily recognised from the linearized Poincaré map, as one of the eigenvalues passes through  $-1$  at criticality. In our example, at criticality, the eigenvalues are  $-1$  and  $-e^{-P_1 T}$  whenever a period-doubling bifurcation occurs.

We now describe the Figure 3.8 Bifurcation Diagram (I). For  $\epsilon < 1$  the origin is locally uniformly asymptotically stable (Lemma 3.2); (it appears to be globally uniformly asymptotically stable). At  $\epsilon_1 = 1$ , a Hopf Bifurcation takes place, giving rise to a limit cycle of approximately period  $T_1 \approx 2\pi$ . This limit cycle persists, decreasing in period, increasing in amplitude until  $\epsilon$  reaches the value  $\epsilon_2 \approx 1.956$ , where a period doubling bifurcation occurs. The orbit becomes unstable, and a new stable periodic orbit of period  $T_2 \approx 11.96$  emerges.

Figure 3.8 Bifurcation Diagram

SN Saddle-Node bifurcation, PDD Period-Doubling Destablizing and PDS Period-Doubling Stabilizing bifurcation



These period doubling bifurcations follow each other very quickly, accumulating toward  $\epsilon_3 \approx 2.55$ , giving rise to periodic orbits, of increasingly larger amplitudes and increasing periods ( $\times 2!$ ). Tracking further the periodic orbit generated by the Hopf Bifurcation, we notice that it becomes again locally uniformly asymptotically stable at  $\epsilon_4 \approx 6.260$ , by a stabilizing period doubling bifurcation. An unstable periodic orbit of double period  $2T_3$  collapses at  $\epsilon_4$  with the unstable periodic orbit of period  $T_3 \approx 5.770$  stabilizing this orbit. Finally at  $\epsilon_5 \approx 6.611$ , this limit cycle undergoes a saddle node bifurcation followed by a second saddle node at  $\epsilon_6 \approx 4.878$ , a further destabilizing period doubling bifurcation at  $\epsilon_7 \approx 5.420$  and a final stabilizing period doubling bifurcation at  $\epsilon_8 \approx 9.87$ .

Remarks:

(R.3.32) We found that this sequence of a first period doubling, destabilizing (PDD) bifurcation, followed by a second stabilizing period doubling (PDS) bifurcation and a saddle node (SN) bifurcation is typical for all periodic orbits in this interval  $\epsilon \in (1, 10)$ . The following table illustrates this (only the first occurrence of the different bifurcations is indicated):

Table 3.1: Some bifurcation values of  $\epsilon$

Period	PDD	PDS	SN
$T(\approx 6)$	1.956	6.260	6.611
$2T$	2.360	2.833	2.934
$4T$	2.507	2.604	2.930
$8T$	2.538	-	-

(R.3.33) Observing the ratio  $(\epsilon_{n-1} - \epsilon_n) / (\epsilon_n - \epsilon_{n+1})$ , where  $\epsilon_n$  is the value of the  $n^{\text{th}}$  period doubling destabilizing bifurcation we observed that its value tend to approximately 4.7 from below. It has been shown that the limit of this ratio as  $n \rightarrow \infty$  is a universal constant of an (infinite) sequence of period doubling bifurcations [2, 9], through numerical experiments and analysis one estimates this

limit by 4.6692016... (the so called Feigenbaum's conjecture [9]).  $\square$

(R.3.34) This covers only partly what happens in this interval of the bifurcation parameter  $\epsilon \in (1, 10)$ . For example, the saddle-node bifurcation at  $\epsilon_6$  implies the existence of an unstable limit cycle which merges at  $\epsilon_6$  with the periodic orbit generated by the Hopf Bifurcation. This orbit could be followed further as a function of  $\epsilon$  and more saddle-node and period doubling bifurcations will be encountered ( $\epsilon_7, \epsilon_8$ ). As this applies for all orbits generated by period doubling bifurcations (in this interval of the bifurcation parameter), we realise that our numerical experiment can go on indefinitely... without yielding new information, and therefore we do not pursue this.  $\square$

It follows from these observations that the asymptotic dynamics of the adaptive control problem (especially for  $\epsilon \in (1, 10)$ ) are extremely complex, as they are characterized by competing sequences of period doubling and saddle-node bifurcations. This is witnessed by the asymptotic behaviour of a generic trajectory for the adaptive system with  $\epsilon = 4.5$ , displayed in Figure 3.9. It appears as if the trajectory never settles down to a periodic orbit, and fills in a two dimensional set, a strange attractor?

Increasing  $\epsilon$  beyond the value  $\epsilon_8$  one stable periodic orbit emerges, which increases in amplitude with  $\epsilon$  (Figure 3.8 Bifurcation Diagram (II)). In Figure 3.10, this orbit is displayed (projection in  $(x_1, x_2)$ - plane and  $(x_1, x_3)$ - plane) for  $\epsilon = 25, 50$ . Three phases can be observed in this orbit. In phase (A),  $\bar{x}_3$  is large and positive, causing an oscillatory decay of  $\bar{x}_1$  and  $\bar{x}_2$  towards  $\bar{x}_1 \cong -(r/a)\bar{x}_3/(\bar{x}_3 + a)$  and  $\bar{x}_2 \cong 0$ . Then,  $\bar{x}_3$  decreases, but very slowly as  $\bar{x}_1 \cong -(r/a)\bar{x}_3/(\bar{x}_3 + a)$  implies that  $d/dt(\bar{x}_3) < 0$  but small (phase B). Because  $\epsilon$  is large  $\bar{x}_3$  decreases, overshooting zero in a negative direction and therefore destabilizes the plant,  $\bar{x}_3$  and  $\bar{x}_2$  increase quickly, therefore bringing  $\bar{x}_3$  back to a large positive value (phase C). The larger  $\epsilon$  the larger  $\bar{x}_3$  becomes, both in negative and positive values, this can be recognised in both projections. In the  $(x_1, x_3)$  projection (output, feedback gain) this is obvious, in the  $(x_1, x_2)$  plane (output, derivative of output) this is reflected by the larger values taken by  $x_1$  and  $x_2$  and by the increasing number of cycles in the spiral. Notice in particular that this asymptotic behaviour for large adaptation gain is qualitatively identical to the transient response for large initial conditions discussed in remark R.3.25! (This remarkable feature reappears in the next case study, where we pay more attention to it.)

Figure 3.9 Complicated Asymptotic Dynamics (  $\epsilon = 4.5$  )

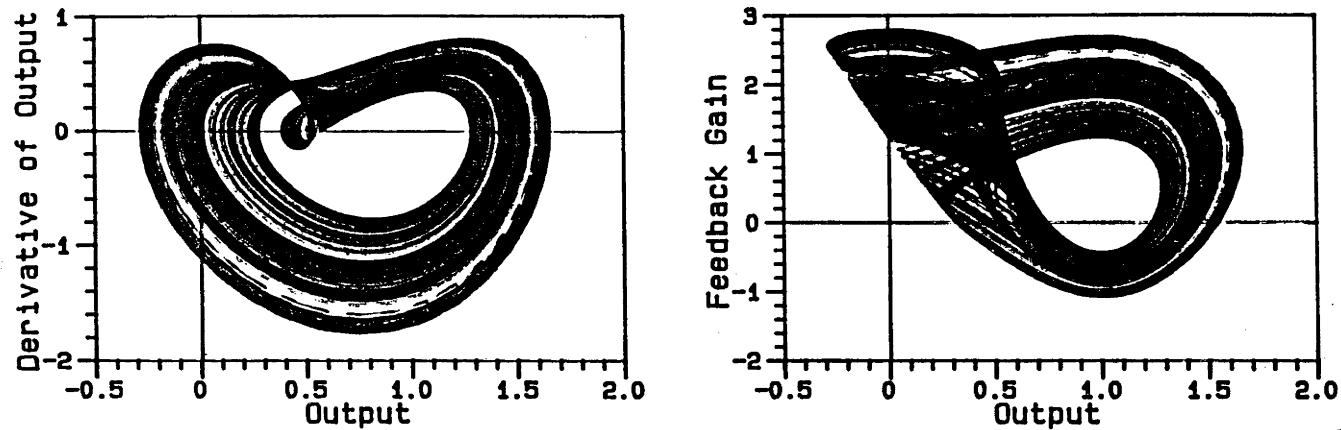
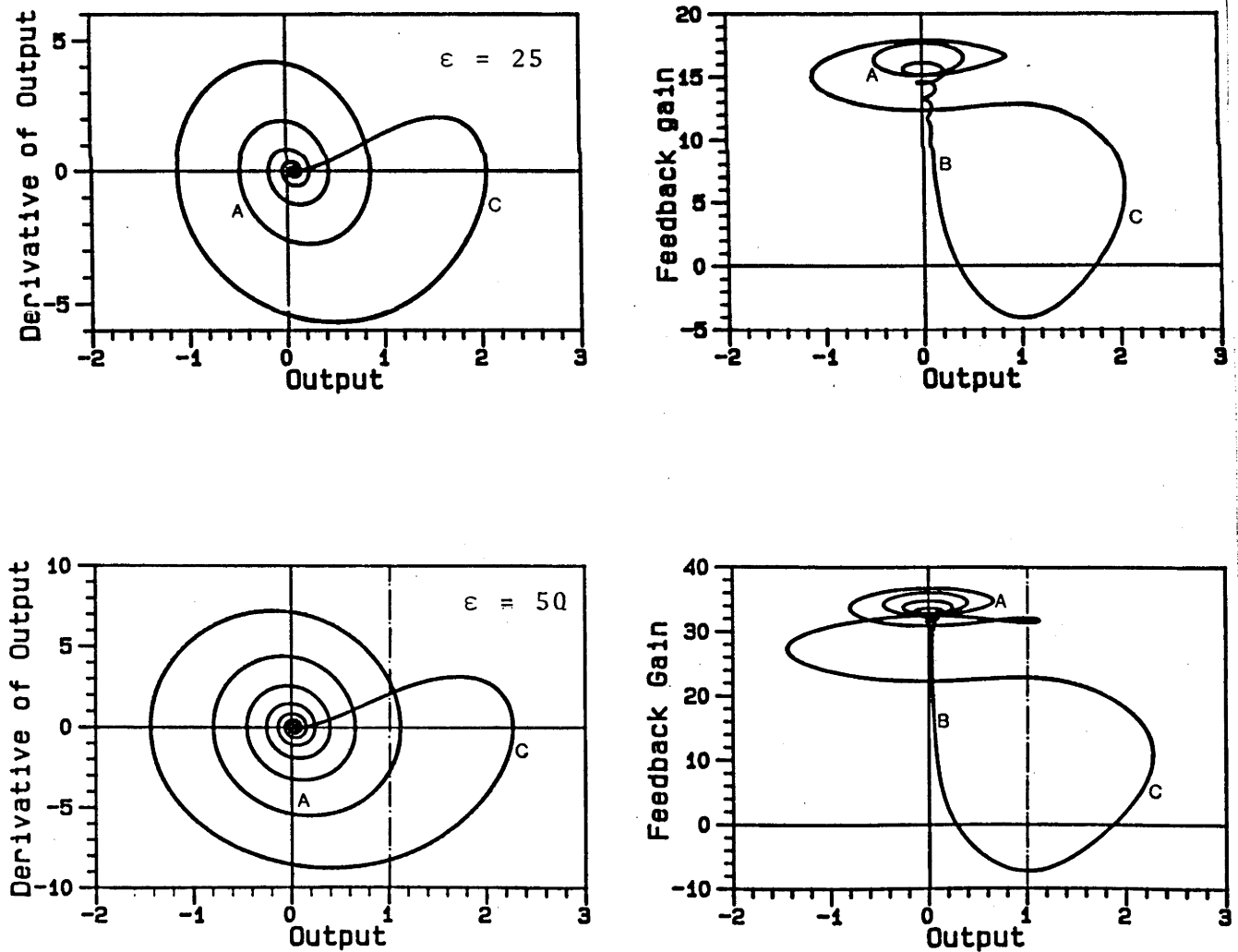


Figure 3.10 Asymptotic Dynamics for Large Adaptation Gain (  $\epsilon = 25, 50$  )





These observations indicate, (do not prove however) the existence of a homoclinic orbit(s) in the flow of the adaptive system for some values  $\epsilon_h$  of the bifurcation parameter  $\epsilon$ . Because the only fixed point is a saddle with eigenvalues  $\alpha \pm i\beta$  and  $\lambda$ , ( $0 < \alpha < -\lambda$  for all  $\epsilon > 1$ ) these homoclinic orbits would explain the presence of the complicated dynamics (period-doubling sequences ...) observed in the above (Silnikov's Theorem cf. [9, pp.318-325]).

Another pointer in this direction is that the global dynamics are hyperbolic which is consistent with the strange invariant sets (see e.g. Figure 3.9) implied by the existence of this type of homoclinic orbit. Also supporting this point is the bifurcation which takes place at  $p_1 = 0$ . This bifurcation is not well understood, but it is known that in its unfolding Hopf Bifurcations and homoclinic orbits are present. [9, pp. 364-376].

We have repeated the same analysis with  $p_1$  for the bifurcation parameter, ranging from  $p_1 = \epsilon r^2/a^3$  to 0, with qualitatively identical results.

Collecting these facts together we conjecture that the bifurcation diagram, with the adaptation gain as bifurcation parameter has the following generic form:

$$(1) \quad 0 < \epsilon < p_1 \quad a^3/r^2$$

The fixed point (origin) is uniformly asymptotically stable (in the large).

$$(2) \quad \epsilon_H = p_1 \frac{a^3}{r^2}$$

A Hopf Bifurcation takes place; an asymptotically stable limit cycle of period around  $2\pi/\sqrt{ap_1}$  emerges.

$$(3) \quad \epsilon_{h1} > \epsilon > p_1 \quad a^3/r^2$$

A first sequence of (destabilizing) period doubling bifurcations.

$$(4) \quad \epsilon = \epsilon_{h1}(p_1, a, r)$$

A homoclinic orbit appears in the flow, in the neighbourhood of the bifurcation value, the flow contains strange attracting sets. (homoclinic explosion, Silnikov Theorem)

...

...

...

$$(5) \quad \epsilon_j > \epsilon > \epsilon_{h1}$$

A last sequence of (stabilizing) period doubling bifurcations,

connected to a sequence of saddle-node bifurcations eliminates all periodic orbits but one.

$$(6) \quad \epsilon > \epsilon_j(p_1, a, r)$$

Only one periodic orbit remains, which is asymptotically stable and grows in amplitude as  $\epsilon$  increases.

Remarks:

(R.3.35) It is possible (and very likely) that there are many, different homoclinic explosions and sequences of period doubling bifurcations in the interval  $(\epsilon_H, \epsilon_j)$ .  $\square$

(R.3.36) A qualitatively identical bifurcation diagram is obtained for the unnormalized model reference control algorithm,  $\mu = 0$  or the algorithm with the more classical normalization (cf. (R.3.3)).  $\square$

It transpires from the above that local and global analyses complement each other. Specifically a global stability result is of little or no value without a guarantee for good local performance. Both local and global issues are nontrivial problems, and the boundary between the two becomes rather fuzzy in the presence of chaotic dynamics. In order to obtain an adaptive controller which performs well it is an absolute necessity to deal with the local issues.

### 3.8 Discussion

We have presented a fairly complete analysis of the asymptotic dynamics of the adaptive control problem set up in section 3.2 under the condition that the input is purely sinusoidal. Only for purely sinusoidal signals can the control objective be achieved, i.e. the plant output can indeed match the model output. Through a perturbation analysis we then demonstrated that this situation yields relevant information even in the case where the input was not purely sinusoidal and/or where the undermodelling was more severe than the one degree mismatch between plant and model as it was set up in section 3.2.

Very concisely, we can extract the following general observations:

- (1) The local dynamics - in the neighbourhood of the desired response - can be analysed, with good control performance in mind, over the complete parameter space. Whenever the local dynamics indicate instability, the nonlinearity of the adaptive algorithm stabilizes (that is if the adaptive algorithm is properly designed with a positive adaptation gain and stable model, sufficient excitation cf.(3.5.2)). The

typical mechanism underlying this soft loss of stability is a Hopf-type Bifurcation.

- (2) The performance of the adaptive system deteriorates quickly, once the parameters of the system indicate unstable local dynamics in the neighbourhood of the desired response. Although the global response is stable, the tracking error increases the more the local dynamics are unstable. The asymptotic dynamics can become extremely complex, characterized by chaotic dynamics and deviate substantially from the desired (zero output tracking error) response.
- (3) The available "global" results are either restricted to a small subset of the parameter space, only discussing singular perturbation type undermodelling error, and/or restricted in the set of initial conditions that can be dealt with. They only discuss stability (bounded input, bounded state response) and neglect largely the influence the design parameters have on the actual asymptotic dynamics. Global stability results are basically established using a Lyapunov function approach. This goes a long way in explaining the shortcomings of the global results we pointed out in the above, as it is very difficult to guess the right form of a Lyapunov function as a function of both the state variables and the parameters of the system. Any Lyapunov argument should be complemented by a characterization of the asymptotic dynamics.
- (4) The transient behaviour is not well understood, very few results are available. In view of the presence of chaotic dynamics this is not really surprising. Also, in view of the simulations (see e.g. Figure 3.4) stating that the transients are bounded is not very useful; unless this bound can be expressed as a function of the parameters and the initial conditions.

It transpires from the above that one has to be very cautious and pessimistic when interpreting global stability (bounded input, bounded state) results. This fact is strengthened by the observations made by F.M.A. Salam and Shi Bai in [16]. They discuss the existence of chaotic dynamics in the transient behaviour of a simple adaptive control problem (first order system, first order model, adaptive model reference control algorithm) generated by a bounded disturbance. Their results differ from the present ones in that here

chaos is generated by the undermodelling error and is therefore present regardless of the input disturbances. Also in [16] the model reference control algorithm with exponential forgetting was considered, whilst here we focussed our attention primarily on the unmodified model reference control algorithm. As a final difference the chaos established in [16] is of the horse shoe type, which is transient chaos, whilst here the chaos determines the asymptotic dynamics.

One big gap in our understanding of adaptive control dynamics remains: transient behaviour. This statement has to be qualified because the transient response of adaptive algorithms with small adaptation gain and initialised close to the desired response is well understood (local analysis); for a detailed analysis see [1]. From the above it is all too clear that this is a hard problem if other than local results are aimed for. The realistic problem of specifying guidelines for the design of an adaptive controller such that its response will remain within pre-specified limits from the desired behaviour and will converge to this behaviour within a certain error margin within a pre-specified settling time is wide open. (In the next chapter we deal with the transient behaviour of an adaptive pole placement law indicating that also the transient can be governed by chaotic dynamics.)

### 3.9 Historical Overview

The model reference control algorithm in its discussed form was a redesign by P.C. Parks [17] of the M.I.T. rule approach to model reference control. This redesign was motivated by the lack of conditions for global stability of the M.I.T. rule and the presence of examples of unstable behaviour of the M.I.T. rule. Using a Lyapunov argument (cf. Section 3.3, Theorem 3.1) P.C. Parks redesigned the adaptive algorithm such that for any plant described by a strictly positive real transfer function the new algorithm was asymptotically stable. (Parks' original algorithm did not include any normalization.)

This early algorithm was amended to cope with the general model output tracking control problem by Monopoli et al. [18]. The first global stability results were available in the late seventies to early eighties, [19,20]. The lack of robustness of these results was succinctly demonstrated by Bo Egardt [21], and C. Rohrs et al. [22]. This spurred a major effort into the analysis of the robustness properties of the model reference control scheme, in order to restore

its damaged reputation. New algorithms were proposed, mainly ad hoc variations of the same theme (normalization [4], exponential forgetting [3], dead zone [23], projection algorithm [4], error scaled exponential forgetting [5], multiple adaptive laws [24]). The aim is to obtain global and robust stability results. A different approach can be found in [1], where the emphasis is on local properties, in search for guidelines to ensure good performance.

Our contribution consists of describing some of the basic mechanisms that govern the dynamics of adaptive model reference control algorithms. Our approach demonstrates the importance of the different design parameters and the way in which they fundamentally change the local and global dynamics ( $(\epsilon/r/a)^2 < p_1 a$  is a key expression). The core of our results indicates how the local theory can break down (soft loss of stability, Hopf Bifurcation), and what the global theory overlooks (limit cycles, chaos...). In particular it follows that changing the adaptive law is no substitute for engineering design. The modified schemes display the same characteristics for different parameter settings. A combination of both approaches is essential in gaining a complete understanding of adaptive control, but it appears that the importance of the local results should not be underestimated.

We are not the first to report the existence of chaotic dynamics in model reference adaptive control. F.M.A. Salam and Shi Bai [16] demonstrated the presence of chaos due to periodic disturbances, in a model reference control algorithm with exponential forgetting (Melnikov-type chaos), which is of a transient nature, and Rubio et al. reported chaos in the adaptive control of a nonlinear plant (containing hysteresis in the actuator) [25]. (In [25] a self tuning regulator (indirect adaptive control) was discussed.) We do believe however that we are the first to analyse an adaptive problem where chaotic dynamics are generated by the adaptive mechanism itself. In the presence of undermodelling, the classical adaptive model reference control scheme exhibits for certain parameter settings chaotic dynamics which determine the asymptotic behaviour of the complete adaptively controlled system!.

3.10 References

- [1] B.D.O. Anderson, R.R.Bitmead, C.R.Johnson Jr, P.V. Kokotovic, R.K. Kosut, I.M.Y. Mareels, L. Praly, B. Riedle, Stability of Adaptive Systems: Passivity and Averaging Analysis, M.I.T. Press, Boston, 1986.
- [2] C. Sparrow, The Lorenz Equations: Bifurcations, Chaos and Strange Attractors, Applied Mathematical Sciences, Vol.41, Springer Verlag, 1982.
- [3] P.A. Ioannou, P.V. Kokotovic, Adaptive Systems with Reduced Models, Lecture Notes in Control and Information Sciences, Vol.47, Springer Verlag, 1983.
- [4] G. Kreisselmeier, K.S. Narendra, 'Stable model reference adaptive control in the presence of bounded disturbances', IEEE Trans. Auto. Control, AC-27, pp1169-11754, December 1982.
- [5] K.S. Narendra, A.M. Annaswamy, 'A new adaptive law for robust adaptation without persistent excitation', IEEE Trans.Auto.Control, to appear.
- [6] A.P. Morgan, 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. Control and Optimization, Vol 15, 1977, pp166-176.
- [7] B.D.O. Anderson, 'Exponential stability of linear equations arising in adaptive control', IEEE Trans.Auto.Control, AC-22, pp82-88, February 1977.
- [8] J.K. Hale, Ordinary Differential Equations, Interscience Series on Pure and Applied Mathematics, Vol.21, Wiley and Sons, 1969.
- [9] J. Guckenheimer, P. Holmes, Nonlinear Oscillations, Dynamical Systems and Bifuractions of Vector Fields, Applied Mathematical Sciences, vol.42, Springer Verlag, 1983.
- [10] G. Iooss, D.D. Joseph, Elementary Stability and Bifurcation Theory, Springer Verlag, 1981.
- [11] A.I. Mees, Dynamics of Feedback Systems, Wiley and Sons, 1980.

- [12] A.B. Poore, 'On the theory and application of the Hopf-Friedrichs bifurcation theory', Archive for Rational Mechanics and Analysis, Vol.60, pp371-393.
- [13] J.E. Marsden, M. McCracken, 'The Hopf Bifurcation and Its Application', Applied Mathematical Sciences Vol. 19, Springer Verlag, 1976.
- [14] W.A. Coppel, 'Dichotomies in Stability Theory', Lecture Notes on Mathematics, Vol 629, Springer Verlag, 1978.
- [15] E.H. Abed, 'Singularly perturbed Hopf bifurcation', IEEE Trans. Circuits and Systems, Vol.CAS-32, No12, December 1985, pp1270-1280.
- [16] F.M.A. Salam, Shi Bai, 'Disturbance generated bifurcation in a simple adaptive control system: Analysis and Simulation', Memorandum No.MSUEESS SM85/01, October 1985.
- [17] P.C. Parks, 'Liapunov redesign of model reference adaptive control systems', IEEE Trans. Auto. Contr., Vol AC-11, No4, August 1966, pp362-367.
- [18] R.V. Monopoli, 'Model reference adaptive control with an augmented error signal', IEEE Trans. Auto. Contr., Vol AC-19, No5, October 74, pp474-485.
- [19] A. Feuer, A.S. Morse, 'Adaptive control of single input single output systems', IEEE Trans. Auto. Contr., Vol AC-23, No4, Aug. 1978, pp557-570.
- [20] K.S. Narendra, L.S. Valavani, 'Stable adaptive controller design - direct adaptive control', IEEE Trans. Auto. Contr., Vol AC-23, No4 Aug. 1978, pp570-583.
- [21] Bo Egardt, 'Stability of adaptive systems', Lecture Notes in Control and Information Sciences, Springer Verlag, 1979.
- [22] C. Rohrs, L. Valavani, M. Athans, G. Stein, 'Robustness of continuous time adaptive control algorithms in the presence of unmodelled dynamics', IEEE Trans. Auto. Contr., Vol AC-30, Sept. 1985, pp881-889.
- [23] B.B. Peterson, K.S. Narendra, 'Bounded error adaptive control', IEEE Trans. Auto. Contr., Vol AC-27, No6, Dec. 82, pp.1161-1169.

- [24] G.C. Goodwin, D.J. Hill, D.Q. Mayne, R.H. Middleton, 'Adaptive Robust Control (convergence, stability and performance)', Technical Report EE8544, Newcastle University.
  
- [25] F.R. Rubio, J. Aracil, E.F. Camacho, 'Chaotic motion in an adaptive control system', Int. J. Control, 1985, Vol42, N0.2, pp353-360.



#### 4. STABILIZING NONLINEAR DYNAMICS IN ADAPTIVE CONTROL

##### 4.1 Introduction

In the previous case studies, we have analysed the dynamics of adaptive control in ideal and nonideal situations highlighting how these are affected by the design variables. Both the local (e.g. in the neighbourhood of a desired response) and global behaviour of an adaptively controlled system changes in a highly nonlinear and nontrivial way with the design variables (adaptation speed, input characteristics...). Though the complete characterization of the (local) dynamics in even the simplest of the problems considered is beyond all available analytical tools, the local response can be effectively and efficiently analysed using linearization techniques combined with the time scale separation principle. As has been illustrated in the previous chapters these methods yield valuable and precise information giving clues as to how one can design an adaptive algorithm that will meet given design criteria. However, adaptively controlled systems are nonlinear in general, which implies that the behaviour in the large may be quite different from the local behaviour. In the previous chapter we exemplified this by demonstrating that the nonlinear dynamics could be globally stabilizing even when the local dynamics in the neighbourhood of the desired response were unstable. Using techniques from bifurcation theory and global analysis we have shown that the larger the adaptation gain becomes the more vividly the adaptive system response displays typical nonlinear phenomena, and the less relevant the local results become. In this last case study we want to expose more clearly these nonlinear dynamics and their potentially stabilizing properties, demonstrating that the complicated nonlinear dynamics of adaptive control indeed have a redeeming feature: amongst other things they make adaptive control work!

Our last adaptive control problem is set up in discrete time. As in the previous chapter, we assume that the plant to be controlled is a second order system, whilst the control law is designed on the basis of a first order model. The control objective is to regulate the plant output to zero. Using the certainty equivalence approach outlined in Chapter 1, Section 1.1, the adaptive controller is constructed on the basis of a very fast (deadbeat) adaptive parameter estimator coupled with a deadbeat feedback control law. The problem is constructed in such a way that the adaptively controlled system

cannot be linearized around its desired trajectory. Hence, the closed loop dynamics are entirely dictated by nonlinear effects, and therefore this example serves its primary purpose of exposing the nonlinear dynamical phenomena more fully. However, the implications of our analysis are more farreaching in that we demonstrate that the particular dynamics associated with this nonlinearizable set up describe the transient behaviour of very standard linearizable adaptive control schemes with large initial conditions.

The analysis is simplified as compared to the previous case study by considering regulation only ( $r \equiv 0$ ) and by using a combination of a deadbeat control with a deadbeat identifier ( $\equiv$  fixed adaptation gain). The only free parameter on which the dynamics depend is the parameter describing the undermodelling (cf.  $p_1$  in the previous chapter). Because of this simplification we are able to give a complete analysis of the nonlinear dynamics (including transient behaviour) over the whole parameter range. As a consequence we are able to describe precisely the robustness properties of the adaptive controller considered with respect to this type of undermodelling. In particular we characterize the set of second order systems which can be stabilized by this adaptive control algorithm.

The chapter is organised as follows. Section 3 is devoted to the explicit formulation of the adaptive control problem studied, where we state the class of plants considered and parametrized models used, the identifier structure and the linear certainty equivalence control strategy. We derive in this section an explicitly nonlinear difference equation which describes the complete closed loop adaptive system. The dynamics of this difference equation fall into three distinct categories depending upon a single parameter ( $b$ ) characteristic of the unmodelled dynamics and the next three sections concentrate on describing the closed loop behaviour for each of these classes. Section 4 briefly describes the properties of the adaptive scheme when no modelling errors are present, i.e.  $b=0$ . In Section 5 we consider negative values for the parameter  $b$  and prove that the feedback gain is asymptotically periodic. For a range of values of negative  $b$  this periodic gain stabilizes the plant. In Section 6 we consider positive  $b$  and show that the feedback gain is chaotic, again stabilizing the plant for a range of values of  $b$ . Section 7 deals more fully with the effects on the closed loop stability of the plant due to these different feedback gains and considers questions of performance and robustness of the adaptive control

scheme. In Sections 8 and 9 we conclude and draw together the threads of the previous sections to discuss the implications of the results of this case study for adaptive control.

#### 4.2 Conventions and Notations

We briefly introduce some of the terminology from dynamical system analysis for difference equations [1].

A homeomorphism is a continuous function with a continuous inverse.

Let  $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$  be a homeomorphism.

Consider the difference equation  $x_{k+1} = f(x_k)$ ,  $x_0$ .  $f$  is called the state transition map.

A trajectory (orbit) is :  $\{f^k(x_0), k \in \mathbb{Z}\}$  a sequence of iterations passing through  $x_0$  at  $k = 0$ . (Slightly abusing this formal definition, we use trajectory also to denote any sequence (finite or infinite) of consecutive points  $x_k$  in state space through which an orbit passes.)

A fixed point is a solution of  $x = f(x)$ .

A periodic orbit of period  $p$  is an orbit  $\{f^k(x), k \in \mathbb{Z}\}$  such that  $f^p(x) = x$ , and  $f^l(x) \neq x$  for  $l = 1, \dots, p-1$ ,  $p > 1$  (a finite orbit).

The stable manifold of a fixed point  $x$  (periodic orbit) is the collection of all points converging to  $x$  (periodic orbit) under the forward iteration of  $f$ .

The unstable manifold of a fixed point  $x$  (periodic orbit) is the collection of all points converging to  $x$  (periodic orbit) under the iteration of  $f^{-1}$ .

An invariant set is  $S \subset \mathbb{R}^n : f(S) \subset S$ ;  $f|_S$  denotes  $f$  restricted to  $S$ .

An invariant attractor is an invariant set  $A$  for which there exists a superset  $D$ , ( $D \supset A$ ) of positive Lebesgue measure such that 
$$\lim_{n \rightarrow \infty} f^n(D) \subset A$$

An indecomposable invariant set is an invariant set such that for any two points  $x, y$  in this set and for any  $\mu > 0$ , there exist an integer  $n$  and a sequence of points  $x = x_0, x_1, \dots, x_n = y$  in this set and time indices  $t_1, \dots, t_n \in \mathbb{N}$  such that  $|f^{t_i}(x_{i-1}) - x_i| < \mu \quad \forall i = 1, \dots, n$ . (Intuitively, any two points can be linked arbitrarily closely by a certain trajectory, completely in the set!)

A hyperbolic invariant set  $S$ , is an invariant set which has a continuous invariant direct sum decomposition on its tangent space such that for any  $x \in S$  the tangent space  $T_x$  is the direct sum of a stable eigenspace  $E_x^s$  and unstable eigenspace  $E_x^u$ . There exist constants  $C > 0$  and  $0 < \lambda < 1$  (independent of

$x \in S$  for uniform hyperbolicity) such that for any  $v \in E_x^s$   $|Df^{-n}(x)v| \leq C\lambda^n|v|$ ; and for any  $w \in E_x^u$   $|Df^n(x)w| \leq C\lambda^n|w|$ .  $E_x^s$  and  $E_x^u$  can be given bases which change in a continuous way with  $x$  (For more details see Definition 5.2.6. in [1]).

A Cantor set is a closed set, such that the largest connected subset is a point, and every point in the set is a limit point.

The study of the dynamics of the difference equation  $x_{k+1} = f(x_k)$ , is concerned with the topology of the space of trajectories:  $\{f^n(x), n \in \mathbb{Z}, x \in \mathbb{R}^n\}$ .

Two state transition maps  $f, g$  are topologically equivalent if there exists a homeomorphism  $h$  that takes orbits of  $f$  to orbits of  $g$ :  $h \circ f = g \circ h$ . As far as dynamical behaviour is concerned we do not distinguish between topologically equivalent state transition maps.

In the sequel we will be mainly concerned with state transition maps defined in  $\mathbb{R}^2$ . In order to be able to illustrate the complete state space we "compactify"  $\mathbb{R}^2$  using the homeomorphism  $H$ :

$$H: \mathbb{R}^2 \rightarrow (-1,1) \times (-1,1) \quad (x,y) \rightarrow \left( \frac{x}{1+|x|}, \frac{y}{1+|y|} \right) \quad (4.2.1)$$

In the future, although we state a result for  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ , the illustrations are for the topologically equivalent map  $g: g = H \circ f \circ H^{-1}$ .

The open first to fourth quadrants of  $\mathbb{R}^2$  are denoted by  $Q_1$  to  $Q_4$ . The closure of a subset  $S$  in  $\mathbb{R}^2$  is denoted as  $S^c$ , the complement of  $S$  in  $\mathbb{R}^2$  is  $S^c$ , the boundary is represented as  $\delta(S)$ , its interior by  $\text{int}(S)$ .

#### 4.3 Problem Description

In this section we sketch the adaptive control problem which we want to study. After setting up the problem, we then distinguish the topologically different types of dynamical behaviour exhibited by this equation.

Suppose that the system to be controlled is given by:

The Plant:

$$y_k = ay_{k-1} + by_{k-2} + u_{k-1}; \quad k \in \mathbb{N}; \quad a, b, u_k, y_k \in \mathbb{R} \quad (4.3.1)$$

Here  $a, b$  are the unknown, but fixed, parameters of the plant. The designer believes that the system can be adequately represented by a first order model:

The Designer's Model:

$$y_k = \hat{a}y_{k-1} + u_{k-1}; \quad k \in \mathbb{N}; \quad \hat{a} \in \mathbb{R} \quad (4.3.2)$$

In order to achieve regulation of the plant output to zero, we use a deadbeat identification scheme coupled with a deadbeat control law, based on the current parameter estimate.

The Deadbeat Parameter Estimator:

$$\hat{a}_{k+1} = \hat{a}_k + \frac{1}{y_k} (y_{k+1} - \hat{a}_k y_k - u_k); \quad y_k \neq 0; k \in \mathbb{N} \quad (4.3.3.1)$$

$$\hat{a}_{k+1} = \hat{a}_k; \quad y_k = 0; k \in \mathbb{N} \quad (4.3.3.2)$$

The Deadbeat Control Law:

$$u_k = -\hat{a}_k y_k; \quad k \in \mathbb{N} \quad (4.3.4)$$

Combining equations (4.3.1)-(4.3.4) we obtain the closed loop description:

$$y_k = \tilde{a}_{k-1} y_{k-1} + b y_{k-2} \quad (4.3.5)$$

$$\tilde{a}_k = \begin{cases} a - \hat{a}_k \\ -b \frac{y_{k-2}}{y_{k-1}} \\ \tilde{a}_{k-1} \end{cases} \quad \begin{matrix} y_{k-1} \neq 0 \\ y_{k-1} = 0 \end{matrix} \quad (4.3.6)$$

Eliminating the parameter error,  $\tilde{a}_k = a - \hat{a}_k$ , the closed loop can be described as a function of the plant output  $y_k$  only:

$$y_k = -b \frac{y_{k-3}}{y_{k-2}} y_{k-1} + b y_{k-2} \quad (4.3.7)$$

Introducing the ratio  $r_k = y_k/y_{k-1}$ , of successive plant outputs, (4.3.5)-(4.3.6) can alternatively be represented as:

Alternative Description of Closed Loop:

$$r_k = b \left[ \frac{1}{r_{k-1}} - \frac{1}{r_{k-2}} \right]; \quad k \in \mathbb{N}; \quad b \in \mathbb{R} \quad (4.3.8)$$

Link equations

$$y_k = r_k y_{k-1}; \quad k \in \mathbb{N} \quad (4.3.9)$$

$$\hat{a}_k = a + b \frac{1}{r_{k-1}}; \quad k \in \mathbb{N}; \quad a, b \in \mathbb{R} \quad (4.3.10)$$

This alternative description only makes sense if the event "division by zero" does not occur. In this situation every trajectory of (4.3.5)-(4.3.6) defines a unique trajectory of (4.3.8) via equation (4.3.9); conversely every trajectory of (4.3.8) defines a one parameter family of trajectories of (4.3.5), (4.3.6) via equations (4.3.9) and (4.3.10). Indeed, given  $r_0, r_1$ , the trajectory  $\{r_k(r_0, r_1), k \in \mathbb{N}\}$  ( $r_0(r_0, r_1) = r_0, r_1(r_0, r_1) = r_1$ ) is uniquely defined, but  $\{y_k, k \in \mathbb{N}\}$  is only specified up to a scaling factor:

$$y_k = \prod_{l=0}^k r_l(r_0, r_1) \cdot y_{-1}; k \in \mathbb{N}.$$

This simply reflects that the original system (4.3.5)-(4.3.6) has a state vector in  $\mathbb{R}^3$ ; but because of the system's specific structure it is sufficient to study (4.3.8) (with a state in  $\mathbb{R}^2$ ) in order to capture the generic dynamics of the adaptive system. It is not too difficult to analyse the behaviour of the trajectories for which a "division by zero" could occur - however this does not add anything fundamental to our knowledge of the system, and moreover this only gives us information about a two dimensional manifold of initial conditions in the state space of the original system. Therefore we do not pursue this here [2].

Remarks:

(R.4.1) The parameter estimator (4.3.3) can be seen as the limit of a normalized least mean square algorithm or recursive least square algorithm with forgetting factor [3]:

$$\hat{a}_{k+1} = \hat{a}_k + \frac{y_k}{c + y_k^2} (y_{k+1} - \hat{a}_k y_k - u_k) \quad (4.3.11)$$

The parameter  $c$  ( $>0$ ) is the inverse of the stepsize of the least mean square algorithm. The deadbeat parameter estimate follows from (4.3.11) by letting the stepsize  $c^{-1}$  tend to infinity, or  $c$  to zero.  $\square$

(R.4.2) Notice that the design followed strictly the classical route of adaptive control algorithm design based on the certainty equivalence principle as outlined in Chapter 1, Section 1.2. Notice also that the implemented control algorithm is the much celebrated minimum variance controller.  $\square$

(R.4.3) The nonlinear dynamics of the complete closed loop system for this particular adaptive control problem are made explicit in terms of the output of the plant only in equation (4.3.7) or equivalently in (4.3.8)-(4.3.9). Notice in particular that it is impossible to linearize the adaptive response around

$\{y_k \equiv 0, k \in \mathbb{N}\}$  which is the desired trajectory.  $\square$

(R.4.4) The closed loop dynamics (4.3.7)-(4.3.8) are independent of the plant parameter  $a$  - the sum of the values of the open loop poles of the plant's transfer function. This is characteristic of the adaptive nature of the control algorithm. An adaptive algorithm eliminates the influence of part of the plant characteristics on the closed loop dynamics (cf. Remark (3.28)). In the previous chapter it was  $1/p_2$  the d.c. gain of the plant which was irrelevant for the closed loop dynamics. In terms of robustness this feature looks most promising, i.e. if the adaptive algorithm works, it works for both stable and unstable plants depending only on the parameter  $b$ . Although the closed loop dynamics are independent of the parameter  $a$ , the signals within the loop, the parameter estimate ( $\hat{a}_k$ ) and the control signal ( $u_{k+1} = -\hat{a}_k y_k$ ), do depend on it. Both feedback gain/parameter estimate and control action adapt to different values for different plants, i.e. the control scheme is really an adaptive control scheme and not merely a robust nonlinear controller.  $\square$

(R.4.5) Observe that provided the plant parameters are known, the plant can be stabilized by constant output feedback (4.3.4) iff  $|b| < 1$ .  $\square$

In the following sections we study the equation (4.3.8) in detail. By way of preliminary analysis we note that as a function of the parameter  $b$  we can at most distinguish three topologically different types of dynamical behaviour for equation (4.3.8). Indeed, rescaling as

$$v_k = r_k / \sqrt{|b|}; \quad b \neq 0 \quad (4.3.12)$$

leads to

$$v_k = \text{sign}(b) \left[ \frac{1}{v_{k-1}} - \frac{1}{v_{k-2}} \right]; \quad k \in \mathbb{N} \quad (4.3.13)$$

So it is possible to consider three situations  $b < 0$ ,  $b = 0$ ,  $b > 0$ . That the topology for these three cases is indeed different follows from the observation that for  $b < 0$  equation (4.3.14) has two periodic orbits of period two  $\{\sqrt{2}, -\sqrt{2}, \dots\}$  or  $\{-\sqrt{2}, \sqrt{2}, \dots\}$ , whilst for  $b = 0$ , the orbits are of the form  $\{v_0, 0, 0, \dots\}$  and for  $b > 0$ , there does not exist a two periodic solution, as can be easily verified.

In the sequel we discuss all three types of dynamical behaviour.

#### 4.4 The Closed Loop Dynamics I: $b = 0$

In this ideal situation, the dynamics are quite obvious, as one could expect from the combination of a deadbeat identifier and deadbeat control law.

The closed loop reduces to:

$$\text{Plant:} \quad y_k = ay_{k-1} + u_{k-1}; \quad y_{-1} \quad \forall k \in \mathbb{N}$$

$$\text{Identifier:} \quad \hat{a}_k = a; \quad (y_{-1} \neq 0); \quad \forall k \geq 0$$

$$\text{Control law:} \quad u_{k-1} = -ay_{k-1}; \quad \forall k \geq 1$$

$$\text{Hence:} \quad y_k \equiv 0; \quad \forall k \geq 1$$

In the situation that  $y_{-1} = 0$ , the identifier does not identify anything useful, but there is also no need to identify anything, as in this case  $y_k \equiv 0 \quad \forall k \in \mathbb{N}$  and  $\hat{a}_k \equiv a_{-1} \quad \forall k \in \mathbb{N}$ ; hence the control objective is met.

In this situation the equation (4.3.8) governing the ratio of successive outputs is redundant, except to highlight a definitional problem with the deadbeat identifier when the deadbeat controller actually works. (Notice however, that it still captures the generic dynamics.)

#### Remarks:

(R.4.6) The adaptive scheme has excellent robustness properties with respect to multiplicative noise and lesser robustness with respect to additive noise. We comment on these aspects in greater detail in Section 4.8.  $\square$

(R.4.7) We illustrate the tracking capabilities of this algorithm with the following example. Assume that the plant is a time varying, first order system which can be described as:

$$\text{Plant:} \quad y_k = a_k y_{k-1} + u_{k-1} \quad a_k \in \mathbb{R}, \quad k \in \mathbb{N}$$

The closed loop responds as:

$$\text{Identifier:} \quad \hat{a}_k = a_k \quad (y_{-1} \neq 0) \quad \forall k \geq 0$$

$$\text{Control Law:} \quad u_k = -a_k y_k \quad \forall k \geq 0$$

$$\text{Closed Loop:} \quad y_k = (a_k - a_{k-1}) y_{k-1} \quad \forall k \geq 0$$

Hence, exponentially fast regulation is obtained if  $|a_k - a_{k-1}| < 1 \quad \forall k! \quad \square$



#### 4.5 The Closed Loop Dynamics II, $b < 0$

In this section we study the equation

$$v_k = -\frac{1}{v_{k-1}} + \frac{1}{v_{k-2}}; \quad k \in \mathbb{N} \quad (4.5.1)$$

which describes the evolution of the normalized ratio of successive plant outputs in the case  $b < 0$  (4.3.12)-(4.3.13). We demonstrate the existence of a "globally" attractive periodic orbit of period two, and investigate the consequence of this periodic behaviour for the closed adaptive loop via equations (4.3.9), (4.3.10) and (4.3.12).

Firstly, we introduce a simple time dependent scaling transformation, which maps the periodic orbits into two fixed points:

$$w_k = (-1)^k \frac{v_k}{\sqrt{2}}; \quad k \in \mathbb{N} \quad (4.5.2)$$

The equation governing  $w_k$  becomes:

$$w_k = \frac{1}{2} \left[ \frac{1}{w_{k-1}} + \frac{1}{w_{k-2}} \right]; \quad k \in \mathbb{N} \quad (4.5.3)$$

Introduce the following state space representation; define the state as:

$$x_k = \begin{bmatrix} w_k \\ w_{k-1} \end{bmatrix}; \quad k \in \mathbb{N} \quad (4.5.4)$$

define the state transition map  $F$  as:

$$F \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \left[ \frac{1}{y_1} + \frac{1}{y_2} \right] \\ y_1 \end{bmatrix}; \quad (4.5.5)$$

the difference equation (4.5.3) can then be represented as

$$x_{k+1} = F(x_k); \quad x_0; \quad k \in \mathbb{N} \quad (4.5.6)$$

$x_0$  is the initial condition. This recursion is properly defined on the domain:

$$D_F = \mathbb{R}^2 \setminus \bigcup_{n \geq 0} F^{-n} \left\{ \begin{bmatrix} 0 \\ y \end{bmatrix}, \begin{bmatrix} y \\ 0 \end{bmatrix}, \quad y \in \mathbb{R} \right\} \quad (4.5.7)$$

In precise terms we are interested in the dynamics of the continuously differentiable map  $F$  restricted to the domain  $D_F$  ( $F|_{D_F}$ ).

The inverse map is given by:

$$F^{-1} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} y_2 \\ \frac{-y_2}{1-2y_1y_2} \end{bmatrix} \quad (4.5.8)$$

In order to obtain a precise characterization of the dynamics of the state transition map  $F$  and therefore, via the link equations (4.3.9)-(4.3.10), of the adaptive control problem in the situation that  $b < 0$ , it is necessary to describe  $D_F$  in detail. This requires us to study the inverse map  $F^{-1}$  more closely. Concisely we demonstrate that  $F^{-1}$  leaves the union of the second and fourth closed quadrants invariant. The origin is an attractor in this set; under the action of  $F^{-1}$  all trajectories converge to the origin. These properties allow us to describe  $D_F$  and to establish that under the action of  $F$  all trajectories, starting in  $D_F$  remain in the open second and fourth quadrant only for a finite number of iterations, after which they remain in the open first and third quadrant. Using a Lyapunov argument we then demonstrate that all trajectories starting in the open first (third) quadrant converge to  $(1,1)$  (respectively  $(-1,-1)$ ), which establishes the claim we made earlier, that all trajectories of (4.5.1) become asymptotically periodic.

#### 4.5.1 Describing $D_F$

$D_F$  is most easily characterized by investigating some basic properties of the family of curves  $\{C_n, n \in \mathbb{N}\}$ :

$$C_n = F^{-n} \left\{ \begin{bmatrix} 0 \\ y \end{bmatrix}, \begin{bmatrix} y \\ 0 \end{bmatrix} ; y \in \mathbb{R} \right\} ; n \in \mathbb{N} \quad (4.5.9)$$

which has to be deleted from the phase plane to obtain  $D_F$ . Notice firstly that it is easy to give  $\{C_n, n \in \mathbb{N}\}$  the structure of a one dimensional manifold in  $\mathbb{R}^2$ ; and that it (therefore) has Lebesgue measure zero. The following result holds:

Lemma 4.1:

- (i) the curves  $C_n$  are a subset of  $Q_2^{cl} \cup Q_4^{cl}$ .
- (ii) the curves  $C_n$  shrink towards the origin as  $n$  increases:

$$C_n \rightarrow \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\} \text{ as } n \uparrow \infty$$

- (iii) the curves  $C_n, n \geq 3$  are the boundaries of a sequence of compact sets,

denoted by  $S_n$ , containing in their interior  $C_{n+1}$ .  $\square$

The proof of this Lemma relies on the following result which investigates the dynamics of  $F^{-1}$  on  $Q_2^c \cup Q_4^c$ :

Lemma 4.2:

The union of the closed second and fourth quadrants ( $Q_2^c \cup Q_4^c$ ) is invariant under  $F^{-1}$ , and is contained in the domain of attraction of the origin, i.e.

$$F^{-n}(x) \rightarrow 0 \text{ as } n \rightarrow \infty \quad \forall x \in Q_2^c \cup Q_4^c \quad \square$$

Proof of Lemma 4.2: From the definition of  $F^{-1}$ , (4.5.10), it follows that  $F^{-1}$  is well defined on  $Q_2^c \cup Q_4^c$  and that

$$F^{-1}(Q_2^c \cup Q_4^c) \subset Q_2^c \cup Q_4^c. \quad (4.5.10)$$

Defining on  $Q_2^c \cup Q_4^c$ , the following Lyapunov function:

$$V(x) = |y_1| + |y_2|; \quad x = (y_1 \ y_2)^T$$

and evaluating  $V$  along an orbit  $\{F^{-n}(x) = x_n; n \in \mathbb{N}\}$  for an initial condition  $x$  in  $Q_2^c \cup Q_4^c$ , we obtain that

$$\begin{aligned} V(F^{-(n+1)}(x)) - V(F^{-n}(x)) \\ &= |y_1|_{n+1} + |y_2|_{n+1} - |y_1|_n - |y_2|_n \\ &= \frac{-2|y_1 y_2|_{n-1}}{1+2|y_1 y_2|_{n-1}} - \frac{2|y_1 y_2|_n}{1+2|y_1 y_2|_n} \\ &< 0. \end{aligned}$$

Equality can hold iff

$$|y_1 y_2|_{n-1} = 0 \text{ and } |y_1 y_2|_n = 0$$

but as this implies that  $x_k \equiv 0 \quad \forall k \geq n+1$ , we conclude that the origin is the unique fixed point of  $F^{-1}$  with a domain of attraction containing  $Q_2^c \cup Q_4^c$ .  $\square$

Proof of Lemma 4.1: Parts (i) and (ii) are direct consequences of the Lemma 4.2. Part (iii) follows readily from the observation that

$$\begin{aligned} C_2 &= \left\{ \left[ -y \quad \frac{y}{1+2y^2} \right]^T, y \in \mathbb{R} \right\} \\ C_3 &= \left\{ \left[ \frac{y}{1+2y^2} \quad \frac{-y}{1+4y^2} \right]^T, y \in \mathbb{R} \right\} \\ C_2 \cap C_3 &= \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\} \end{aligned}$$

$C_3$  encloses a compact set in  $Q_2^1 \cup Q_4^1$  (denoted by  $S_3$ ). Because  $F^{-1}$  is a homeomorphism on  $Q_2^1 \cup Q_4^1$ , it follows the  $C_n \cap C_{n+1} = \{(0, 0)^T\} \quad \forall n \in \mathbb{N}$  and because  $F^{-1}$  contracts  $Q_2^1 \cup Q_4^1$  to the origin,  $C_{n+1}$  must be contained in the compact set enclosed by  $C_n$  denoted by  $S_n$  ( $\delta(S_n) = C_n$ ,  $S_n \supset C_{n+1}$ ).  $\square$

The previous lemmata allow us to partition the domain  $D_F$  as follows:

$$D_0 = Q_3 \cup Q_1$$

$$D_1 = \left\{ \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} : y_1 y_2 < 0, |y_2| - |y_1| > 0 \right\}$$

$$D_2 = \left\{ \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} : y_1 y_2 < 0, |y_2| - |y_1| < 0, |y_2| - \frac{|y_1|}{2y_1^2 + 1} > 0 \right\}$$

$$D_3 = \left\{ \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} : y_1 y_2 < 0, |y_2| - \frac{|y_1|}{2y_1^2 + 1} < 0 \right\} \setminus S_3$$

$$D_{n+1} = \text{int}(S_n \setminus S_{n+1}); \quad n \geq 3.$$

By construction,  $D_F = \bigcup_{n \geq 0} D_n$ , the boundaries of these sets being

$$\delta(D_0) = C_0$$

$$\delta(D_n) = C_n \cup C_{n-1} \quad n \geq 1 \quad (4.5.11)$$

One can easily verify that:

$$F(D_n \cap Q_2) = D_{n-1} \cap Q_4$$

$$F(D_n \cap Q_4) = D_{n-1} \cap Q_2 \quad (4.5.12)$$

$$F(D_n) = D_{n-1}$$

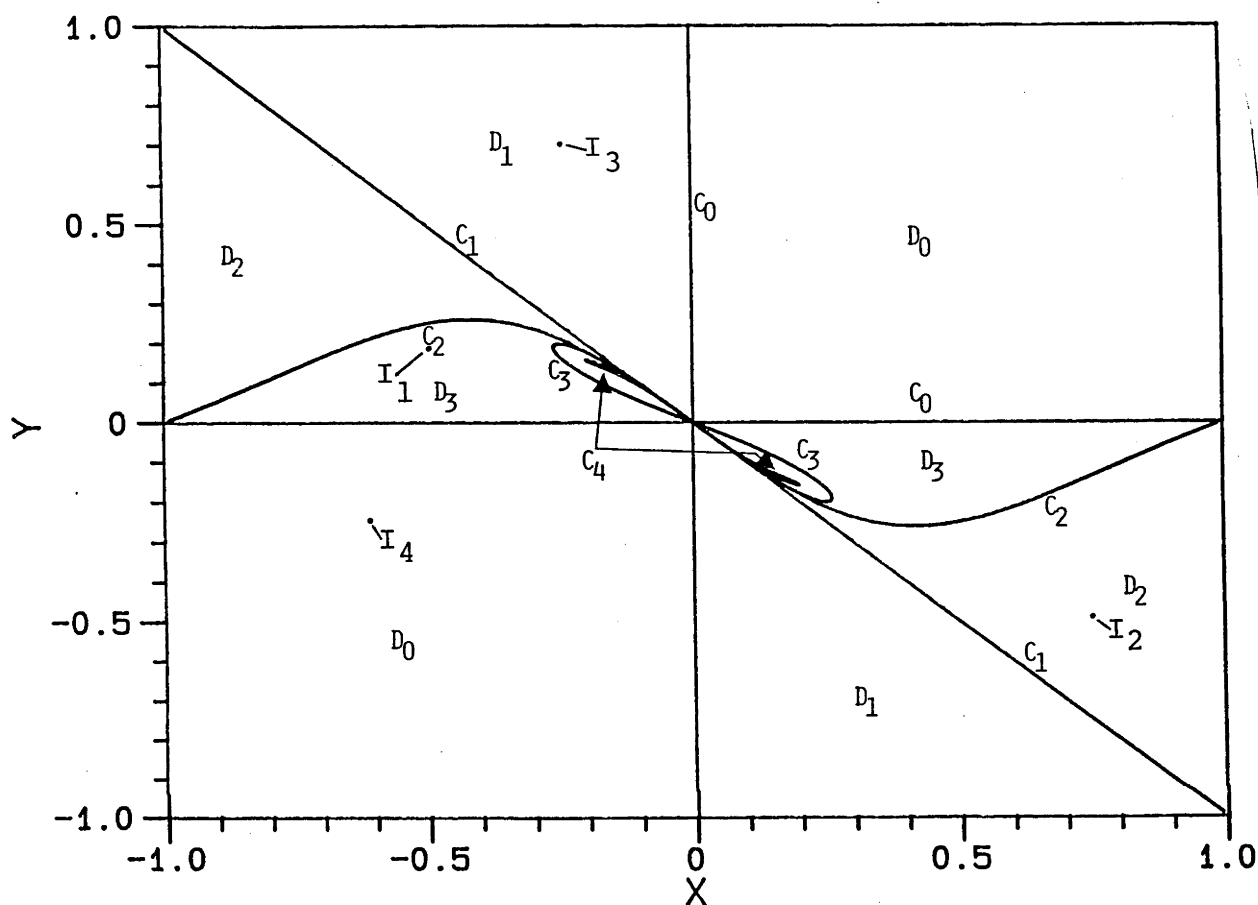
$$F(D_0) \subset D_0$$

Summarising the expressions (4.5.12) we have:

Lemma 4.3:

For (Lebesgue) almost all initial conditions - excluding initial conditions on  $D_F^c = \bigcup C_n$  - the orbits are well defined. Initial conditions in  $D_n$  will travel through the sets  $D_k$   $0 \leq k \leq n$ , alternating between  $Q_2$  and  $Q_4$ , finally reaching  $D_0 = Q_1 \cup Q_3$  and remain there.  $\square$

This lemma is illustrated in Figure 4.1, which displays some of the curves  $C_n$  ( $C_0$ - $C_4$ ) regions  $D_n$  ( $D_0$ - $D_4$ ) and an orbit starting in  $D_3 \cap Q_2$ , travelling through  $D_2 \cap Q_4$ ,  $D_1 \cap Q_2$ , finally arriving in  $D_0 \cap Q$ . (Notice that we have first applied  $H$ , cf. Section 4.2, (4.2.1)).

Figure 4.1 Domain of Definition of  $F$ 

Having established that the asymptotic dynamics are restricted to the first and third open quadrants in  $\mathbb{R}^2$ , we now turn to the analysis of  $F$  restricted to these open quadrant.

#### 4.5.2 The Dynamics of $F$ Restricted to $Q_1$ or $Q_3$

Because of the symmetry, we limit ourselves to  $F$  restricted to  $Q_1$ . Clearly,  $F(Q_1) \subset Q_1$ , and also  $F((1 \ 1)^T) = (1 \ 1)^T$ . The local stability properties of the fixed point  $(1 \ 1)^T$  are summarized in

##### Lemma 4.4:

The fixed point  $(1 \ 1)^T$  is locally exponentially stable in the sense of Lyapunov.  $\square$

Proof: Directly from linearization of  $F$  in the neighbourhood of  $(1 \ 1)^T$ :

$$DF \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} \\ 1 & 0 \end{bmatrix} \quad (4.5.13)$$

which is a stability matrix (its eigenvalues are less than one in modulus) with eigenvalues:

$$\lambda_{1,2} = \frac{-1 \pm i\sqrt{7}}{4} \quad (4.5.14)$$

$$|\lambda_{1,2}| = \frac{1}{\sqrt{2}} < 1 \quad (4.5.15)$$

□

The global stability properties of the fixed point  $(1 \ 1)^T$  are stated in:

Lemma 4.5:

For any initial condition  $x \in Q_1$ , the orbit  $\{F^n(x), n \in \mathbb{N}\}$  tends to  $(1 \ 1)^T$ . □

In the proof we make use of the following lemma:

Lemma 4.6:

For all strictly positive numbers  $\alpha, \beta \in \mathbb{R}^+$ , we have that:

$$\begin{aligned} \max \left[ \alpha, \beta, \frac{1}{\alpha}, \frac{1}{\beta} \right] &\geq \max \left[ \alpha, \frac{\alpha+\beta}{2\alpha\beta}, \frac{1}{\alpha}, \frac{2\alpha\beta}{\alpha+\beta} \right] \geq 1 \\ \min \left[ \alpha, \beta, \frac{1}{\alpha}, \frac{1}{\beta} \right] &\leq \min \left[ \alpha, \frac{2\alpha\beta}{\alpha+\beta}, \frac{1}{\alpha}, \frac{\alpha+\beta}{2\alpha\beta} \right] \leq 1 \end{aligned} \quad (4.5.16)$$

□

Proof of Lemma 4.6: Because of symmetry it is sufficient to look at the situations  $\alpha \geq \beta \geq 1$  and  $\beta^{-1} \geq \alpha \geq 1$  and  $\alpha^{-1} \geq \beta \geq 1$ . In the first case

$$\alpha = \max \left[ \alpha, \beta, \frac{1}{\alpha}, \frac{1}{\beta} \right] = \max \left[ \alpha, \frac{\alpha+\beta}{2\alpha\beta}, \frac{1}{\alpha}, \frac{2\alpha\beta}{\alpha+\beta} \right] \geq 1$$

$$\frac{1}{\alpha} = \min \left[ \alpha, \beta, \frac{1}{\alpha}, \frac{1}{\beta} \right] = \min \left[ \alpha, \frac{2\alpha\beta}{\alpha+\beta}, \frac{1}{\alpha}, \frac{\alpha+\beta}{2\alpha\beta} \right] \leq 1$$

in the second case

$$\frac{1}{\beta} = \max \left[ \alpha, \beta, \frac{1}{\alpha}, \frac{1}{\beta} \right] \geq \max \left[ \alpha, \frac{\alpha+\beta}{2\alpha\beta}, \frac{1}{\alpha}, \frac{2\alpha\beta}{\alpha+\beta} \right] \geq 1$$

$$\beta = \min \left[ \alpha, \beta, \frac{1}{\alpha}, \frac{1}{\beta} \right] \leq \min \left[ \alpha, \frac{2\alpha\beta}{\alpha+\beta}, \frac{1}{\alpha}, \frac{\alpha+\beta}{2\alpha\beta} \right] \leq 1$$

in the last case

$$\frac{1}{\alpha} = \max \left[ \alpha, \beta, \frac{1}{\alpha}, \frac{1}{\beta} \right] = \max \left[ \alpha, \frac{\alpha+\beta}{2\alpha\beta}, \frac{1}{\alpha}, \frac{2\alpha\beta}{\alpha+\beta} \right] \geq 1$$

$$\alpha = \min \left[ \alpha, \beta, \frac{1}{\alpha}, \frac{1}{\beta} \right] = \min \left[ \alpha, \frac{2\alpha\beta}{\alpha+\beta}, \frac{1}{\alpha}, \frac{\alpha+\beta}{2\alpha\beta} \right] \leq 1$$

From this the lemma follows.  $\square$

Observe also that equality (in 4.5.16) can hold iff

$$\max(\alpha, \frac{1}{\alpha}) = \max(\alpha, \frac{1}{\alpha}, \beta, \frac{1}{\beta})$$

$$\min(\alpha, \frac{1}{\alpha}) = \min(\alpha, \frac{1}{\alpha}, \beta, \frac{1}{\beta})$$

Proof of Lemma 4.5: In the domain  $Q_1$ , we define the following "Lyapunov" function

$$V(x) = \max(y_1, y_2, \frac{1}{y_1}, \frac{1}{y_2}) - 1$$

$$x = (y_1 \ y_2)^T \in Q_1$$

This function is positive definite:

$$V(x) \geq 0; \ x \in Q_1 \quad V((1 \ 1)^T) = 0$$

and radially unbounded:

$$\lim_{x \rightarrow \partial(Q_1)} V(x) = +\infty$$

Evaluated along an orbit  $\{F^n(x) = x_n, \ n \in \mathbb{N}\}$ ,  $x \in Q_1$ ,

we obtain: (denote  $x_k = (y_{1k} \ y_{2k})^T$ )

$$V(x_{k+1}) - V(x_k) = \max(y_{1k+1}, y_{2k+1}, \frac{1}{y_{1k+1}}, \frac{1}{y_{2k+1}})$$

$$- \max(y_{1k}, y_{2k}, \frac{1}{y_{1k}}, \frac{1}{y_{2k}})$$

$$= \max\left(\frac{y_{1k}+y_{2k}}{2y_{1k}y_{2k}}, \frac{2y_{1k}y_{2k}}{y_{1k}+y_{2k}}, y_{1k}, \frac{1}{y_{1k}}\right)$$

$$- \max(y_{1k}, \frac{1}{y_{1k}}, y_{2k}, \frac{1}{y_{2k}})$$

Hence, from the previous lemma we obtain:

$$V(x_{k+1}) - V(x_k) \leq 0$$

with equality only holding if

$$\max(y_{1k}, y_{2k}, \frac{1}{y_{1k}}, \frac{1}{y_{2k}}) = \max(y_{1k}, \frac{1}{y_{1k}})$$

Consequently  $V(x_{k+1}) - V(x_k) \equiv 0$  iff  $y_{2k} \equiv y_{1k} \equiv 1$ , this proves the result.  $\square$

Remark:

(R.4.8) It is possible to obtain good estimates for local (but not  $\epsilon$ -small) regions containing  $(1 \ 1)^T$  wherein  $F$  is a contraction in some norm. Because this does not add substantially to our knowledge about the dynamics of  $F$ , we do not pursue this result in the sequel, local exponentially stability will prove to be sufficient.  $\square$

We have now obtained all the necessary information to describe the global dynamics of the state transition map  $F$  and to characterize the adaptive response for the situation  $b < 0$ .

#### 4.5.3 Global Dynamics - Consequences for the Adaptive System

Linking the previous lemmata, we obtain the following picture for the global dynamics of  $F$  on  $D_F$ .

Theorem 4.1:

- (i) For any initial condition  $x_0$  in  $D_F$  ( $\equiv$  for Lebesgue almost all initial conditions in  $\mathbb{R}^2$ ) the orbit  $\{F^k(x_0), k \in \mathbb{N}\}$  (see equations (4.5.5), (4.5.6)) is well defined: there exist positive numbers  $m, M$  depending on  $x_0$ , such that

$$0 < m < |F^k(x_0)| < M < \infty, \quad \forall k \in \mathbb{N}$$

- (ii) Any orbit converges to either  $(1 \ 1)^T$  or  $(-1 \ -1)^T$ . The domain of attraction  $A_1$  of the fixed point  $(1 \ 1)^T$  is:

$$A_1 = (D_0 \cap Q_1) \cup \bigcup_{n \geq 1} \{(D_{2n-1} \cap Q_4) \cup (D_{2n} \cap Q_2)\} \quad (4.5.17)$$

The domain of attraction  $A_{-1}$  of the fixed point  $(-1 \ -1)^T$  is:

$$A_{-1} = (D_0 \cap Q_3) \cup \bigcup_{n \geq 1} \{(D_{2n-1} \cap Q_2) \cup (D_{2n} \cap Q_4)\} \quad (4.5.18)$$

- (iii) Locally at the fixed points the convergence is exponential.  $\square$



This result has immediate consequences for the adaptive closed loop system. Recall that, the parameter estimate/feedback gain is given by (cf. equations (4.5.2), (4.5.1), (4.3.12), and the link equations (4.3.9), (4.3.10)) :

$$\hat{a}_k = a + \left[ \frac{|b|}{2} \right]^{\frac{1}{2}} \frac{(-1)^{k-1}}{w_{k-1}} \quad (4.5.19)$$

and that the output of the plant  $y_k$  is given by:

$$y_k = (-1)^k \sqrt{2|b|} w_k y_{k-1} \quad (4.5.20)$$

From these expressions and the previous theorem we obtain:

**Theorem 4.2:**

For almost all initial conditions, except for a set of Lebesgue measure zero, and for all parameter values  $a, b < 0$  the adaptive closed-loop system described by (4.3.8)-(4.3.10) produces a bounded parameter estimate feedback gain, which exponentially becomes periodic with period two:

$$\hat{a}_k \rightarrow a \pm (-1)^k \left[ \frac{|b|}{2} \right]^{\frac{1}{2}} \quad \text{as } k \uparrow \infty \quad \square$$

Proof: Follows directly from Theorem 4.1 and equation (4.5.19).  $\square$

**Theorem 4.3:**

For Lebesgue almost all initial conditions, and for all parameter values  $a$ , the adaptive closed loop system is stable for all  $b : 0 \geq b \geq -\frac{1}{2}$ , in the sense that the state vector is bounded. Moreover for  $0 \geq b > -\frac{1}{2}$  the output is regulated exponentially to zero:

$$y_k \rightarrow 0 \text{ as } k \uparrow \infty \quad 0 \geq b > -\frac{1}{2}$$

For  $b < -\frac{1}{2}$  the adaptive system is unstable, and  $y_k$  diverges exponentially.  $\square$

Proof: Because,

$$y_k = (-1)^k \sqrt{2|b|} w_k y_{k-1} \quad (4.5.20)$$

and because  $|w_k|$  converges exponentially to 1, we have that, for almost all initial conditions  $y_0, y_{-1}, y_{-2}$ :

$$|y_k| < C(y_0, y_{-1}, y_{-2})(\sqrt{2|b|})^k \quad (4.5.21)$$

where  $C$  is a positive constant, depending on the initial conditions, but bounded. From (4.5.21), the result is obvious.  $\square$

Remark:

(R.4.9) It is possible to give an explicit estimate for the rate of exponential convergence, and to prove that the convergence is more than "asymptotical" (see (R.4.8)).  $\square$

(R.4.10) Notice that it is possible to obtain both plant parameters  $a, b$  from the observation of the  $\hat{a}_k$  sequence! Indeed, using the conclusion of Theorem 4.2 it follows that:

$$\lim_{N \uparrow \infty} \frac{1}{N} \sum_{k=1}^N \hat{a}_k = a \quad (4.5.22)$$

$$\lim_{N \uparrow \infty} \frac{1}{N} \sum_{k=1}^N (\hat{a}_k - a)^2 = b/2 \quad (4.5.23)$$

(The estimates converge to their limiting values as  $1/N$ .)  $\square$

These results and their counterparts of the next section which treat the case  $b > 0$ , are commented upon in Section 7, because we prefer to discuss the implications of our findings when we have a more complete picture of the closed loop dynamics, described over the whole parameter range.

#### 4.6 The Closed Loop Dynamics III: $b > 0$

In this section we present both analytical and numerical evidence for the presence of chaos in the equation:

$$v_k = \frac{1}{v_{k-1}} - \frac{1}{v_{k-2}} \quad k \in \mathbb{N} \quad (4.6.1)$$

which describes the dynamics of the normalized ratios of the successive plant outputs ( $v_k = y_k/(b^{1/2}y_{k-1})$ ) in the situation  $b > 0$ , (see (4.3.12)-(4.3.13)) and investigate what this implies for the closed loop adaptive system.

We use the following state space representation. Define the state vector as:

$$x_k = \begin{bmatrix} v_k \\ v_{k-1} \end{bmatrix} \quad (4.6.2)$$

and the state transition map as:

$$G \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{y_1} - \frac{1}{y_2} \\ y_1 \end{bmatrix} \quad (4.6.3)$$

the difference equation (4.6.1) can then be represented as

$$x_{k+1} = G(x_k); \quad x_0 \quad k \in \mathbb{N} \quad (4.6.4)$$

The recursion (4.6.4) is well defined on:

$$D_G = \mathbb{R}^2 \setminus \bigcup_{n \geq 0} G^{-n} \left\{ \begin{bmatrix} y \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ y \end{bmatrix}; y \in \mathbb{R} \right\}$$

the recursion (4.6.4) can be inverted:

$$x_k = G^{-1}(x_{k+1}) \quad (4.6.5)$$

where  $G^{-1}$  is defined by:

$$G^{-1} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} z_2 \\ \frac{z_2}{1 - z_1 z_2} \end{bmatrix} \quad (4.6.6)$$

the backward recursion is well defined on the set

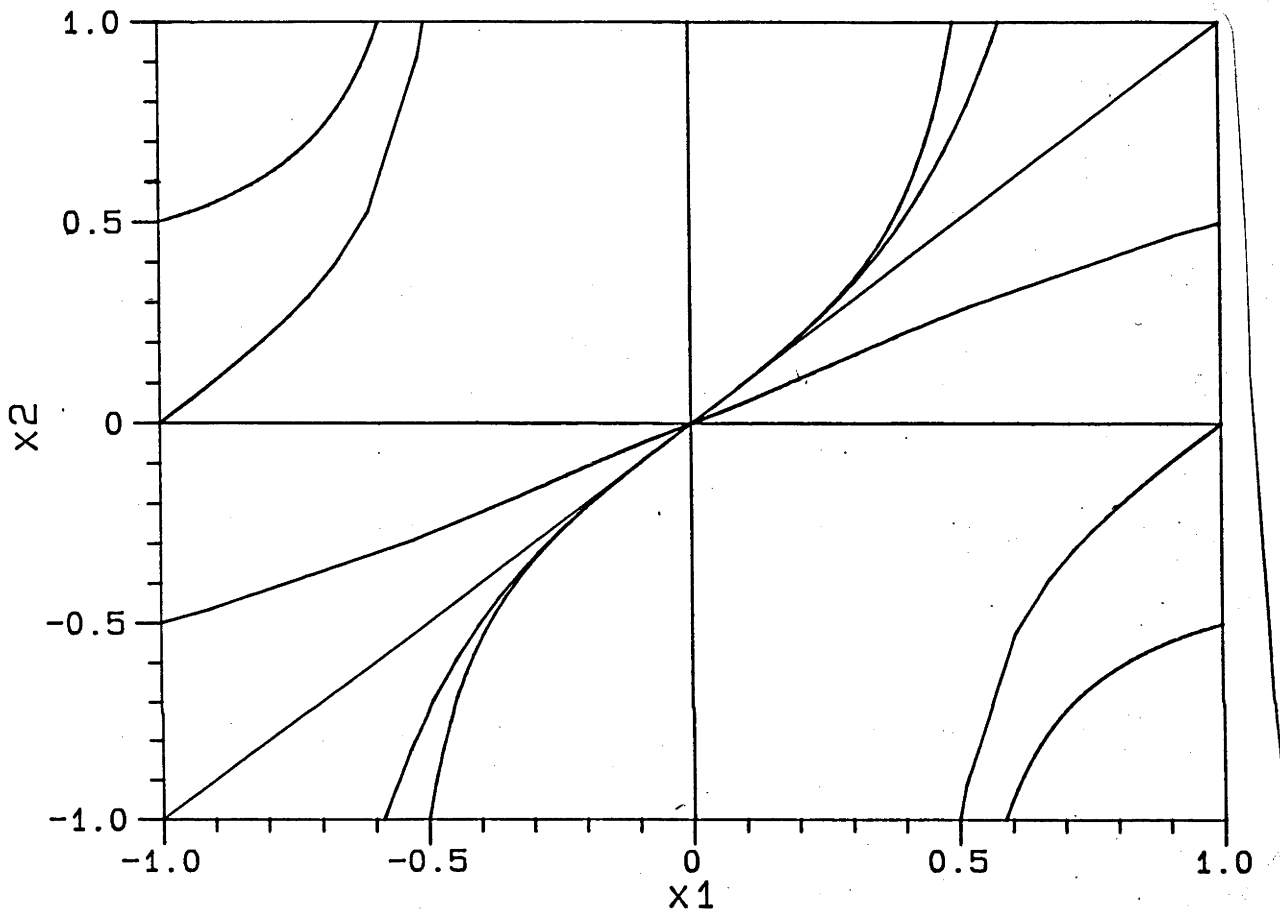
$$D_{G^{-1}} = \mathbb{R}^2 \setminus \bigcup_{n \geq 0} G^n \left\{ \begin{bmatrix} z \\ 1/z \end{bmatrix}, \begin{bmatrix} 1/z \\ z \end{bmatrix}; z \in \mathbb{R} \right\}$$

We will describe the dynamics of (4.6.4) on the set  $D = D_{G^{-1}} \cap D_G$

In precise terms we are interested in the dynamics on the two dimensional open set  $D$  of the map  $G : D \rightarrow D$ . On  $D$ ,  $G$  is a diffeomorphism.

#### Remarks:

(R.4.11) The complement of  $D$  in  $\mathbb{R}^2$ , is a set of measure zero, because it is the union of a countable number of curves in  $\mathbb{R}^2$ .  $D^c$  is partly illustrated in Figure 4.2: Domain of Definition. The figure shows the collection of the first three curves  $G^{-n}((0 \ y)^T, (y \ 0)^T; y \in \mathbb{R})$   $n=0,1,2$  which have to be deleted from  $\mathbb{R}^2$ . (We remind the reader about the compactification of  $\mathbb{R}^2$  by the homeomorphism  $H$  (4.2.1)).  $\square$

Figure 4.2 Domain of Definition

In the sequel we demonstrate that the difference equation (4.6.1) displays chaos. These complicated dynamics are exhibited in Figures 4.3 and 4.4. Figure 4.3 gives the time portrait of the  $v_k$ -sequence, while Figure 4.4 contains the corresponding state space portrait. (Figure 4.3 contains only the first 500 samples, while Figure 4.4 contains 40,000 samples.) (Notice the distortion of the scale due to the application of the homeomorphism  $H$  (4.2.1).)

Figure 4.3 Time Series  $\{v_k; k \in \mathbb{N}\}$

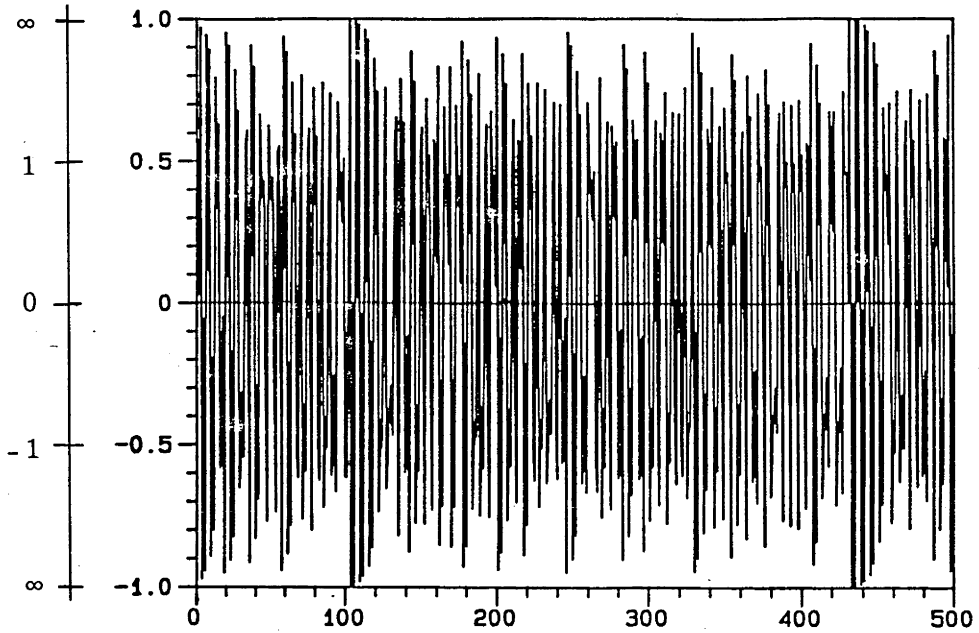
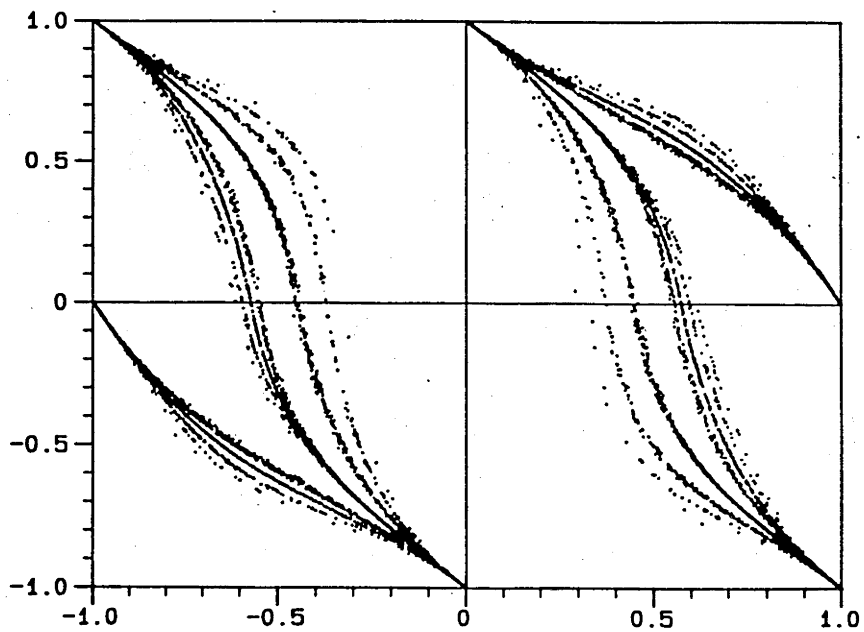


Figure 4.4 Gumleaf Attractor  $\{x_k = (v_k \ v_{k-1})^T, k \in \mathbb{N}\}$



Before proceeding, we list here one of the possible definitions of chaotic dynamics. We refer to [1,4] for more information.

Definition 4.1:

The difference equation (4.6.4) is chaotic if there exists an invariant set  $S \subset \mathbb{R}^2$  ( $G(S) \subset S$ ) containing sets  $P$ ,  $A_1$  and  $A_2$  with the properties:

- (i) (4.6.4) has a countably infinite number of periodic solutions with all periods above a certain integer;  $P$  is the collection of all the points visited by these trajectories
- (ii) (4.6.4) has an uncountably infinite number of aperiodic solutions which never become asymptotically periodic;  $A_1$  is the collection of all the points visited by these trajectories

$$\forall u_0 \in P, \forall y_0 \in A_1 : \limsup_{k \rightarrow \infty} \|G^k(u_0) - G^k(y_0)\| > 0$$

- (iii)  $\exists \epsilon > 0 \forall u_0, y_0 \in A_1 : u_0 \neq y_0 : \limsup_{k \rightarrow \infty} \|G^k(u_0) - G^k(y_0)\| > \epsilon$

(all aperiodic orbits separate)

- (iv)  $A_1$  contains an uncountable subset  $A_2$  such that

$$\forall u_0, y_0 \in A_2 : \liminf_{k \rightarrow \infty} \|G^k(u_0) - G^k(y_0)\| = 0$$

Remarks:

(R.4.12) Observe that the definition of chaos is rather academic. Any finite wordlength implementation of a difference equation - even if this difference equation is chaotic - has a finite state space and hence cannot be chaotic in the sense of Definition 4.1. The practical importance however is that whenever an algorithm has chaos according to Definition 4.1, its output from any reasonable implementation will demonstrate an effective unpredictability of future behaviour without infinite precision knowledge of the present state. In this way it may closely resemble a random process.  $\square$

Properties (iii) and (iv) express an extreme sensitivity of the trajectories of (4.6.4) to changes in initial conditions. Trajectories (belonging to  $A_2$ ) merge and separate consecutively and continuously in time; coming closer to each other but then being forced to separate to at least a distance  $\epsilon$  from each other. The phenomenon is easily understood when the only finite orbits of the difference equation are of saddle type - i.e. have one dimensional stable and unstable manifolds, and when their manifolds form a dense web. Close to stable manifolds trajectories are attracted to each other, whilst along the unstable

manifolds they are separated. This observation is really the key to an intuitive understanding of the complete dynamical behaviour of the present closed loop adaptive system. We will expand upon this idea, but first we demonstrate the presence of chaos.

The approach is standard for the analysis of any particular dynamical system. First we search for fixed points and periodic orbits and investigate their local stability properties. Having convinced ourselves that there are apparently a countably infinite number of periodic orbits, we look for the presence of a homoclinic orbit which would explain their presence. A transversal homoclinic orbit also indicates the presence of chaotic dynamics as defined above. Finally we try to understand in intuitively appealing terms what chaos means in this context and discuss some of its implications for the underlying adaptive control problem.

#### 4.6.1 Fixed Points, Periodic Orbits, Local Stability Properties

The first result lists some of the more elementary properties of the trajectory  $\{v_k, k \in \mathbb{N}\}$  as generated by (4.6.1).

**Lemma 4.7:**

For any initial condition  $(v_0, v_1)$  the trajectory of (4.6.1)  $\{v_k(v_0, v_1), k \in \mathbb{N}\}$  ( $v_0(v_0, v_1) = v_0$  and  $v_1(v_0, v_1) = v_1$ ) has the properties:

- (i)  $v_k(-v_0, -v_1) = -v_k(v_0, v_1)$
- (ii) alternating sign patterns cannot occur; precisely if  $v_k > 0$  and  $v_{k+1} < 0$  then  $v_{k+2} < 0$ , or if  $v_k < 0$  and  $v_{k+1} > 0$ , then  $v_{k+2} > 0$ .
- (iii) the sequence either terminates (i.e.  $\exists k^*: v_{k^*} = 0$ ) or has indefinitely many sign changes. □

**Proof:** (i) and (ii) are immediately clear from (4.6.1). For (iii) suppose that for the initial condition  $(v_0^*, v_1^*)$  along the trajectory  $v_k(v_0^*, v_1^*) > 0 \forall k$ . Using (4.6.1) it follows that  $v_k(v_0^*, v_1^*)$  is a strictly decreasing sequence bounded below by zero, hence converges. But it is clear from (4.6.1) that there does not exist a fixed point. Hence the assumption was invalid, which demonstrates the result. □

Searching for fixed points and periodic orbits requires solving  $G^P(x) = x$ ,  $p \in \mathbb{N}_0$ , or alternatively solving a set of algebraic equations (derived from (4.6.1)):

$$v_1 = \frac{1}{v_p} - \frac{1}{v_{p-1}} \quad (4.6.7.1)$$

$$v_2 = \frac{1}{v_1} - \frac{1}{v_p} \quad (4.6.7.2)$$

$$v_k = \frac{1}{v_{k-1}} - \frac{1}{v_{k-2}} \quad k = 3, \dots, p \quad (4.6.7.3)$$

From a careful examination of sign patterns of potential fixed points and Lemma 4.7 it follows that  $G^p(x) = x$  has no real solutions for  $0 < p < 4$ . There is a unique orbit of period four, which can be found analytically by solving (4.6.7) with  $p = 4$ :

$$\begin{aligned} x_1 &= \begin{bmatrix} \sqrt{2 + \sqrt{2}} \\ -\sqrt{2 - \sqrt{2}} \end{bmatrix}, \quad x_2 = \begin{bmatrix} -\sqrt{2 - \sqrt{2}} \\ -\sqrt{2 + \sqrt{2}} \end{bmatrix}, \quad x_3 = \begin{bmatrix} -\sqrt{2 + \sqrt{2}} \\ \sqrt{2 - \sqrt{2}} \end{bmatrix} \\ x_4 &= \begin{bmatrix} \sqrt{2 - \sqrt{2}} \\ \sqrt{2 + \sqrt{2}} \end{bmatrix} \end{aligned} \quad (4.6.8)$$

$x_i$   $i=1, \dots, 4$  are the fixed points of  $G^4$ . Numerically, we verified the existence of periodic orbits for periods up to 55, by solving the set of algebraic equations (4.6.7). Despite the apparent symmetry, we were not able to prove analytically that there indeed exist periodic orbits of all periods greater than four.

Through linearization, we analysed the local stability properties of these periodic orbits. We verified numerically that all periodic orbits we established (up to period 55) are of saddle type, i.e. they possess a one dimensional stable and a one dimensional unstable manifold. In particular, for the periodic orbit of period four; we have that the Jacobian of  $G^4$  at  $x_1$  (which determines the local stability properties) is given by

$$DG^4(x_1) = \begin{bmatrix} \frac{3 - \sqrt{2}}{2} & -\frac{3 + 2\sqrt{2}}{2} \\ -\frac{2 - \sqrt{2}}{2} & \frac{2 + \sqrt{2}}{2} \end{bmatrix}^2 \quad (4.6.9)$$

$$\lambda_u(DG^4(x_1)) = \left[ \frac{5 + \sqrt{17}}{4} \right]^2 > 1 \quad (4.6.10.1)$$



$$\lambda_S(DG^4(x_1)) = \left[ \frac{5 - \sqrt{17}}{4} \right]^2 < 1 \quad (4.6.10.2)$$

Locally, at  $x_1$ , the stable manifold is tangent to:

$$\xi_S = \begin{bmatrix} -0.043115 \\ -0.999070 \end{bmatrix}$$

the eigenvector of  $DG^4(x_1)$  corresponding to  $\lambda_S$ ; and the unstable manifold is tangent to:

$$\xi_U = \begin{bmatrix} 0.758652 \\ -0.651497 \end{bmatrix}$$

the eigenvector of  $G^4(x_1)$  corresponding to  $\lambda_U$ .

#### 4.6.2 A Horseshoe in an Iteration of $G$

The global stable and global unstable manifold are defined respectively as the union of all backward iterations of the local stable manifold and as the union of all forward iterations of the local unstable manifold [1] (see also Chapter 3, Section 3.3). Because the stable manifold is attractive under the inverse map, and the unstable manifold is attractive under the forward map, these manifolds can be computed in a numerically stable way, using the definition to construct them.

In this way we constructed partially the stable and unstable manifolds of the fixed point  $x_1$  of  $G^4$ . (This is also part of the stable and unstable manifold of the periodic orbit of period 4 of  $G$ .) We iterated backwards, under  $G^{-4}$ , three neighbourhoods  $S_{-1}$ ,  $S_0$ ,  $S_1$ ; intervals on three parallel straight lines oriented along the stable eigenvector  $\xi_S$ ,  $S_0$  centred on  $x_1$ ,  $S_{-1}$  and  $S_1$  centred respectively on  $x_1 - \epsilon_S$ ,  $x_1 + \epsilon_S$ . Analogously, we iterated forwards, under  $G^4$ , three neighbourhoods  $U_{-1}$ ,  $U_0$ ,  $U_1$ , intervals on three straight lines oriented along the unstable eigenvector  $\xi_U$ ,  $U_0$  centred on  $x_1$  and  $U_{-1}$ ,  $U_1$  centred respectively on  $x_1 - \epsilon_U$ ,  $x_1 + \epsilon_U$ . (See Figure 4.5.1 for the precise configuration. Figure 4.5.1 displays a neighbourhood of the fixed point  $x_1$  enclosed by  $S_{-1}$ ,  $S_{+1}$ ,  $U_{-1}$  and  $U_{+1}$ ). We find that the curves

$$GS_i = G^{-4n_S}(S_i) \quad i = -1, 0, 1$$

$$GU_j = G^{4n_u}(U_j) \quad j = -1, 0, 1$$

intersect each other, for  $n_s$  and  $n_u$  sufficiently large, transversally in "new" points, called homoclinic points. (See Figure 4.5.2 for precise configuration, the transversal intersections are denoted by HP. Notice in particular that the separate curves  $GS_i$  are indistinguishable as are the curves  $GU_j$ . This illustrates the fact that the described method for constructing the manifolds is (numerically) stable. In view of this fact, and in view of the continuity properties of  $G$  and because "transversal intersection" is a property which persists under slight perturbations, i.e. is structurally stable, we conclude that the stable and unstable manifolds of the periodic orbit of period four intersect transversally in a homoclinic point. With the existence of a homoclinic point established, the Smale-Birkhoff Homoclinic Theorem (Guckenheimer and Holmes, section 5.3, [1]) asserts that:

Theorem 4.4:

There exists a zero dimensional hyperbolic invariant set on which an iteration of  $G^4$  is topologically equivalent to a subshift of finite type.

□

Remarks:

(R.4.13) Decoded, this Theorem states that the difference equation  $z_{k+1} = G^l(z_k)$  is chaotic in the sense of Definition 4.1, where  $l$  is a multiple of 4. The chaotic dynamics established by the Smale-Birkhoff Homoclinic Theorem are of the horseshoe-type. (The Smale horseshoe is the prototype example for chaotic dynamics, for a discussion see e.g. [1] section 5.1. The horseshoe map was originally defined by Smale in terms of stretching and bending of an area in  $R^2$  [5]). ) Notice that at best  $\{v_k, k \in \mathbb{N}\}$  is a collection of  $l$  interleaved chaotic processes.

□

(R.4.14) It is clear from the presence of the asymptotic lines in the manifolds, Figure 4.5.2 that arbitrarily close to the fixed point  $x_1$ , there are points belonging to  $D^c$ , i.e. for which not all iterations of  $G$  or its inverse are defined.

Figure 4.5.1 Neighbourhood of the Fixed Point  $x_i$ : Local Manifolds

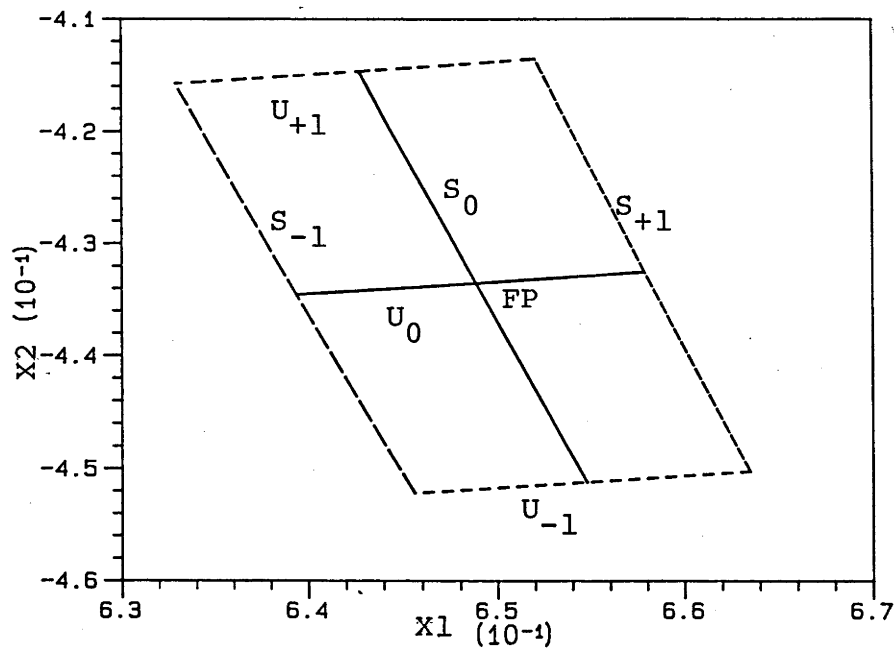
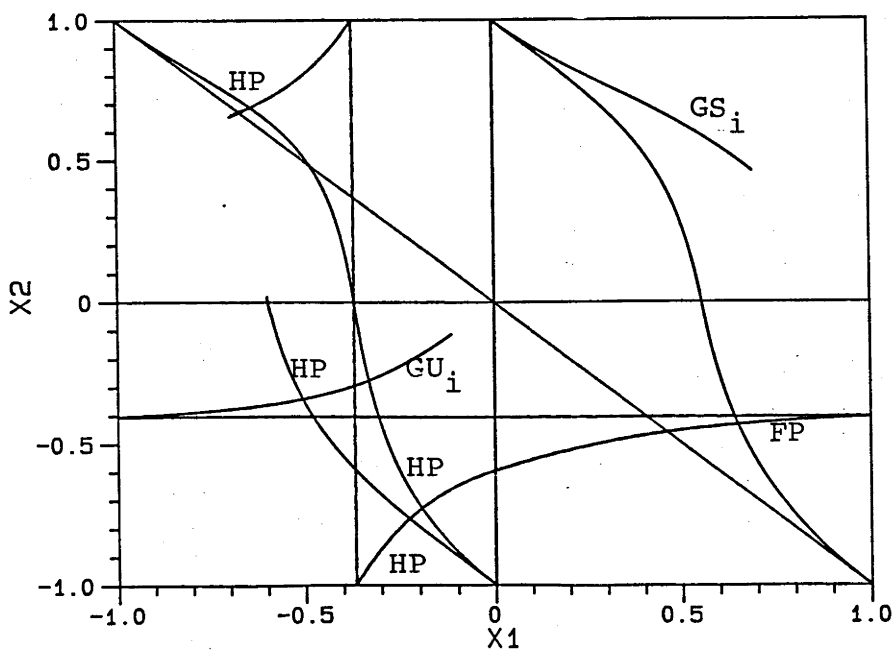


Figure 4.5.2 Global Manifolds and Homoclinic Points



More precisely, for any given  $\epsilon$ -neighbourhood of  $x_1$ , there exist points  $x \in D^c$ , belonging to this neighbourhood for which there exist positive integers  $n_1, n_2$  depending on  $\epsilon$ :

$$x \in G^{-n_1} \left\{ \begin{bmatrix} 0 \\ y \end{bmatrix}, y \in \mathbb{R} \right\} \cap G^{n_2} \left\{ \begin{bmatrix} z \\ 1 \\ z \end{bmatrix}, z \in \mathbb{R} \right\} \quad \square$$

(R.4.15) In the construction of the horseshoe, using the Smale-Birkhoff Homoclinic Theorem, it is clear that the local area around the fixed point is stretched to infinity under the action of  $G$  before it is bent over itself. Hence the divergence (convergence) along the unstable manifold under the action of  $G$  (of  $G^{-1}$ ) or along the stable manifold under the action of  $G^{-1}$  (of  $G$ ) must be faster than exponential. Only "locally" at  $x_1$ , do points separate (converge) exponentially. The phenomenon is of course due to the highly nonlinear features of the map  $G$ , and its peculiar behaviour at infinity.  $\square$

(R.4.16) In a similar manner, we can proceed for the other periodic orbits (period  $> 4$ ). The results are the same, yielding more horseshoes for different iterations of  $G$ . Such a procedure cannot convince us of the chaotic nature (in the sense of Definition 4.1) of  $G$  itself as there is no stopping rule. Therefore it seems pointless to pursue this.  $\square$

#### 4.6.3 A Period Doubling Route to Chaos in $G$

Having established the existence of chaotic behaviour in some iterations of  $G$ , and noting the symmetry of the problem and the importance of the behaviour at infinity, it is not hard to believe that  $G$  itself must be chaotic (cf. (R.4.13)). In this section we strengthen this by looking at the bifurcation diagram (cf. Chapter 3, Sections 3.3 and 3.7) for the one parameter family of maps:

$$G_c \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \frac{y_1}{c+y_1^2} - \frac{y_2}{c+y_2^2} \\ y_1 \end{bmatrix}; \quad c > 0 \quad (4.6.11)$$

On  $D$ ,  $G_c$  converges pointwise to  $G$  as  $c$  approaches zero from above:

$$G = \lim_{c \downarrow 0} G_c|_D \quad (4.6.12)$$

Some easily established results are:

- (i) The origin is a fixed point  $\forall c > 0$ , which is globally attractive

for  $c > 1$  and a locally unstable node for all  $c < 1$ .

(ii) For  $c < 1$  there is an orbit of period six

$$\left\{ \begin{bmatrix} \alpha \\ \alpha \end{bmatrix}, \begin{bmatrix} 0 \\ \alpha \end{bmatrix}, \begin{bmatrix} -\alpha \\ 0 \end{bmatrix}, \begin{bmatrix} -\alpha \\ -\alpha \end{bmatrix}, \begin{bmatrix} 0 \\ -\alpha \end{bmatrix}, \begin{bmatrix} \alpha \\ 0 \end{bmatrix} \right\} \quad (4.6.13)$$

$$\alpha = \sqrt{1 - c}$$

which locally is exponentially stable for all  $c$ :  $0.40775 \leq c < 1$ , and is unstable for all  $c$ :  $0 < c < 0.40774$ . (The correct boundary value is the zero of:

$$\frac{2c-1}{c} (6c^2-7c+3) + \frac{(2c-1)^2}{c^2} (8c^2-11c+4) + 1 = 0 \Big].$$

(iii) For  $c < 0.25$  there is a four periodic orbit,

$$\left\{ \begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \begin{bmatrix} -\beta \\ \alpha \end{bmatrix}, \begin{bmatrix} -\alpha \\ -\beta \end{bmatrix}, \begin{bmatrix} \beta \\ -\alpha \end{bmatrix} \right\} \quad (4.6.14)$$

which is initially,  $c$  close to 0.25, locally exponentially stable; then bifurcates into a saddle periodic orbit, forming two asymmetric stable periodic orbits of period four. This (symmetric) orbit converges to the periodic orbit of period four of  $G$ .

Figure 4.6 gives the numerically established bifurcation diagram. Horizontally, the bifurcation parameter  $c$  is represented and vertically the  $\omega$ -limit set of stable periodic orbits is represented by the first co-ordinate of the state (e.g. the six periodic orbit is represented as  $\{\alpha, 0, -\alpha, -\alpha, 0, \alpha\}$ ). This bifurcation diagram indicates the period doubling route to chaos (cf. section 6.8 in [1]).

Notice that a symmetric periodic orbit first bifurcates into two stable asymmetric periodic orbits of the same period (symmetry breaking bifurcation), then for a smaller value of the parameter these asymmetric periodic orbits undergo a period doubling bifurcation. For small  $c$  values, after a series of period doubling bifurcations, finally chaos emerges. (This information is not obtainable from the picture in Figure 4.6, due to the projection of the state onto its first co-ordinate; but is apparent from the way these bifurcations have to operate [1] and from the simulations.)

The diagram is obtained by running the difference equation for consecutive

values of  $c$ , detecting periodic behaviour and then plotting out the  $\omega$ -limit set for that value of the parameter. This is repeated for different initial conditions in order to obtain the two copies of every asymmetrical periodic orbit. (Because of the symmetry of the equation it suffices to take an initial condition out of the first and third quadrant in order to obtain the whole picture.) (Notice that the vertical axis is transformed into the interval  $(-1,1)$  as in the other figures.)

The diagram in Figure 4.6.1 contains the information for 1000  $c$ -values. the only easily recognisable bifurcations are the bifurcations involving the orbits of period six and four. If we used a finer resolution in  $c$ , and magnified the scale of both axes, the now dark patches would show similar bifurcations (symmetry breaking followed by period doubling) of periodic orbits of periods five, seven, eleven ... (e.g. Figure 4.6.2).

Remarks:

(R.4.17) Although the family  $G_c$  does not depend continuously on  $c$  at  $c=0$ , (which can best be seen from the fact that the symmetric six periodic orbit cannot exist for  $c=0$ , at least not with finite amplitude) we believe that this period doubling sequence indeed indicates that  $G$  is chaotic. It does provide substantial evidence for the earlier observation that all periodic orbits are unstable!  $\square$

Figure 4.6.1 Bifurcation Diagram

PD period doubling, SB symmetry breaking  
HB Hopf bifurcation

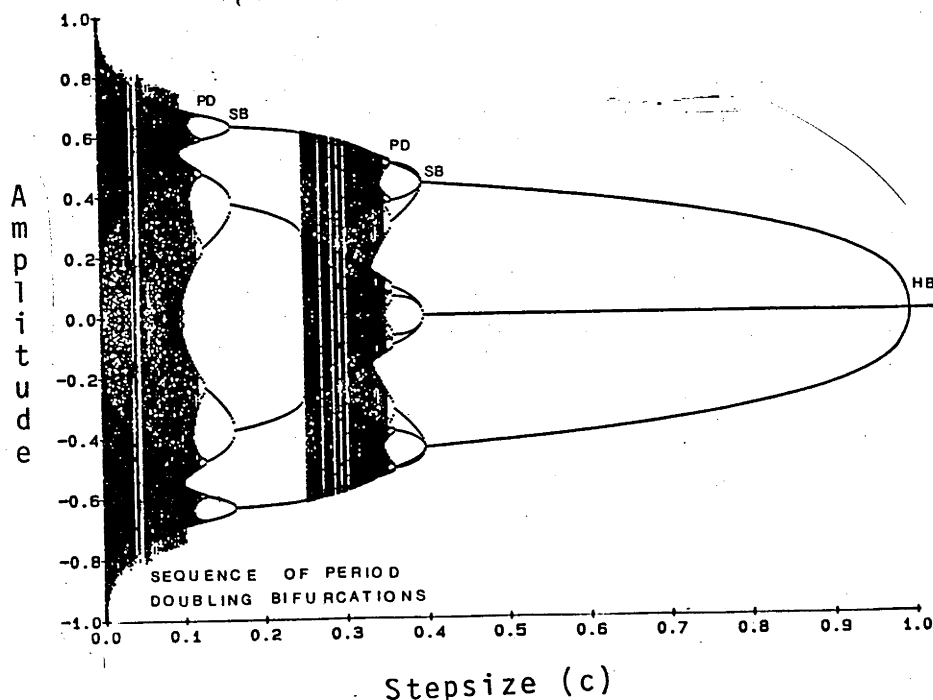
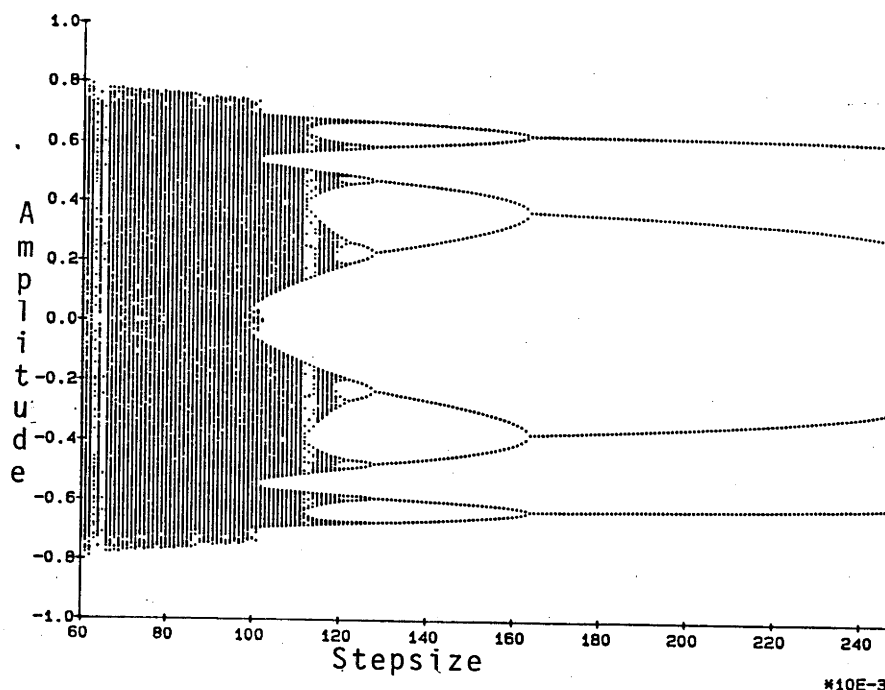


Figure 4.6.2 Bifurcation Diagram

#### 4.6.4 "Cycle Slipping"

Analysing the time sequence  $\{v_k, k \in \mathbb{N}\}$  (as defined by 4.6.1) one observes that it consists of certain segments of apparent almost periodic behaviour, separated by a short transition characterized by large deviations. (See Figure 4.3, which displays a sample of the time sequence of  $\{v_k, k \in \mathbb{N}\}$ .)

Assuming that all periodic orbits are either of saddle type (or completely unstable), of which we are strongly convinced in the light of the previous observations, this phenomenon becomes easy to understand. Orbits "close" to a stable manifold belonging to a certain periodic orbit approach this periodic orbit, hence approach the unstable manifold of this periodic orbit and are consequently repelled away from it, to be captured by a stable manifold belonging to another periodic orbit. This whole cycle keeps repeating itself.

This intuitive picture, "cycle slipping", gives us the impression that the union of all the unstable manifolds of the periodic orbits of saddle type might be a one dimensional strange attractor.

Remark:

(R.4.18) In principle it would be possible that there are strictly stable periodic

orbits, with such a small basin of attraction that they are unobservable from a computer simulation. However, we have never established such orbits, and also the period doubling sequence seems to exclude this possibility - although, this is not analytically established. In any practical situation, the effect such orbits could have on the control objective is nil.  $\square$

#### 4.6.5 The "Gumleaf Attractor"

So far we have established chaos (in the sense of Definition 4.1) using the Smale-Birkhoff Homoclinic Theorem. This Theorem leads to chaos of the horseshoe type. The invariant set established by this Theorem is typically a Cantor set  $\times$  Cantor set; which need not be attractive. Its domain of attraction (due to the hyperbolicity of the unstable periodic orbits contained in the invariant set) is typically a Cantor set  $\times$  curve, and therefore this type of chaos is called transient - it essentially dies out. In order to address the stability question for the underlying adaptive problem we need to characterize the asymptotic dynamics, because it is precisely the asymptotic behaviour (of a generic orbit) which determines the stability properties of the adaptive control problem. What kind of dynamics govern the asymptotic behaviour of (4.6.1) or (4.6.4)? In view of the above facts, especially that all periodic orbits are unstable, we are convinced that chaos is generically present and governs both the transient and the asymptotic dynamics. (In the absence of any stable (attracting) periodic orbit, it is quite obvious that almost all trajectories are doomed to wander aperiodically around in the state space for ever.)

The Figure 4.4 represents the orbit for a "typical" initial condition in the phase plane. It consists of 40 000 iterations. Comparing the "curves" traced by this orbit with the unstable manifold of the periodic orbit of period four, we find that they are virtually identical. Therefore combining, all previous observations, we conjecture:

#### Conjecture 4.1:

The union of all unstable manifolds of the periodic orbits of saddle type form a one dimensional hyperbolic strange attractor, denoted by S.  $\square$

#### Discussion:

It is clear that the union of all the unstable manifolds of the periodic



orbits of saddle type is a one dimensional set, which is attracting. Moreover, if it is an attractor it is a strange attractor because it contains several transversal homoclinic orbits - and the corresponding horseshoes. The conjecture would be proven if all periodic orbits were of saddle type having transversally intersecting stable and unstable manifolds. The closure of the stable manifolds of the periodic orbits would then form the basin of attraction for this attractor, as well as the foliation of stable manifolds of this attractor. All numerical, and previous analytical evidence points this way. (An important fact to note is that it appears as if the unstable manifold of the periodic orbit of period four is dense in this attractor. This is a property of an hyperbolic, connected attractor.) It even appears as if the whole of  $D$ , the open manifold on which  $G$  is defined, forms the basin of attraction of this attractor.  $\square$

Asymptotically the trajectories are in the strange attractor. It is therefore necessary to be able to describe the behaviour on this invariant set in some manner in order to discuss the stability of the underlying adaptive control problem. (We remind the reader that the plant output is related to the product of the  $v_k$ 's, via (4.3.9) and (4.3.12)). The "averaged behaviour" on the invariant set can be characterized using concepts familiar from stochastic process theory.

A Theorem of Sinai-Bowen-Ruelle [1,6,7] states that for a diffeomorphism  $F$  defined on a compact manifold possessing a hyperbolic strange attractor  $S$  there exists a measure  $\mu$ , invariant with respect to the diffeomorphism, supported on the hyperbolic attractor, such that for all initial conditions  $x$  in the basin of attraction of the attractor, and for all real valued continuous functions  $g$  the césaro-mean evaluated along the orbit generated by  $x$  exists and converges to the ensemble average of  $g$  over the attractor:

$$\lim_{N \uparrow \infty} \frac{1}{N} \sum_{k=1}^N g(F^k(x)) = \int_S g d\mu \quad (4.6.15)$$

In other terms  $F$  is ergodic. We conjecture that this result also holds in the present situation.

#### Conjecture 4.2:

$G$  is ergodic.  $\square$

The technical difficulties - apart from those encountered for the previous

conjecture, stem from the fact that it is not clear whether  $G$  can be extended to a diffeomorphism  $G^1$  defined on a compact manifold  $D^1$  containing  $D$  as a sub-manifold. If this were possible, then this conjecture follows from the previous one by the quoted theorem of Sinai, Bowen and Ruelle. A direct proof, using the definition of  $G$ , could be envisaged, but was beyond our capabilities (so far it escapes us).

Notice that the "cycle slipping" idea, strongly suggests that the césaro-mean of continuous functions evaluated along typical orbits should exist and be independent of the particular orbit. Indeed trajectories spend most of their time in the neighbourhood of periodic orbits. Clearly, on the periodic orbits the césaro-mean is well defined. The invariant measure would then simply attribute different weights to different periodic orbits according to the relative amount of time spent by a typical orbit in their neighbourhood.

#### 4.6.6 Implications for the closed loop adaptive system

Firstly, because of (4.3.10), (4.3.11) it is clear that the parameter estimate feedback gain behaves chaotically. What does a chaotic feedback gain imply for the stability of the closed loop?

Recall that

$$y_k = \sqrt{b} v_k y_{k-1} \quad (4.6.16)$$

(see equations (4.3.9) and (4.3.10)). Hence, we are interested in the products of "chaotic" signals:

$$y_k = \left( \prod_{l=1}^k v_l \right) (\sqrt{b})^k y_0 \quad (4.6.17)$$

Using the ergodicity property, we can immediately investigate the césaro-mean of the logarithm of the absolute values of the  $v_k$ 's; evaluating this numerically we obtain:

$$\lim_{N \uparrow \infty} \frac{1}{N} \sum_{k=1}^N \log |v_k| \approx \frac{1}{2} \log 2 \quad (4.6.18)$$

#### Remark:

(R.4.19) We have exhaustively evaluated (4.6.18) for different initial conditions and for different sample sizes. For sufficiently large  $N$ ,  $N > 10,000$  appears

adequate, the same value  $(\log 2)/2$  was obtained in a "statistically" consistent way. This observation strongly supports the ergodicity property.  $\square$

(R.4.20) The result (4.6.18) yields yet another confirmation of the fact that the dynamics of the difference equation (4.6.1) are hyperbolic. Notice that

$$DG^N(x) = \prod_{k=1}^N DG(x_k)$$

where  $x_1 = x$ ,  $x_k = G^{k-1}(x)$ ,  $k = 1, \dots, N$ . By definition of the state and the state transition map (cf. (4.6.2) and (4.6.3)), we have that

$$\det DG(x_k) = (1/v_{k-1})^2$$

with

$$x_k = (v_k \quad v_{k-1})^T \quad (4.6.2)$$

Hence, from (4.6.18) we obtain that

$$\lim_{N \uparrow \infty} \det (DG^N(x))^{2N} = 1.$$

This demonstrates that the map  $G$  is on average area contracting, and therefore the asymptotic dynamics are restricted to, at most, a one dimensional set. It follows also that almost any orbit is either locally completely stable or is of saddle type. (Cf. [8] for more details about the implications of this type of result for the map  $G$ .)  $\square$

Consequently, from (4.6.18)

$$\lim_{N \uparrow \infty} \frac{1}{N} \log |y_N| \approx \frac{1}{2} \log 2b \quad (4.6.19)$$

We conclude therefore that  $y_k$  converges exponentially to zero for all  $b : 0 < b < \frac{1}{2}$ , in the sense that, for (Lebesgue) almost all initial conditions, there exists a constant  $\alpha > 1$ , independent of the initial conditions such that:

$$\alpha^k |y_k| \rightarrow 0 \quad \text{as } k \uparrow \infty \quad (4.6.20)$$

and diverges exponentially for all  $b > \frac{1}{2}$ , in the sense that there exists a constant  $\beta < 1$ , such that for almost all initial conditions:

$$\beta^k |y_k| \rightarrow +\infty \quad \text{as } k \uparrow \infty$$

$\alpha, \beta$  only depend on  $b$ ;

$$\begin{aligned} 1 < \alpha < \frac{1}{(2b)^{\frac{1}{2}}}; & \quad 0 < b < \frac{1}{2} \\ \frac{1}{(2b)^{\frac{1}{2}}} < \beta < 1; & \quad \frac{1}{2} < b \end{aligned} \quad (4.6.21)$$

Remarks:

(R.4.21) The present definition of exponential convergence is different from the classical definition, however it is frequently used in the context of stochastic processes [9]. Notice in particular that (4.6.20) does not imply the existence of a (uniform in the initial conditions) exponentially decaying overbound for  $|y_k|$ .  $\square$

(R.4.22) The adaptive feedback gain  $\hat{a}_k$  given by

$$\hat{a}_k = a + \frac{(b)^{\frac{1}{2}}}{v_k} \quad (4.6.22)$$

(see equations (4.3.10) and (4.3.12)) behaves in a chaotic fashion because  $v_k$  does, hence does not yield directly any information about the system's parameters. However, observing that:

$$\lim_{N \uparrow \infty} \frac{1}{N} \sum_{k=1}^N \log\left(\frac{|\hat{a}_k - a|}{(b/2)^{\frac{1}{2}}}\right) \cong 0 \quad (4.6.23)$$

(which follows from (4.6.21) and (4.6.18)) one could envisage a statistical test to extract both plant parameters  $a$  and  $b$  from the observed  $v_k$  or  $\hat{a}_k$  sequences. Using the assumption that  $G$  is ergodic, this would lead to estimates for  $a$  and  $b$  as obtained in the periodic case (cf. (4.5.22), (4.5.23)).  $\square$

#### 4.7 The Adaptive Control Problem: Discussion

In this section we return to the adaptive closed loop, described in Section 4.3 and state the main robustness results, summarising the previous sections 4.4, 4.5 and 4.6. The obtained results are discussed and interpreted in the light of the available theories for establishing robustness of adaptive schemes. In particular we argue that chaotic parameter estimates are not necessarily a bad thing to have. We illustrate our discussion at the end of this section with some

simulation examples.

We collect first our main observations about the dynamics of this particular adaptive control problem in the following two Corollaries. The first result describes the behaviour of the parameter estimate:

Corollary 4.1:

For all initial conditions, except possibly for a set of Lebesgue measure zero, the parameter estimate  $\hat{a}_k$ ,  $k \in \mathbb{N}$ , defined in (4.3.3)-(4.3.10) has the properties:

- (i) if  $b < 0$ , the parameter estimate is bounded, and exponentially becomes periodic with period 2:

$$\hat{a}_k \rightarrow a + (\pm)(-1)^k \left(\frac{|b|}{2}\right)^{\frac{1}{2}} \quad \text{as } k \rightarrow \infty \quad (4.7.1)$$

- (ii) if  $b=0$ , then  $\hat{a}_k \equiv a \quad \forall k \geq 2$

- (iii) (subject to the veracity of the Conjectures 4.1 and 4.2): if  $b > 0$ ,  $\hat{a}_k$  behaves chaotically, and along almost all trajectories the césaro-mean of the logarithm of  $\tilde{a}_k = a - \hat{a}_k$  is defined and is given by:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{l=1}^N \log |\tilde{a}_l| \equiv \frac{1}{2} \log \frac{|b|}{2}, \quad b \neq 0 \quad (4.7.2) \quad \square$$

Notice that these results are valid independent of the size of  $b$ . Observe also that in the case that  $\tilde{a}_k$  behaves chaotically, one does not have a boundedness result, on the contrary, almost certainly  $\tilde{a}_k$  will exceed any given bound (!), however, as we shall argue, this is not dramatic, not even bad since the control signal and the plant output are bounded.

The robustness result, which gives a sharp stability-instability boundary in the  $(a,b)$  parameter plane is:

Corollary 4.2:

For all initial conditions, except for a set of Lebesgue measure zero, and conditioned on the veracity of the Conjectures 4.1 and 4.2, the plant output  $y_k$  defined in (4.3.1)-(4.3.3) or (4.3.4) is:

- (i) bounded and regulated to zero if  $|b| < \frac{1}{2}$   
(ii) the rate of convergence is exponential, in the sense that

$$\text{for } b < 0 \quad |y_k| < C(2|b|)^{k/2} \quad (\text{some } C > 0)$$

$$\text{for } b > 0 \quad \frac{|y_k|}{(2b)^{k/2}} \rightarrow C \text{ as } k \uparrow \infty \quad (\text{some } C > 0)$$

$$\text{for } b = 0 \quad y_k \equiv 0 \quad \forall k \geq 2$$

(iii) unbounded and diverges exponentially if  $|b| > \frac{1}{2}$ .  $\square$

Remarks:

(R.4.23) Observe that this algorithm achieves exponentially fast regulation (in an appropriate sense) of the output in the presence of model-errors, without using - obviously - an external input. This demonstrates that there are alternatives to obtain robust adaptive schemes, other than forcing exponentially fast identification by the use of an external sufficiently exciting input.  $\square$

(R.4.24) Here we have another example of an adaptive scheme which regulates the plant output to zero - i.e. achieves its desired purpose - whilst the parameter error does not converge. Actually, because of the undermodelling it would be rather surprising if the parameter error did converge to zero. Notice however that this "non-convergence" limits the robustness margin of the adaptive loop. Indeed, if it were possible to estimate  $a$  correctly, then the proposed controller would regulate the output to zero for all  $|b| < 1$ .  $\square$

(R.4.25) Notice that the model by no means needs to be a good approximation for the real plant. Even in the situation  $b \neq 0$ ,  $a = 0$ , where the first order model does not make any sense at all, good control action is obtained. This is clearly an advantage of the fast adaptive loop, which slow adaptation never can obtain.  $\square$

(R.4.26) Averaging techniques in adaptive control are able to handle slow adaptation and fast parasitic plant states for model errors. Simplified, the theory states that as long as the adaptive algorithm together with the dominant slow part of the plant is slow compared to the parasitic states - even after closing the adaptive loop, all is well provided that the parameter estimates in the adaptive loop are close to the real parameters. The theory is conditioned on the assumption that the controller indeed can stabilize the closed loop system (including the parasitic part) for parameters belonging to some set containing the real parameter (cf.[10]). These ideas don't work in this environment - as the adaptive observer and control law have deadbeat response in the ideal circumstances. But, averaging can still teach us something. Indeed, assuming that the time-average of the parameter estimate  $\hat{a}_k$  exists i.e.

$$\langle \hat{a}_k \rangle = \lim_{N \uparrow \infty} \frac{1}{N} \sum_{k=\ell}^{N+\ell-1} \hat{a}_k = \bar{a}, \text{ uniformly in } \ell \quad (4.7.3)$$

it is possible to obtain an "averaged" equation which has the same stability properties as the original system:

$$\bar{y}_k = \langle \bar{a}_k \rangle \bar{y}_{k-1} + b \bar{y}_{k-2} \quad (4.7.4)$$

As indicated in section 6, the assumption "the time average exists" is a non trivial one, which we have been able to remove in our analysis.

In order to be able to answer the stability question of the non-linear loop through analysis of (4.7.4) it is necessary to require that  $b$  is small. It is possible to obtain estimates for how small  $b$  has to be, typically  $|b| \ll 1$ , which has to be compared with the correct robustness margin of  $|b| < \frac{1}{2}$ . The philosophy behind this kind of averaging is precisely the opposite of the one discussed in Chapters 2 and 3, (cf. also [10,11,12,13]); it is not the state of the fast part of the plant which is averaged out, but the fast adaptive estimate which is averaged by the plant! Why these results are not available in the literature is precisely due to the difficulties encountered in establishing the existence of time averages for the parameter estimate - caused by the complex dynamics of these estimates.  $\square$

(R.4.27) It is impossible by using a constant output or state feedback to stabilize the class of "uncertain" systems:

$$y_k = ay_{k-1} + by_{k-2} + u_{k-1}$$

with constant parameters  $a$  and  $b$ , which are unknown, but satisfy the following bounds:

$$|a - a_1| < a_0, \quad a_0 > \frac{1}{2}, \quad a_1 \text{ arbitrary}$$

$$|b| < \frac{1}{2}$$

( $a_0, a_1$  are given constants). Notice that for a control input defined by

$$u_{k-1} = f_1 y_{k-1} + f_2 y_{k-2}$$

where  $f_1, f_2$  are the gains of the controller, there is not a single gain setting which can stabilize all systems in the above class. This demonstrates that the adaptive controller is, in a sense, superior to a more complicated controller based on fixed gain design for uncertain systems.  $\square$

(R.4.28) For  $b < 0$ , the periodic feedback gain situation, the stability is in the sense of Lyapunov, however for  $b > 0$ , chaotic feedback gain, this is not the case! The present technique (via ergodicity and césaro-mean convergence) is closely related to techniques used to establish asymptotic properties of products of sequences of (ergodic) stochastic matrices [8,14]. Nothing too surprising, as deterministic chaos seems to have a lot in common with random processes [1,6,7,8].  $\square$

(R.4.29) The problem of a possible division by zero in the parameter estimator does not cause any great difficulty. The only trace of this problem in the Theorem statements is the qualifier "except for a set of Lebesgue measure zero". Hence our analysis captures the generic properties of the adaptive scheme. Note that this is common practice in the discussion of the adaptive system in a stochastic context, where results are only almost surely (at their best) valid. See for example [3,15] - and the references mentioned therein where the event of a division by zero is treated and disposed of by noting that it is an event of zero probability - much in the same style as our analysis. Moreover, the qualifier 'for Lebesgue almost all initial conditions' appears to be a product of our methods. Our analysis does not deal with those initial conditions for which a division-by-zero events occurs. From a further analysis we suspect that the control algorithm as introduced in section 4.3 regulates the output to zero for all initial conditions provided that  $|b| < 0.5$ .  $\square$

(R.4.30) In both cases  $b > 0$  and  $b < 0$ , the adaptive response is extremely sensitive to small changes in the initial conditions (at least for part of  $\mathbb{R}^3$ ). This fact is accentuated here due to the discontinuity in the equation (4.3.8) describing the evolution of the ratio of two successive outputs. In the chaotic case  $b > 0$ , this is obvious and the sensitivity is uniform over the state space, i.e. for any initial condition  $(y_0, y_{-1}, y_{-2})$  there is an initial condition arbitrarily close which has a completely different (transient) response. Also in the periodic case this phenomenon can be observed, but only for initial conditions  $(y_0, y_{-1}, y_{-2})$  which yield a trajectory  $\{y_k, k \in \mathbb{N}\}$  which terminates, i.e.  $y_k = 0$  for some  $k$ . (Cf. the discussion of the transient behaviour of trajectories  $\{w_k, k \in \mathbb{N}\}$  in Lemma 4.3.) This implies that the transient analysis is extremely difficult in the presence of undermodelling errors, and although effectively completed in our analysis, there seems little hope that we will be able to repeat this analysis in more general situations.  $\square$



Some Examples:

Figures 4.7, 4.8 and 4.9 illustrate the adaptive loop's response.

Figure 4.7 Periodic Stabilization  $a = 2$ ,  $b = -0.3333$

Figure 4.7.1 Output and Control Input

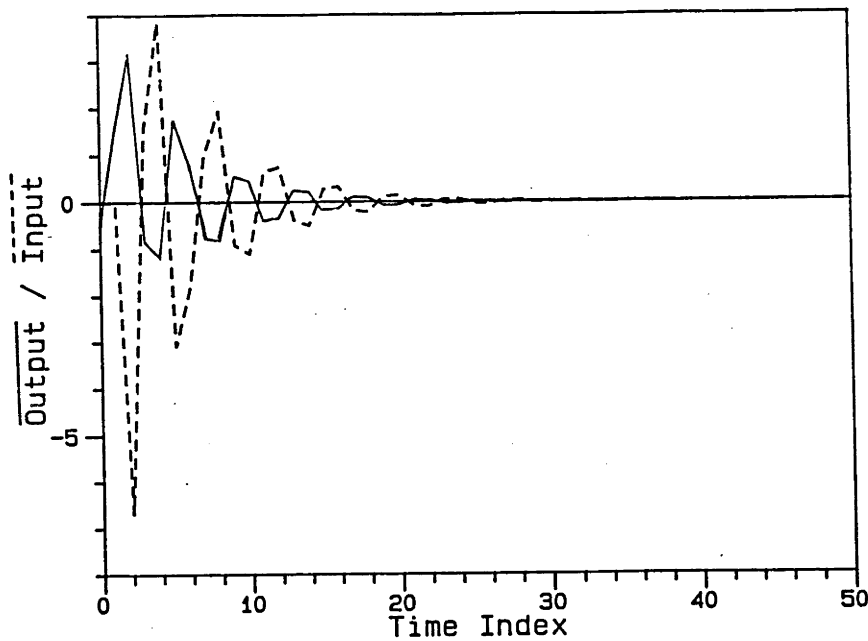


Figure 4.7.2 Feedback Gain/Parameter Estimate

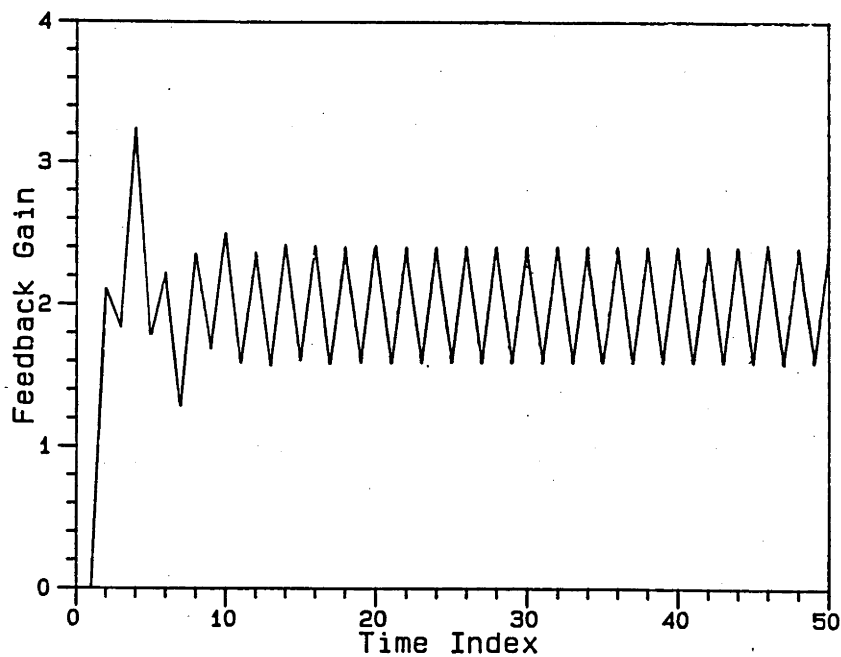


Figure 4.8 Chaotic Stabilization  $a = 2$ ,  $b = 0.3333$

Figure 4.8.1 Output and Control Input

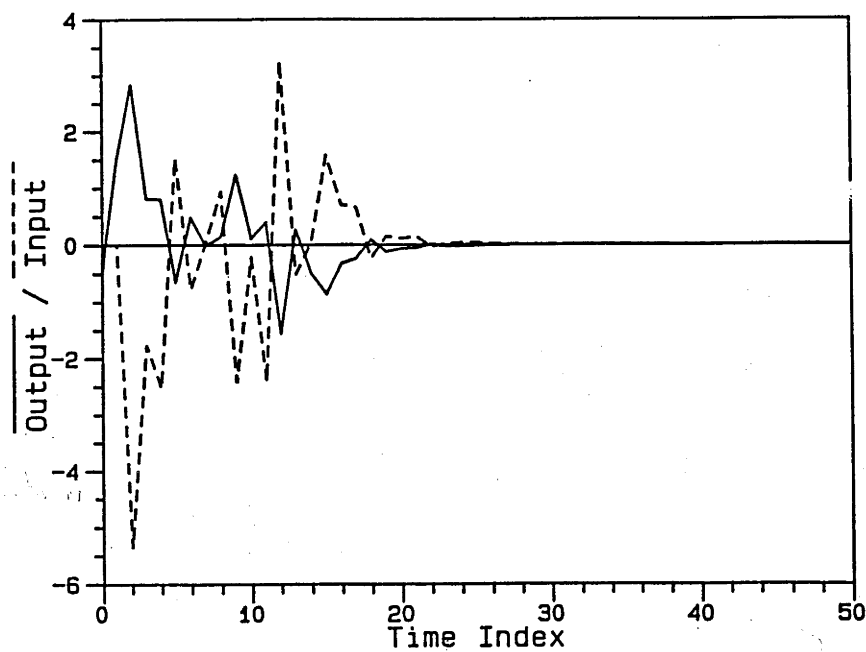


Figure 4.8.2 Feedback Gain/Parameter Estimate

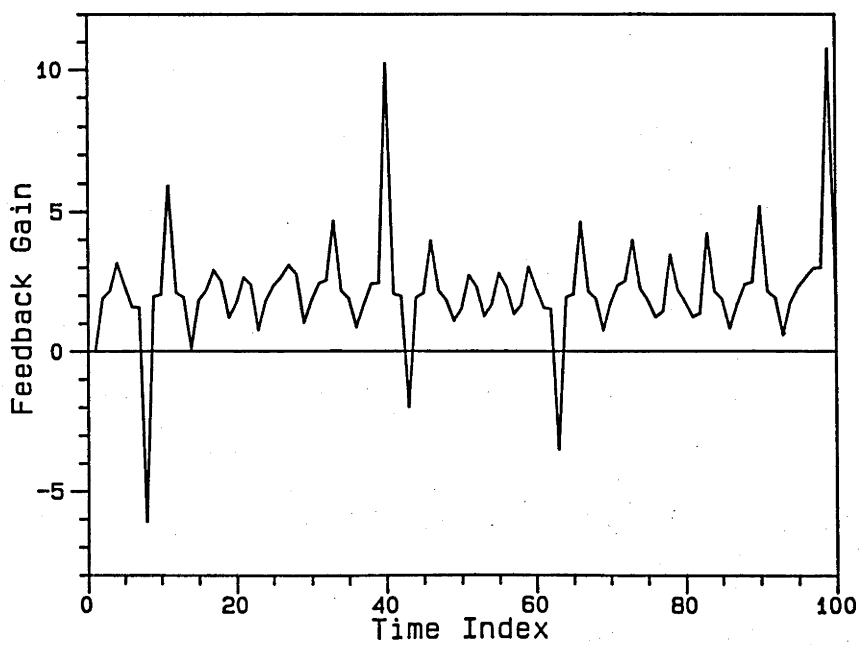


Figure 4.9 Unstable Response  $a = 0$ ,  $b = -0.8$

Figure 4.9.1 Output and Control Input

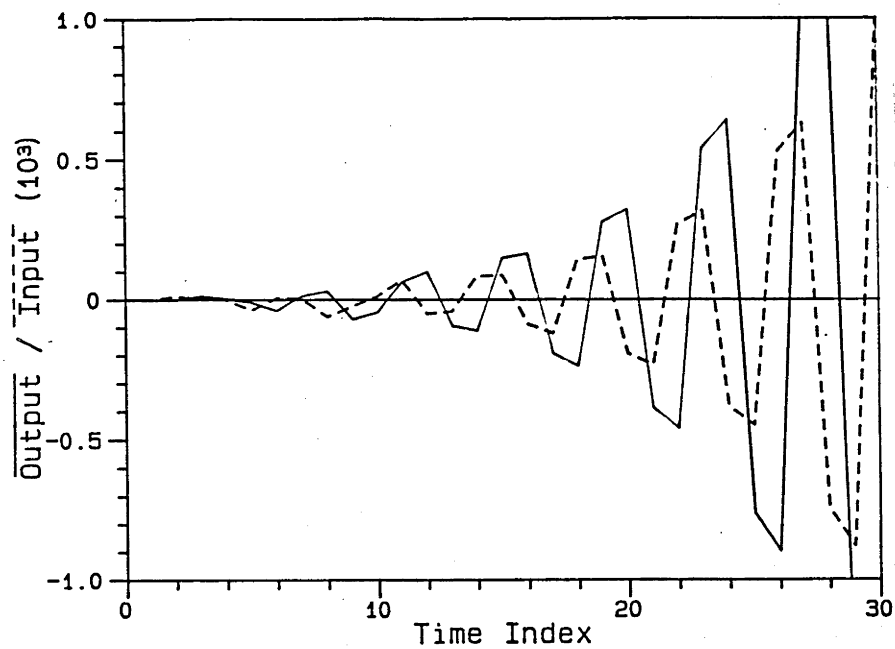
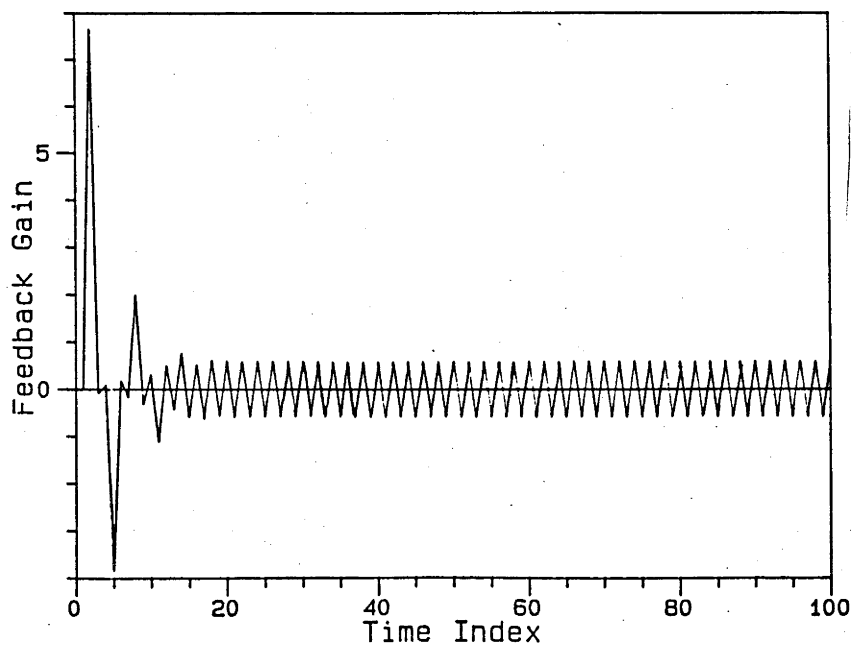


Figure 4.9.2 Feedback Gain/Parameter Estimate



Figures 4.7.1 and 4.7.2 contain respectively the trajectories of the plant output  $y_k$ , the control action  $u_k = -\hat{a}_k y_k$ , and the parameter estimate/feedback gain  $\hat{a}_k$  for the open loop unstable plant with parameters  $a = 2$ ,  $b = -.3333$ . The plant output is quickly regulated to zero, whilst the controller gain behaves asymptotically periodically. Figures 4.8.1 and 4.8.2 display respectively the same trajectories for the open loop unstable plant with parameters  $a = 2$  and  $b = +.3333$ . Again the output is regulated to zero, and also the control action disappears quickly, but the controller gain behaves quite erratically. (For both simulations the initial conditions were  $y_{-2} = 1.5$ ,  $y_{-1} = -.5$ ,  $\hat{a}_0 = 0$ .) Figure 4.9 illustrates an unstable response. The plant parameters are  $a = 0$ ,  $b = -0.8$ , which characterizes an open loop stable plant. The plant parameter estimate/feedback gain is periodic (Figure 4.9.2) but input and output diverge exponentially. Notice the different time scales on the horizontal axes in Figures 4.8 and 4.9. The feedback gain/parameter estimate is displayed over a longer period of time than plant output and input are.

#### 4.8 Complements

Before presenting the final conclusions, we briefly point out some further features without going into details, mainly referring to simulation experience.

The comments and claims of this section follow exhaustive simulation experiments. Given the simplicity of the problem formulation (discrete time, low dimension, rational calculations) and the inherent scepticism of the readers of the adaptive control literature, these statements are more convincingly checked by the readers themselves. Because, the modifications we discuss in this section lead to much more complicated analyses without promising more insight, we feel justified in not exploring these avenues too closely.

In the simulations the influences of the various disturbances are easily identified in the "periodic case  $b < 0$ " whilst in the "chaotic case  $b > 0$ ", the chaos tends to obscure any regularity. In this situation the effects are best identified by altering the initial conditions slightly and comparing the different trajectories.

##### 4.8.1 Higher order problem, time-varying parameters

The observed dynamical features are not miraculously due to the

combination of a second order plant with a first order model e.g. if the plant is of third order:

$$y_k = ay_{k-1} + by_{k-2} + cy_{k-3} + u_{k-1} \quad (4.8.1)$$

we obtain qualitatively the same results. Moreover, it appears as if the size of  $c$  is immaterial. (It is not difficult to demonstrate that for  $-\frac{1}{2} < b < 0$ , the periodic orbit of period two for the feedback gain, indeed is independent of  $c$ , and is asymptotically stabilizing!)

The adaptive control scheme, because of its ultra fast identification law, promises good tracking properties in the case the plant is time-varying. Supposing, that the plant can be represented by:

The plant:

$$y_k = a_{k-1}y_{k-1} + by_{k-2} + u_{k-1} \quad (4.8.2)$$

and using the same adaptive scheme as outlined in section 2, we obtain

The closed loop:

$$r_k = b \left[ \frac{1}{r_{k-1}} - \frac{1}{r_{k-2}} \right] + (a_{k-1} - a_{k-2}) \quad (4.8.3)$$

$$y_k = r_k y_{k-1} \quad (4.8.4)$$

$$\hat{a}_k = a_{k-1} + \frac{b}{r_{k-1}} \quad (4.8.5)$$

From these equations and from the previous discussion of the homogeneous part of equation (4.8.3) it transpires that small time-variations pose no threat to this controller. Simulations with  $a_k$  being periodic, or stationary random or even being a random walk process or a ramp function can be conducted without experiencing difficulties.

#### 4.8.2 Effects of additive noise, rounding errors and clipping

The proposed adaptive scheme is sensitive to measurement errors, though not as sensitive as one would suspect. Assume that the plant can be represented by:

$$y_k = ay_{k-1} + by_{k-2} + u_{k-1} + v_k \quad (4.8.6)$$

where  $v_k$  is additive measurement noise. In this situation the parameter

estimate is governed by:

$$\hat{a}_k = a + b \frac{y_{k-2}}{y_{k-1}} + \frac{v_k}{y_{k-1}} \quad (4.8.7)$$

If the signal to noise ratio is negligible,  $\hat{a}_k$  jumps crazily, driving  $y_k$  to large values, but therefore restoring the signal to noise ratio to acceptable levels, and hence bringing the adaptive scheme back to good behaviour, driving the output to zero and so on. The resulting "cycle" gives acceptable control performance.

In the case  $b > 0$ , the feedback parameter is chaotic and occasionally the feedback parameter can take on astronomical values, as can the control input; the more so if  $b$  is closer to  $\frac{1}{2}$ , because then the stabilization requires more time and those "rare" events become more visible. Therefore it is natural to introduce clipping in the parameter estimator to limit its possible range. Done with care and conditioned on the "return" strategy this does not alter the nature of the dynamics of the closed loop. More precisely, the true parameter should be within the allowed parameter range, and this range should be large enough to accommodate for (most) of the periodic points of the adaptive algorithm. If these obvious criteria, which are easy to meet, are indeed met, clipping has a beneficial effect on the dynamics of the algorithm, certainly for values of  $|b|$  close to  $1/2$ , without upsetting the earlier picture of the dynamics.

Multiplicative noise does not cause any difficulties, this is a consequence of the adaptive system's good tracking properties. Assume that the plant is of the form

$$y_k = (a + \gamma_k)y_{k-1} + (b + \mu_k)y_{k-2} + (1 + \rho_k)u_{k-1} \quad (4.8.8)$$

where  $\gamma_k$ ,  $\mu_k$  and  $\rho_k$  are disturbances, (e.g. due to rounding errors). For small  $\gamma$ ,  $\mu$ ,  $\rho$  all of the above conclusions remain qualitatively valid. The presence of  $\mu$  tends to decrease the robustness margin. The noise source  $\gamma$  makes the adaptive deadbeat control exponentially decaying instead of deadbeat. The noise source  $\rho$  introduces a small offset ( $\hat{a} \approx (a/(1+\rho f))$ ) (but such as to effect the deadbeat control law) and slows down the estimation speed of the algorithm, instead of having deadbeat response it converges exponentially.

#### 4.8.3 Transient response of more standard algorithms

An important facet of our analysis is that it appears to capture the transient behaviour of the adaptive scheme with a finite step size estimator ( $c > 0$ )

in 2.5):

$$\hat{a}_k = \hat{a}_{k-1} + \frac{y_{k-1}}{c+y_{k-1}^2} (y_k - \hat{a}_{k-1}y_{k-1} - u_{k-1}) \quad (4.8.9)$$

For  $c$  small compared to  $y_{k-1}^2$ , our analysis is valid. In this "transient region" of the state space, all the above discussed dynamical properties are present. In particular, for  $|b| < 1/2$ , our theory predicts convergence of  $y_k$  to zero, hence, after a transient period,  $y_{k-1}^2$  becomes of the order of magnitude of  $c$  and then the dynamics become essentially linear, driving  $y_k$  further towards zero and  $\hat{a}_k$  towards some constant, depending on  $b$ . (Note that for the specified range of  $b$  the linearized equation predicts asymptotic stability.) On the other hand if the initial conditions (for  $y$ ) were large (compared to  $c$ ) and the parameter  $b$  is larger than  $1/2$  in modulus the algorithm behaves unstably, despite the fact that the linearized system could well be stable. Figure 4.10 illustrates this point. The initial condition were  $y_0=10$ ,  $y_{-1}=-0.34$  and  $\hat{a}_0=0$ , the stepsize was chosen as  $c=0.0001$ , the other parameters were set as  $a = 5$ ,  $b = -0.4$  for Figure 4.10.1 and  $c=0.001$ ,  $a = 0$ ,  $b = -0.6$  for Figure 4.10.2. One can clearly recognize the "transient periodicity" in the parameter estimate. As chaos obscures any regularity we do not present any figures for  $b>0$ , but the readers may convince themselves that the transient behaviour does contain the features of "chaos" - which become especially clear when one alters the initial conditions slightly and tries to compare trajectories! It is possible to trade smaller  $1/c$  against larger initial conditions.

This observation indicates that our analysis describes possible transient phenomena for the more standard algorithm (with the estimator (4.8.9)) in the presence of this particular type of undermodelling. For  $c > 0$ , we have locally exponential stability if  $|b| < 1$ , but the large scale behaviour is unstable for  $|b| > \frac{1}{2}$  and exponentially stable for  $|b| < \frac{1}{2}$ .

Figure 4.10 Finite Stepsize Adaptive Algorithm

Figure 4.10.1 Stable Response  $|b| < 0.5$  ( $a = 5, b = -0.4, c = 0.0001$ )

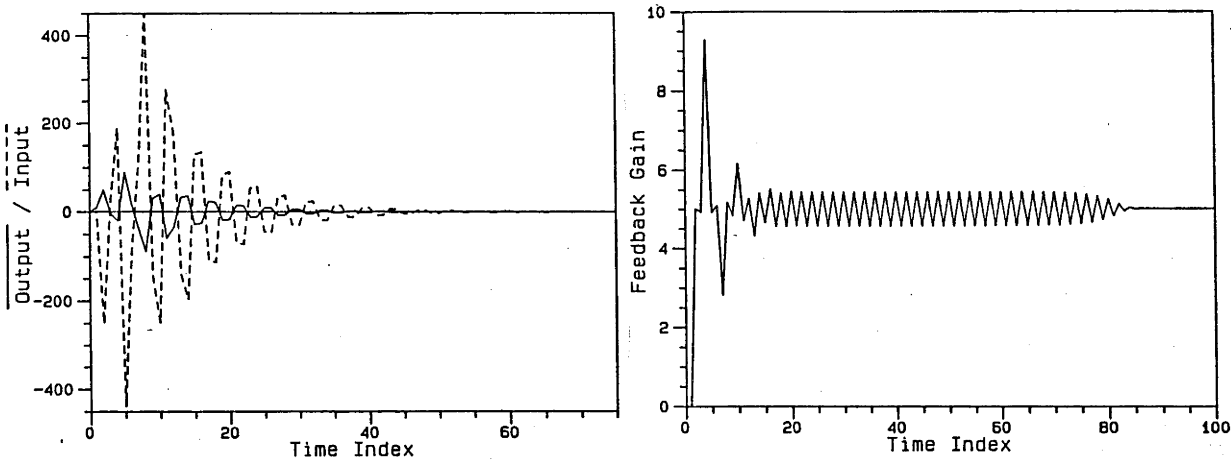
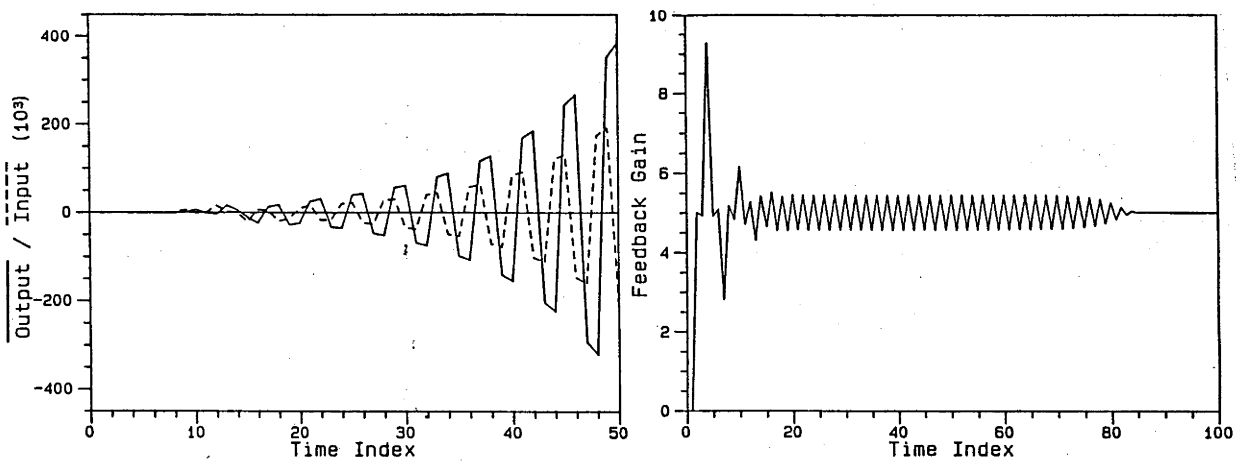


Figure 4.10.2 Unstable Response  $|b| > 0.5$  ( $a = 0, b = -0.6, c = 0.001$ )





#### 4.9 Conclusions

We have presented a detailed analysis for a particular adaptive control problem in the presence of a restricted form of undermodelling. The minimum variance controller for a first order system connected to a variant of the least mean square estimation algorithm on the basis of the certainty equivalence principle can adaptively stabilize any second order system with transfer function  $z^{-1}/(1 - az^{-1} - bz^{-2})$  with  $|b| < \frac{1}{2}$ . The example highlights the complexity of the nonlinear behaviour in both transient and asymptotic response.

In particular we point out that whenever chaotic dynamics are present in the transient response (as here is the case for  $b > 0$ ), a Lyapunov function approach to the stability/convergence question is bound to fail, as along the solutions the difference  $\Delta V/V$  itself behaves chaotically. It is part of the folklore of adaptive control to design and redesign [16,17] adaptive algorithms on the basis of a Lyapunov function, both in the exact modelling situation or in the presence of bounded disturbances. In the light of the present case study it is clear that such an approach, although helpful to obtain a reasonable algorithm, may break down in the presence of undermodelling.

Although we do neither promote nor advocate the present ultra-fast adaptive controller, we do stress that the algorithm has very good robustness properties, both with respect to undermodelling and to time variations in the plant parameters (and multiplicative noise) without losing its control objective. Hence it deserves further analysis, not in the least in the direction of extending the present results to higher order models and controllers. Although this promises to be a particularly hard problem, we do believe it is possible.

"It is hard to adapt to chaos, but it can be done.

I am living proof of that: It can be done."

Kurt Vonnegut Jr,

Breakfast of Champions

(Ch19, p210)

#### 4.10 References

- [1] J. Guckenheimer and P. Holmes, Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields, Springer Verlag, NY, 1983.
- [2] I.M.Y. Mareels, R.R. Bitmead, 'Nonlinear dynamics in adaptive control: chaotic and periodic stabilization', to appear Automatica, 1986.
- [3] G.C. Goodwin and K.S. Sin, Adaptive Filtering, Prediction and Control, Prentice Hall, Englewood Cliffs, NJ, 1984.
- [4] P.E. Kloeden, 'Chaotic difference equations in  $\mathbb{R}^n$ ', J. Austral. Math. Soc., XXXIA, pp. 217-225, 1981.
- [5] S. Smale, 'Differential dynamical systems', Bulletin of the American Mathematical Society, 1967, Vol 73, pp741-817.
- [6] R. Bowen and D. Ruelle, 'The ergodic theory of axiom A flows', Invent. Math., Vol.79, pp. 181-202, 1975.
- [7] D. Ruelle, 'Small random perturbations of dynamical systems and the definition of attractors', Comm. Math. Phys., pp. 137-151, 1982.
- [8] V.I. Oseledec, 'A multiplicative ergodic theorem', Trans. Moscow Math. Soc., Vol.19, pp. 197-231, 1968.
- [9] R.R. Bitmead, 'Persistency of excitation conditions and the convergence of adaptive schemes', IEEE Trans. Information Theory, IT-30, No.1, pp. 183-191, 1984.
- [10] B.D.O. Anderson, R.R. Bitmead, C.R. Johnson Jr, P.V. Kokotovic, R.L. Kosut, I.M.Y. Mareels, L. Praly and B.D. Reidle, Stability of Adaptive Systems: Passivity and Averaging Analysis, M.I.T. Press, Boston, MA, 1986.
- [11] B.D. Riedle, P.V. Kokotovic, 'Integral manifold of slow adaptation', IEEE Trans. Auto. Contr., AC-31, No.4, pp. 316-324, April 1986.
- [12] B.D. Riedle, B. Cyr and P.V. Kokotovic, 'Disturbance instabilities in an adaptive system', IEEE Trans. Auto. Control, AC-29, No.9, pp.822-824, Sep. 1984.

- [13] M. Bodson, S. Sastry, B.D.O. Anderson, I.M.Y. Mareels, R.R. Bitmead, 'Nonlinear averaging theorems, and the determination of parameter convergence rates in adaptive control', Systems and Control Letters, Vol. 7, No.3, pp. 145-159, June 1986.
- [14] R.R. Bitmead and R.K. Boel, 'On stochastic convergence of infinite matrix products and its application to adaptive estimation', Proc. 7th IFAC Symposium on Parameter Estimation and System Identification, York, UK, 1985.
- [15] P.E. Caines and S.P. Meyn, 'The zero division problem of stochastic adaptive control', Proc. 7th IFAC Symposium on Parameter Estimation and System Identification, York, UK, 1985.
- [16] P.C. Parks, 'Liapunov redesign of model reference adaptive control systems', IEEE Trans. Auto. Contr., AC-11, No.4, pp362-367, 1966.
- [17] R.V. Monopoli, 'Model reference adaptive control with an augmented error signal', IEEE Trans. Auto. Contr., AC-19, No.5, pp474-485, Oct. 1974.

## 5. GENERAL DISCUSSION

### 5.1 Introduction

Having presented three rather detailed case studies into the dynamics of adaptive control, it is time to draw the threads together in order to get a broader picture. What have we learned about adaptive control?

In view of our examples the perplexing question 'What makes adaptive control "adaptive"?' arises. From a dynamical systems point of view there is no test on a controller for "adaptivity". The same nonlinear controller and nonlinear behaviour could have been reached by a number of routes, one of which is adaptive. Consequently it makes little sense to speak of "the adaptive control problem", or to search for a universal or monolithic adaptive control theory. It is just not so well defined. It is our opinion that adaptive control describes a design methodology. This point of view has been introduced in Chapter 1, where we considered adaptive control as an automation of the identify-then-control procedure (cf. Section 1.1, equations (1.1.1-1.1.3)), and its validity is confirmed by our examples. In section 5.2 we address this question in more detail.

In the search for the robust control algorithm, regardless of what this may precisely stand for, two schools have emerged. There is the global approach (cf. [1,2,3]) in which global stability is pursued through algorithm modifications. While asymptotic results and boundedness are addressed, little or no attention is paid to actual control performance, especially its transient aspect is neglected. In the local (cf. [4]) approach the emphasis is on the control performance, at the expense of disregarding large deviations from the desired (ideal) behaviour. In this current work the no man's land between the local and global theories is explored. Some mechanisms for the local theory to stop being valid and for global stability to occur in conjunction with unacceptable transient and/or asymptotic behaviour are identified. This local versus global issue is concisely discussed in section 5.3. We arrive at the disappointing conclusion that it is hard to guarantee good performance without local restrictions. Suitable (global) performance is not only unlikely but is also very hard to establish due to the presence of nonlinear effects exemplified by chaotic dynamics.

The transient behaviour, even when the plant to be controlled is exactly

matched, is not at all well understood. The transient behaviour may be chaotic and sensitive to changes in the initial conditions. Some observations about transient behaviour and analysis based on our case studies are discussed in section 5.4.

We end this chapter on general discussion with some remarks about sufficient information to construct a reasonably performing adaptive control algorithm.

### 5.2 What Makes Adaptive Control Adaptive?

Is it because an adaptive controller can stabilize any member in a large class of systems and track a changing plant whilst maintaining its control objective? No doubt, this quality is the ultimate aim of any adaptive control design and is its *raison d'être*. However, only very few analytic results discussing the tracking capabilities of adaptive algorithms are available [5,6]. Usually, one exploits the exponential convergence of the identification scheme to conclude that sufficiently slow time variations can be followed, without too great a loss in the control performance. Most explicit results of this type, which do quantify the amount of time variations that can be tracked and which characterize the resulting loss in control performance, use the condition of slow adaptation. These results are of the same nature as the tracking result obtained for the M.I.T. rule in Chapter 2. The only tracking result (known to us) which allows for "fast" time variations, without loss of control performance concerns the deadbeat control deadbeat identification adaptive controller discussed in Chapter 4. Although this result demonstrates that fast tracking can be achieved by adaptive controllers designed using the classical approach (identification + certainty equivalence + linear control...), there is not enough information available to conclude that adaptive controllers can be distinguished on the basis of their tracking capabilities.

This observation can be strengthened by noticing that an adaptive control algorithm is just one nonlinear controller which can stabilize a large class of linear systems. Many nonlinear control schemes have this property and have the advantage of being very insensitive with respect to parameter fluctuations in the description of the systems. For example compare the periodic feedback law obtained by the fast adaptive regulator of Chapter 4 (the case  $b < 0$ ) with a vibrational control law [7], the similarity is striking. Is vibrational control

adaptive? We give two further illustrations. Consider the first order plant:

$$\dot{x} = -ax + gu, \quad a, g \in \mathbb{R} \quad g > 0 \quad (5.2.1)$$

It is easy to show that with the nonlinear control law

$$u = -bx^3, \quad b > 0 \quad (5.2.2)$$

the system (5.2.1) is stabilized for any  $a, g \in \mathbb{R} \quad g > 0$  fixed, and for any bounded function  $a(t)$ . If regulation is the control objective then one can adjust the control error by increasing  $b$ , as the remaining error is less than  $(|a|/gb)^{\frac{1}{2}}$  in magnitude. An alternative to (5.2.2) is the nonlinear control law

$$\dot{u} = -x^2u - bx^5, \quad b > 0. \quad (5.2.3)$$

A Lyapunov argument quickly reveals that for any bounded function  $a(t)$  the closed loop system (5.2.1)-(5.2.3) is stabilized (see [8] Chapter II). Also it is not hard to demonstrate that both (5.2.2) and (5.2.3) are able to stabilize the second order system :

$$\ddot{x} + a_1\dot{x} + a_2x = u, \quad a_1, a_2 \in \mathbb{R} \quad a_1 > 0. \quad (5.2.4)$$

(See examples 1.9 and 2.8 in [8, Chapter II] for suitable Lyapunov functions respectively for the control law (5.2.2) and (5.2.3).) It is instructive to compare (5.2.3) with the model reference control algorithm discussed in Chapter 3. On the basis of simulation results of both control algorithms only it would be hard to tell which one is adaptive.

It appears that adaptive control should be regarded as a methodology to control: automating the identify-then-control procedure of classical (linear) control theory and practice (Cf. Chapter 1, Section 1.1). The resulting algorithm should produce a different nominal control law for different plants, eliminating in part the plant's influence on the dynamics. In our case studies this property appeared in the following form:

- (P1) For the M.I.T. Rule (Chapter 2): The plant  $k_p Z_p(s)$  influences the dynamics through the time average of  $[Z_m(s)r](t)[Z_p(s)r](t)$ . In the exact modelling case  $Z_m \equiv Z_p$ , its influence disappears. The nominal control law is  $k_c = 1/k_p$ , which changes with the plant.
- (P2) Adaptive Model Reference Control (Chapter 3): The plant  $Z_p(s) = p_1/(s^2 + p_1s + p_1p_2)$  only influenced the dynamics via  $p_1$ ; its effect disappeared for  $p_1 = +\infty$ , which corresponds to the exact

modelling situation. The nominal (constant, linear, output) feedback law is  $u_p = -(p_2 - a)y_p$  which depends on the particular plant.

- (P3) For the Deadbeat Adaptive Regulator (Chapter 4): The plant  $Z_p(z^{-1}) = z^{-1}/(1 - az^{-1} - bz^{-2})$  had only an impact on the dynamics via  $b$ ,  $b \neq 0$  ( $b = 0$  corresponds to exact modelling). The nominal control law is  $u_k = -ay_k$  (which is however never achieved); both the nominal and the actual control change with changing plant.

Although the plant only affects the dynamics whenever model errors are present, the impact of these model errors is enormous: from uniform asymptotic stability in the large to chaos. This implies that the actual control action can differ completely from the nominal one. Also the response of the adaptively controlled system might well be indistinguishable from the system controlled by a "nonadaptive" algorithm.

By these observations, we believe to be justified in stating that the name tag "adaptive" mainly serves an historical purpose. It identifies an approach to control rather than a property of tracking of or adapting to the environment.

### 5.3 Dynamics of Adaptive Control: Global Versus Local Analyses

The examples presented by Rohrs et.al. [9] spurred a major research effort to re-establish the damaged reputation of adaptive control theory. In the style of the first "global" convergence results [10,11,12], valid for ideally modelled plants and challenged by the examples in [9], new results were developed for (modified) adaptive algorithms applied to an incorrectly modelled plant [1,2,3]. In the presence of model errors emphasis is placed on global stability and ultimate boundedness as opposed to asymptotic optimality or achieving exactly the control objective characteristic of the earlier results. Typical for this class of results is the use of a Lyapunov function tailored to some form of model errors in order to design a modified algorithm with enhanced stability properties. As opposed to this global approach a local theory [4] has been developed discussing the stability properties of adaptive algorithms (modified or not) locally at a desired or nominal response. Under the condition of slow adaptation, by using averaging techniques this theory yields design guidelines for good (local) control performance.

Here we have explored the no man's land between these approaches by focussing attention on dynamical behaviour rather than on stability properties

only. The discrepancy between these theories necessitates such an analysis.

Because our examples demonstrate that both local and global response of an adaptively controlled plant may be governed by chaotic dynamics, we believe that it is insufficient to establish only a global stability result. Also it follows (especially from the second case study (Chapter 3)) that modifying the adaptive algorithm cannot eliminate engineering intuition in designing a well performing controller. A careful selection of the main design variables (adaptation gain, input spectrum, control objective) is essential to obtain an adaptive algorithm with a performance which is robust with respect to its design assumptions. The local theory yields useful design guidelines. However guaranteeing good local performance does not exclude the possibility of chaotic and unacceptably large transients for large initial deviations from the desired response. (The last two case studies indicate that large initial conditions may induce all the complicated dynamical phenomena characteristic of large adaptation gains.)

The local theory relies on the slow adaptation condition. This is not an essential prerequisite for a good adaptive response in the presence of model errors, as our deadbeat adaptive regulator demonstrates (Chapter 4). However the "slow adaptation" qualifier is important to ensure that the local results are significant. For the deadbeat adaptive regulator the local results were absent because the adaptive gain was effectively infinite. But in both the M.I.T. rule and the adaptive model reference examples this condition was essential as it avoids (local) instabilities due to high gain or resonance. The important lesson from these examples is that the "slowness" is quantified in terms of the model errors. The more important the model errors the smaller the adaptation gain should be for the local results to be significant.

Remark:

(R.5.1) This property, although expected, could not be revealed by the averaging theory [4,11], but is very clearly expressed in the adaptive model reference example by:

$$\epsilon \left( \frac{r}{a} \right)^2 < p_1 a. \quad (3.5.3)$$

Which can be interpreted as: the effective adaptive gain  $\epsilon(r/a)^2$  (product of algorithm gain  $\epsilon$  and excitation level  $(r/a)^2$ ) must be smaller than a constant



scaled by the amount of modelling errors ( $p_1 = \infty$ , no model errors).  $\square$

It transpires that the local theory is conservative in estimating the domain in state space and in quantifying the magnitude of the adaptive gain for which the adaptive system performs well.

Clearly conditions guaranteeing good global control performance for adaptively controlled systems will not be easily established.

#### 5.4 Towards a Transient Analysis for Adaptive Control

It is a well known feature of nonlinear dynamic systems that their behaviour may depend in a very fundamental way on the initial conditions. It comes therefore as no surprise that the response, both transient and asymptotic behaviour, of an adaptive system depends crucially on the initial conditions.

This is nicely illustrated in our case studies. For the model reference control scheme of Chapter 3, good transient response is obtained for parameter settings which initially stabilized the plant and extremely large transients are observed for parameter settings which destabilize (Figure 3.4). Moreover, for certain parameter settings (adaptive gain  $\epsilon$ , undermodelling parameter  $p_1$ , model pole  $a$ , reference input  $r$ ) the asymptotic dynamics are extremely complex, and very sensitive to changes in the initial conditions (Section 3.7.3). In our last case study we demonstrated that the transient response, in the presence of undermodelling (Sections 4.5 and 4.6) was extremely sensitive to small changes in the initial conditions due to the presence of chaotic dynamics (cf. R.4.28).

In view of the presence of these complicated dynamics caused by modelling errors it is very unlikely that we could obtain general analytic results about the transient response of an adaptive system in the presence of undermodelling; except for the not very useful, but highly nontrivial, result that the transient response might be bounded. In Chapter 4 we were able to solve the transient problem for the deadbeat adaptive algorithm. We were indeed able to describe completely the behaviour of all trajectories of the adaptive system, but only in part analytically. In the chaotic situation (section 4.6) we had to rely on a (well motivated) conjecture and numerically obtained results (cf. R.4.28 and section 4.8). It looks indeed almost impossible to repeat the detailed analysis of Chapter 4 in a more complicated situation.

Even in the case of exact modelling very little is known about the actual

transient behaviour and the effect of the initial conditions. A first attempt to try to answer the questions: "How can we specify sets of initial conditions in state space such that the transient response will be bounded by a given constant and such that the control objective is achieved within a certain allowable error margin in a given time?", and "How do design parameters influence these sets?" is presented in [14]. The methods are based on nonlinear averaging over spatial variables in order to obtain approximations for the transient response over short time intervals.

### 5.5 Necessary Conditions for Adaptive Control?

Mårtensson, in his paper "The order of a stabilizing regulator is sufficient a priori knowledge for adaptive stabilization" [15], demonstrates that the knowledge of the degree of a stabilizing (time invariant) controller is sufficient information to build an adaptive control system to regulate a linear, time invariant (finite dimensional) causal plant. In essence, the controller discussed in [15] is based on a classical linear control law, coupled to an algorithm which performs an exhaustive search through the parameter space of the control law (using an everywhere dense trajectory) cunningly exploiting the exponential decaying or diverging response of a linear system. The search algorithm is the equivalent of the estimator in the parametric approach to adaptive control discussed in this thesis. This type of adaptive control belongs (ironically) to the class of nonparametric adaptive control algorithms (Cf. Chapter 1). Mårtensson concludes his paper with the remark that this nonparametric approach is not practical due to the possibly catastrophic transient response and because the search algorithm continues indefinitely in the presence of disturbances, which implies unstable (unbounded) response.

In all three case studies presented here the knowledge that the controller structure chosen by the designer can indeed stabilize the plant if its parameters were known, is a necessary prerequisite for the success of the algorithm. However, this is not a sufficient condition for good (stable) adaptive response. This necessary condition is expressed in the following form:

(N1) For the M.I.T. Rule (Chapter 2): The sign of the plant's gain  $k_p$  must be known, and  $k_p$  must be bounded below and above in magnitude.

(N2) For the model reference algorithm (Chapter 3): The second order

plant with transfer function  $Z_p(s) = p_1/(s^2+p_1s+p_1p_2)$  must be stabilizable by constant output feedback, i.e.  $p_1 > 0$ . (cf. (R.3.8))

- (N3) For the dead beat adaptive regulator (Chapter 4): The second order plant with transfer function  $Z_p(z^{-1}) = z^{-1}/(1-az^{-1}-bz^{-2})$  must be stabilizable by constant output feedback, i.e.  $|b| < 1$ .

That this condition is not sufficient comes about in the following way:

- (S1) For the M.I.T. Rule (Chapter 2): In order to obtain good adaptive control performance, the input spectrum and the adaptive algorithm gain have to be restricted.
- (S2) For the model reference algorithm (Chapter 3): Good local performance is guaranteed under the same conditions as above (S1).
- (S3) For the dead beat adaptive regulator (Chapter 4): In this case we were able to find a necessary and sufficient condition for globally stabilizing and optimal adaptive performance  $|b| < \frac{1}{2}$ , as opposed to the necessary condition  $|b| < 1$ .

We conclude that our case studies indicate that the knowledge that the chosen control structure indeed can stabilize the plant if its parameters were known, may be a necessary but is in general not a sufficient condition for the success of the parametric adaptive control approach. This is in sharp contrast to the nonparametric adaptive control methodology.

### 5.6 Some final observations

The dynamics of adaptive control algorithms, in particular in the presence of modelling errors but also in the situation that the plant is correctly modelled, are complicated and difficult to characterize. In view of our case studies this is an understatement. This is not surprising; an adaptive system is a highly nonlinear system whose dynamics depend in a very nonlinear fashion on the design parameters and the input characteristics.

Although we have demonstrated a wide variety of possible dynamical phenomena in adaptively controlled linear systems, we did not exhaust the possibilities. Glancing over simulation results readily available in the literature we quickly realize that there is a wealth of different nonlinear phenomena with a disturbing frequency of occurrence [16]. In particular we did not discuss any of the effects of insufficient (external) excitation, linked to drift instabilities and bursting phenomena [17,18].

This should however not frustrate us or discourage the use of adaptive control algorithms. The overwhelming volume of literature on chaotic behaviour in "real" world problems is simply another indicator of the insufficiency of our mathematics to describe the "real" world in detail; a fact engineers have always been able to live with. Our observations however urge us to be cautious when using classical tools such as Lyapunov stability theory in the adaptive control context and when interpreting any stability conclusions arrived at in this way. It also follows from our analysis that the results and techniques of the geometric theory and global analysis of nonlinear dynamical systems can yield valuable (indispensable) information about the performance of an adaptive algorithm which complements the more classical Lyapunov type results (global results) and the local results obtained for slow adaptation via linearization and/or averaging. If the reader is convinced and aware of these facts we have attained what we set out to do.

5.7 References

- [1] P. Ioannou, P.V. Kokotovic, 'Adaptive Systems with Reduced Models', Lecture Notes In Control and Information Sciences, Springer-Verlag, 1983, No.47.
- [2] B.B. Peterson, K.S. Narendra, 'Bounded error adaptive control', IEEE Trans. Auto. Contr., Vol AC-27, No6, Dec 82, pp.1161-1169.
- [3] G.C. Goodwin, D.J. Hill, D.Q. Mayne, R.H. Middleton, 'Adaptive robust control, (Convergence, stability and performance)', Technical Report EE8544, Newcastle University.
- [4] B.D.O. Anderson, R.R. Bitmead, C.R. Johnson Jr., P.V. Kokotovic, R.L. Kosut, I.M.Y. Mareels, L. Praly, B.D. Riedle, Stability of Adaptive Systems: Averaging and Passivity Analysis, MIT Press, Boston, MA, 1986
- [5] A. Benveniste, G. Ruget, 'A measure of the tracking capability of recursive stochastic algorithms with constant gains', IEEE Trans. Auto. Contr., Vol AC 27, No3, Jun. 82, pp639-650
- [6] K.S. Sin, G.C. Goodwin, S.W.Chan, 'Performance of adaptive filters for time -varying systems', Inversity of Newcastle (N.S.W. Australia), Technical Report EE8010, 1980
- [7] S.M. Meerkov, 'Principle of vibrational control: theory and applications' IEEE Trans. Auto. Contr., Vol AC 25 , No4, Aug. 1980, pp755-762.
- [8] N. Rouché, P. Habets, M. Laloy, 'Stability Theory by Lyapunov's Direct Method', Applied Mathematical Sciences Vol22, Springer-Verlag, 1977
- [9] C. Rohrs, L. Valavani, M. Athans, G. Stein, 'Robustness of continuous time adaptive control algorithms in the presence of unmodelled dynamics', IEEE Trans.Auto. Contr., AC-30, Sept.1985, pp.881-889.
- [10] G.C. Goodwin, P.J. Ramadge, P.E. Caines, 'Discrete time multivariable adaptive control', IEEE Trans. Auto. Contr., Vol.AC-25, No.3, June 1980, pp.449-456.
- [11] A.S. Morse, 'Global stability of parameter adaptive control systems', IEEE Trans. Auto Contr., Vol.AC-25, No.3, June 1980, pp. 433-440.

- [12] K.S. Narendra, Y.H. Lin, L. Valavani, 'Stable adaptive controller design, Part II: Proof of stability', IEEE Trans. Auto Contr., Vol.AC-25, No.3, June 1980, pp.440-449.
- [13] R.L. Kosut, B.D.O. Anderson, I.M.Y. Mareels, 'Stability theory for adaptive systems: methods of averaging and persistency of excitation', Proc. 24th IEEE Conf. on Decision and Control, Ft. Lauderdale, FL, pp478-483; IEEE Trans. Auto. Control, to appear.
- [14] R.L. Kosut, I.M.Y. Mareels, B.D.O. Anderson, R.R. Bitmead, C.R. Johnson Jr., 'Transient Analysis for Adaptive Systems', submitted to the 10th IFAC World Congress, Munich, 1987
- [15] B. Mårtensson, 'The order of a stabilizing regulator is sufficient a priori knowledge for adaptive stabilization', Systems&Control Letters, Vol.6, July 1985, pp87-91.
- [16] R.R. Bitmead, I.M.Y. Mareels, "Nonlinear Dynamics of Adaptive Control", Proc. of the 3rd I.E. (Aust.) Conf. on Contr. Engineering, Sydney Australia, May 1986, pp.35-39.
- [17] B.D.O. Anderson, 'Adaptive systems, lack of persistency of excitation and bursting phenomena', Automatica, Vol.21, No.3, May 1985, pp. 247-259.
- [18] B.D. Riedle, P.V. Kokotovic, 'Stability-instability boundary for disturbance-free slow adaptation and unmodeled dynamics', IEEE Trans. Auto. Contr., Vol. AC-30, pp.1027-1030.