# STABILITY AND ROBUSTNESS OF ADAPTIVE POLE-ZERO PLACEMENT ALGORITHM

A thesis
presented for the degree of
Doctor of Philosophy in Electrical Engineering
in the
University of Canterbury
Christchurch, New Zealand

by

Chan Chok You

1987

*To every thing*
*there is a season,*
*and a time*
*to every purpose*
*under the heaven.*

*Ecclesiastes 3:1*

ABSTRACT

A computationally efficient pole-zero placement algorithm for explicit adaptive control of discrete-time plants is presented. It is effectively an implicit algorithm in the sense that the controller design stage is trivial. Although the algorithm is restricted to open-loop stable plants, it is applicable to nonminimum phase plants.

Results concerning the adaptive control of linear, time-invariant plants having purely deterministic or stochastic disturbances are given. In the deterministic case, it is shown that the adaptive control algorithm ensures that the purely deterministic disturbances are removed from the system output and that asymptotic perfect tracking is achieved. In the stochastic case, it is shown that the adaptive control algorithm leads to the required stability properties of the closed-loop system.

The robustness properties of the modified adaptive control algorithm in the presence of bounded external disturbances and unmodelled dynamics are also given. It is shown that if the modelling error is sufficiently small relative to the normalizing signals, then the algorithm ensures the boundedness of the input-output signals.

ACKNOWLEGEMENTS

I would like to thank my supervisor, Dr. H.R. Sirisena, for his guidance and advice during the course of my study.

I would also like to thank Dr. L. Praly of CAI -Ecole des Mines de Paris for his helpful discussions.

Finally, I would like to thank my family, without whose support and understanding, this study would not have been possible.

CONTENTS

# CHAPTER 1

# INTRODUCTION

The idea of adaptive control had its origin in the early 1950's. However, due to lack of theory and implementation difficulties interest waned. It was not until the early 1970's that these stumbling blocks were overcome and interest was renewed. An excellent survey of the adaptive control field is provided by Astrom (1983).

But what is adaptive control ? It is not easy to define. In fact, a good defintion seems hard to come by in spite of the vast literature that exists on adaptive control. Astrom (1983) takes the view that it is a special type of nonlinear feedback control. For Goodwin and Sin (1984), it is a type of control problem when the plant models and disturbances are not completely specified. However, the concepts involved are simple. An adaptive controller generally consists of two elements - one that identifies the plant parameters and the other which adjusts the controller parameters.

The two different approaches to adaptive control that have attracted much interest are model reference adaptive control and self-tuning control.

The basic idea in model reference adaptive control is to cause the system to behave like a given model. Landau (1979) gives a comprehensive study of work up to 1977.

The idea of self-tuning control seems due to Kalman (1958). It was later revived by Astrom and Wittenmark (1973) who developed the self-tuning regulator in a stochastic framework and showed that the algorithm has some nice properties based on the convergence of the estimated parameters. Subsequently, Clark and Gawthrop (1975) provided some modifications to the original self-tuning regulator. Since then, numerous papers on the subject have appeared.

Adaptive control algorithms can be divided into two broad classes: implicit (direct) and explicit (indirect) algorithms. In an implicit algorithm, the controller parameters are directly estimated. In an explicit algorithm, the parameters of the plant are estimated instead.

Most of the earlier works on adaptive control have concentrated on algorithmic implementation and simulation of linear systems. Later, proofs of convergence for the different versions of the adaptive controller appeared. By convergence, it is meant that the control objective is asymptotically achieved and all input and output signals remain bounded for any bounded initial conditions. Various assumptions on the plant are required to develop a convergent adaptive control scheme.

There are several surveys on the convergence theory of adaptive control algorithms ( e.g., Goodwin et al 1984, Kumar 1985 ). The latter provides a comprehensive survey of results on stochastic adaptive control.

Complete stability and convergence results have now been obtained for minimum phase deterministic and stochastic systems ( see the surveys by Goodwin et al 1984 and Kumar 1985 ). The results are restricted to the model reference type control algorithms, which used pole-zero cancellation methods.

Model reference adaptive control algorithms have been developed either in a continuous-time or in a discrete-time framework. It would appear that the continuous-time algorithms can be implemented digitally with little error if the sampling period is sufficiently small. However, Astrom et al (1984) have demonstrated that all continuous-time systems with poles excess larger than two will always give sampled systems with unstable zeros. This leads to the conclusion (M'saad et al 1985) that nonminimum phase characteristic are much more prevalent for sampled data systems than for continuous-time systems. A consequence of this is that discrete-time model reference control is ruled out. On the other hand, there are many results pertaining to continuous-time model reference control of such systems.

Recently, Goodwin et al (1986) have shown that the apparent paradox between continuous-time and discrete-time model reference control can be resolved by a slight modification to the usual discrete-time model format. The new model format is shown to retain the key features of the continuous-time model and thereby allows a model reference control law to be designed which is guaranteed to be exponentially stable.

However, the problem of the adaptive control of nonminimum phase systems based on more complicated design procedures has not been completely solved. The principal difficulty with the proposed approaches has been that a stabilizability or controllability condition over the estimated model may arise. A comprehensive review of the problem of designing adaptive controllers for nonminimum phase systems is given by M'saad et al (1985).

Despite this difficulty, a number of stability results have been obtained. For example, in Elliot, Cristi and Das (1985), Anderson and Johnstone (1985), Goodwin and Teoh (1985) global stability has been established on the basis of persistency of excitation, or parameter convergence. This requirement may be removed by introducing a correction procedure to the parameter estimates ( e.g., De Larminat 1984, Lozano and Goodwin 1985 ). In Kreisselmeier (1985), a continuous time adaptive control approach is made which requires the plant parameters to lie in a known convex set where no unstable pole-zero cancellation occurs.

It should be stressed that the preceding stability results have been established only for deterministic nonminimum phase systems. For stochastic systems, however, very few results have been obtained to date. Nevertheless, Hersh and Zarrop (1986) have developed a convergent stochastic adaptive control algorithm by imposing some requirements on the asymptotic behaviour of the parameter estimates.

Until now, the emphasis has been on explicit adaptive pole placement schemes. A major drawback of the adaptive pole placement algorithm is that the controller design step is nontrivial. Specifically, it involves the solution of a pole placement Diophantine equation, and so suffers from computational complexity and numerical stability problems.

Motivated by such considerations, some recent studies have been made to derive adaptive control algorithms applicable to nonminimum phase systems which do not require the solution of a polynomial identity. In Astrom and Wittenmark (1980), and Praly (1984) implicit algorithms for deterministic systems have been proposed, but they involve bilinear estimation problems. Implicit algorithms which use linear parameter estimation techniques for directly identifying the controller parameters are presented in Elliot (1982), Karam, Warwick and Farsi (1986). Global stability for the deterministic adaptive control algorithm in Elliot (1982) has been established in Elliot, Cristi and Das (1985) subject to the assumption that the reference signal is a sum of sinusoids. As the authors there remarked, the limitation to sinusoids is severe. However, no self-contained proof of convergence of the stochastic adaptive control algorithm in Karam et al (1986) is available. Explicit algorithms for deterministic systems have also been proposed ( e.g. Sirisena and Teng 1986, Lin and Chen 1986 ). In Lin and Chen (1986), a spectral factorization problem is introduced in the controller design step. Despite the fact that no solution of an identity is required, the estimated model stabilizability problem remains. In Sirisena and Teng (1986), no such factorization is involved and the stabilizability problem is avoided by restricting the algorithm to open-loop stable plants. Convergence proofs of the algorithms have yet to appear.

Global convergence of adaptive model reference control schemes for minimum phase systems having purely deterministic disturbances has been established (Goodwin and Chan 1983). An extension of Elliot's implicit adaptive pole placement algorithm (1982) for systems having purely deterministic disturbances has also been presented by Janecki (1987). However, the problem of extending the explicit adaptive pole placement algorithm to systems having purely deterministic disturbances is, as yet, unresolved. The difficulty has been that purely deterministic disturbances give rise to common factors having roots on the unit circle.

It seems fair to say that the sole objective in the synthesis of adaptive control algorithms has been the global stability of the closed-loop system. However, it should be noted that all the stability proofs of the adaptive control algorithms in the literature have been established under idealized conditions. The typical assumptions in these stability proofs are that the plant is linear and time invariant. Moreover, the upper bound of the system order is assumed to be known. Such requirements are restrictive, and hence impractical. In fact, it has been shown in Rohrs et al (1985) that a stable adaptive control algorithm is not necessarily stable in the presence of unmodelled dynamics. That is, it is not robustly stable.

Achieving stability of adaptive control algorithms in the presence of bounded external disturbances is a first step towards resolving the robustness problem. Most of the proposed approaches involve modifications to the adaptive laws. For example, in Egardt (1979), Peterson and Narendra (1982), Samson (1983), a dead zone is used in the adaptive law. The motivation is to switch off the parameter adaptation algorithm when the prediction error is less than a certain threshold. In Egardt (1979), Kreisselmeier and Narendra (1982), a projection in the adaptive law is used to restrict the parameter estimates to lie within a bounded region. In Ioannou and Kokotovic (1984), a $\sigma$-modification, i.e., an adaptive law with an additional term $-\theta\sigma$, is suggested. Another approach (Narendra and Annaswamy 1986) to deal with bounded disturbances requires the reference input to be persistently exciting inorder to attain exponential stability of the adaptive system. In an another paper, Narendra and Annaswamy (1987) propose a new adaptive law in which the constant $\sigma$ in Ioannou and Kokotovic (1984) is replaced by a term proportional to $|e_1|$ where $e_1$ is the output error. This $e_1$-modification is shown to improve the performance of the system in all respects.

However, when the true plant is not accurately modelled by an assumed linear model, the unmodelled dynamics become an external disturbance which cannot be assumed to be bounded. Therefore, these approaches for bounded disturbances do not necessarily solve the robustness problem.

Despite this difficulty, a number of robustness results are now available for a wide class of adaptive control algorithms. The first robustness results were obtained in Gawthrop and Lim (1982) and Lim (1982). They are derived in terms of the design parameters available. This work has one drawback in that the results derived use an a priori plant signal-boundedness assumption. In Kosut and Friedlander (1985) conditions under which a class of continuous-time adaptive controllers will preserve stability despite unmodelled dynamics are derived. There, the concept of a tuned system ( i.e., a control system that could be obtained if the plant were known ) is introduced. In Ioannou and Kokotovic (1984), it is shown that modification of the adaptive law for a continuous-time model reference adaptive control scheme can give tracking to within a uniform bound in the presence of unmodelled dynamics. The result has been extended in

Ioannou (1986) where a smaller residual set for the tracking error can be obtained and which reduces to zero when the unmodelled dynamics disappear. Later, a continuous-time direct adaptive control algorithm is proposed ( Ioannou and Tsakalis 1986 ) which is robust with respect to additive and multiplicative plant unmodelled dynamics. It is shown that, subject to gain and frequency restrictions on the unmodelled part, the algorithm guarantees boundedness of the input-output signals. It is also shown in Narendra and Annaswamy (1987) that further modification of the new adaptive law (which utilizes the output error magnitude) renders it applicable to the problem of adaptively controlling a plant in the presence of a class of unmodelled dynamics.

Other interesting approaches to the robustness problem have also been suggested. For example, in Praly (1983) an explicit adaptive control approach is made which uses signal normalization and projection in the adaptive law. Robust stability is established under the assumptions that bounds on the unknown plant parameters are known and that the estimated system is uniformly controllable and observable. In Ortega, Praly and Landau (1985) conditions for stability are derived for an implicit adaptive controller where the parameter estimator is modified in terms of normalized signals. The stability conditions require the existence of a linear controller such that the closed-loop transfer function satisfies certain conic conditions. In this work, the use of normalized signals in the parameter estimation algorithm for improved robustness completes the results of Gawthrop and Lim (1982) and Lim (1982). Furthermore, a robust approach is proposed in Kreisselmeier (1986) for an explicit continuous-time adaptive control scheme which uses signal normalization in combination with a dead zone and projection in the adaptive law. Robust stability is then shown on the assumption that the unknown plant parameters lie in a known convex set where no unstable pole-zero cancellation occurs. The ideas have been extended to the model reference adaptive control problem in Kreisselmeier and Anderson (1986). Robust stability is shown subject to the assumptions that bounds on the plant parameters and the exponential bounds on the impulse response of the inverse plant transfer function are known. In Hsu and Costa (1987), a continuous-time adaptive control law based on saturation and inclusion of a discontinuous $\sigma$-factor is proposed to improve the robustness of the adaptive system with respect to fast unmodelled dynamics. The introduction of the discontinous $\sigma$-factor represents an extension of an earlier work (e.g.

Ioannou and Kokotovic 1984). The advantages of this extension include global stability, with less restrictive assumptions. In Cluett et al (1987), a stable discrete-time adaptive controller in the presence of unmodelled dynamics and bounded disturbances is presented. The proposed formulation uses an augmented plant representation that incorporates weighting polynomials for the system-output, input and setpoint respectively into the predictive control law. The algorithm also includes a normalized least squares estimation scheme and a parameter adaptation stopping criterion.

In this thesis, a pole-zero placement algorithm for the adaptive control of plants which are not necessarily minimum phase is presented. Although technically an explicit algorithm, it is effective -ly an implicit algorithm in the sense that the controller design step is trivial. Moreover, it is applicable to systems having purely deterministic disturbances. The controller parameters can be obtained trivially from the estimated plant parameters, with no solution of an identity necessary. Consequently, the adaptive algorithm is not only computationally efficient but also side steps the estimated model stabilizability problem. The only technical weakness of this approach is that the plant must be stable. The stability of the resulting closed-loop adaptive control systems subject to external disturbances and unmodelled dynamics is also addressed.

The main contributions of the thesis are the following:

(1) A computationally efficient algorithm for the adaptive control of a class of nonminimum phase systems is developed;

(2) Establishes stability and convergence of the adaptive control for a class of nonminimum phase systems having purely deterministic disturbances;

(3) A solution to the problem of the stochastic control of a class of nonminimum phase systems is provided;

(4) A robust adaptive controller for a class of nonminimum phase systems is proposed.

The content and arrangement of the thesis are as follows.

Chapter 2 presents the development of the pole-zero placement algorithm for SISO (single-input single-output) and MIMO (multi-input multi-output) servo systems for the nonadaptive case.

Chapter 3 presents the deterministic adaptive control and theoretical aspects of the algorithm. Simulation examples are given to illustrate the results.

Chapter 4 presents the stochastic adaptive control and theoretical aspects of the algorithm. Both white noise and coloured noise disturbances are considered.

Chapter 5 presents the following: some mathematical background concerned with the input-output properties of dynamical systems; stability criteria; development of design guidelines for improving robustness of the nonadaptive algorithm in the presence of modelling errors. Examples are provided to illustrate the results.

Chapter 6 presents a modified adaptive control algorithm which is robust with respect to unmodelled dynamics and bounded disturbances.

Chapter 7 presents the conclusions and further areas for future research.

# CHAPTER 2

# POLE-ZERO PLACEMENT

## 2.1 Introduction

Pole-zero placement self-tuning controllers were developed for various reasons, the main ones being the following. In the regulator case ( Wellstead et al 1979 ), they provide a means of overcoming the restriction to minimum phase systems of the original self-tuning regulator (Astrom and Wittenmark 1973). In the servo case, they provide the ability to introduce bandwidth and damping ratio as tuning parameters (Astrom and Wittenmark 1980).

However, there is a major drawback of the standard pole-zero placement strategy as compared to the minimum variance strategy (Astrom and Wittenmark 1973). Specifically, the plant parameters rather than the controller parameters are estimated. Moreover, the subsequent control law computation stage involves the solution of a set of equations satisfying a pole placement identity. Thus it suffers from the much increased computational effort required to calculate the control signal. Moreover, its stability is compromised by possible unstable pole-zero cancellations in the estimated model.

In Prager and Wellstead (1981) a multivariable version of the pole placement self-tuning regulator (Wellstead et al 1979) is presented. Besides suffering from the drawbacks of being an explicit algorithm, an additional computation step is incurred in the control synthesis stage as a result of having to transform from a right to a left matrix-fraction description. Since the algorithm is developed for systems with random noises, it is essentially a regulator. Reference tracking is handled by (arbitrarily) incorporating an integrator in each loop.

Recently, efforts have been made to develop pole-zero placement algorithms where no control synthesis step is necessary. For example, in Elliot (1982), Karam et al (1986) implicit pole placement algorithms for deterministic and stochastic systems, respectively have been presented. In Sirisena and Teng (1986), Lin and Chen (1986) explicit pole-zero placement algorithms have also been presented.

In the following section, a review of the aforementioned pole-zero placement approaches for explicit adaptive control is given, since this thesis is concerned with adaptive methods based on estimation of the parameters in an explicit process model.

In this chapter, a pole-zero placement algorithm for SISO and MIMO servo systems is developed. A key feature of the algorithm is that it avoids the solution of a Diophantine equation. As a result, it is not only computationally efficient but also side steps the system stabilizability problem. It can also handle steady-state errors in the output(s) in a straightforward manner, without using feedforward elements or arbitrarily cascading integrators. The only limitation is that the plant must be stable, though not necessarily minimum phase.

The organization of this chapter is as follows. Section 2.2 provides a review of the existing pole-zero placement approaches for SISO systems. Section 2.3 - 2.4 presents the pole-zero placement for SISO systems. Section 2.5-2.7 presents the MIMO versions of an existing and the new pole-zero placement algorithms.

## 2.2 Review of Existing Pole-zero Placement Approaches

In this section, a review of the classical pole placement, the pole-zero placement of Astrom and Wittenmark (1980), and the pole-zero placement of Lin and Chen (1986) is given. The pole-zero placement of Sirisena and Teng (1986) is covered in a later section. The discussion given here is only limited to SISO systems.

### 2.2.1 Pole placement

Consider a system described by the following representation

$$A(d)y(t) = B'(d)u(t-k) \tag{2.1}$$

where $y(t)$ and $u(t)$ denote the system output and input, respectively. $k \geq 1$ is an integer time delay corresponding to the system time delay. $A(d)$, $B'(d)$ are polynomials in the unit delay operator $d$ defined as

$$A(d) = 1 + a_1 d + \ldots\ldots + a_{na} d^{na}$$

$$B'(d) = b_0 + b_1 d + \ldots\ldots + b_{nb} d^{nb}$$

Consider a control law of the form

$$F(d)u(t) = G(d)[y^*(t) - y(t)] \tag{2.2}$$

where $y^*(t)$ is the desired value, and $F(d)$ and $G(d)$ are polynomials in $d$ given by

$$F(d) = f_0 + f_1 d + \ldots\ldots + f_{nf} d^{nf}$$

$$G(d) = g_0 + g_1 d + \ldots\ldots + g_{ng} d^{ng}$$

Combining (2.1) and (2.2) yields the closed-loop equation

$$\left[A(d)F(d) + d^k B'(d)G(d)\right]y(t) = d^k B'(d)G(d)y^*(t) \tag{2.3}$$

The closed-loop poles are assigned by $P(d)$ such that

$$A(d)F(d) + d^k B'(d)G(d) = P(d) \tag{2.4}$$

where $P(d)$ is an arbitrary stable polynomial with

$$P(d) = 1 + p_1 d + \ldots\ldots + p_{np} d^{np}$$

If $A(d)$ and $B'(d)$ are relatively prime, $F(d)$ and $G(d)$ can be solved from (2.4). A unique solution of (2.4) exists if

$$nf = nb + k - 1$$

$$ng = na - 1$$

$$np \leqq na + nb + k$$

(2.4) can be expressed in the following matrix form

$$\begin{bmatrix} 1 & & & & & \\ a_1 & \ddots & & b_0 & & \\ \vdots & & 1 & \vdots & \ddots & b_0 \\ a_{na} & & \vdots & b_{nb} & & \vdots \\ & \ddots & & & \ddots & \\ & & a_{na} & & & b_{nb} \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_{nf} \\ g_0 \\ \vdots \\ g_{ng} \end{bmatrix} = \begin{bmatrix} 1 \\ p_1 \\ \vdots \\ \vdots \\ \vdots \\ p_{np} \end{bmatrix} \tag{2.5}$$

If the polynomials $A(d)$ and $B'(d)$ contain common roots, then the matrix on the left-hand side is singular. Near singularity of the matrix must be also be avoided because the coefficients of $F(d)$ and $G(d)$ are required to be bounded for numerical stability in practical applications.

Thus the main difficulty with this algorithm is that it is not possible to evaluate the control law when the system (2.1) contains unstable pole-zero cancellations, i.e. the system is non-stabilizable.

Special attention has to be given to the problem of steady-state errors in the output. If necessary, an integrator may be incorporated with the system.

2.2.2 Pole-zero placement (Astrom and Wittenmark 1980)

The pole placement algorithm considered in the last section only specifies the closed-loop poles. In Astrom and Wittenmark (1980) the control law is more complicated in that zeros can also be placed.

Consider a system described by

$$A(d)y(t) = B'(d)u(t-k) \tag{2.6}$$

where A(d) and B(d) are defined as in (2.1). It is assumed that A(d) and B(d) are relatively prime.

A control law of the form

$$H(d)y^*(t) = G(d)y(t) + F(d)u(t) \tag{2.7}$$

is considered. The resulting closed-loop transfer function is given by

$$T_f(d) = \frac{d^k\ B'(d)H(d)}{A(d)F(d) + d^k\ B'(d)G(d)} \tag{2.8}$$

This transfer function is made equal to a prespecified model

$$T_{fp}(d) = \frac{d^k\ B_m'(d)}{A_m(d)} \tag{2.9}$$

where $A_m(d)$ and $B_m'(d)$ are assumed relatively prime, and may be normalized such that

$$\frac{B_m'(1)}{A_m(1)} = 1 \tag{2.10}$$

The control signal is given by

$$u(t) = \frac{A(d)}{B'(d)}\ \frac{B_m'(d)}{A_m(d)}\ y^*(t) \tag{2.11}$$

If there are roots of $B'(d)A_m(d)$ outside some region, then this control has undesirable modes. Hence, the roots of $B'(d)A_m(d)$ must lie within some restricted region. B'(d) may be factored as

$$B'(d) = B_u(d)B_s(d) \tag{2.12}$$

where $B_s(d)$ has roots corresponding to well-damped modes and all roots of $B_u(d)$ correspond to unstable or poorly damped modes. From (2.11) $B_m'(d)$ must be of the form

$$B_m'(d) = B_{m_1}(d)B_u(d) \qquad (2.13)$$

Now, since degree of $A_m \leq$ degree of $(AF + d^k B'G)$ there are factors in (2.9) which cancel. It can be shown that this cancelling factor $A_o(d)$ is an observer in state-space theory. Also, $A_o(d)$ has roots in the restricted stability region. The design method thus consists of the following steps:

(1) Choose $A_m(d)$, $B_m'(d)$ and $A_o(d)$

(2) Solve for $F_1(d)$ and $G(d)$ from

$$A(d)F_1(d) + d^k B_u(d)G(d) = A_m(d)A_o(d) \qquad (2.14)$$

(3) Use the control law (2.7) with

$$F(d) = F_1(d)B_s(d) \quad \text{and} \quad H(d) = A_o(d)B_{m_1}(d)$$

The pole placement equation (2.14) has infinitely many solutions, further discussion is given in Astrom and Wittenmark (1980).

To perform the design, the polynomial $B'(d)$ must be factorized so that the decomposition $B_u(d)B_s(d)$ can be made( in the self-tuning case, this decomposition has to be performed at each step). The decomposition problem can be avoided in cases where all process zeros are cancelled (as in model reference control) or where no process zero is cancelled( pole placement).

In the self-tuning case, the strategy where all process zeros are cancelled leads to what is often called 'implicit' algorithms, where the design calculations are simplified considerably. In this type of algorithms the parameters of the controllers are updated directly. However, the resulting algorithms are only applicable to minimum phase systems.

There are some drawbacks associated with the proposed self-tuning control algorithms. The strategy where no process zero is cancelled leads to the solution of a Diophantine equation. Integral action relies on the use of feedforward (which is an open-loop strategy). Thus, when the parameter estimates are biased, good control is not attainable.

2.2.3 Pole-zero placement (Lin and Chen 1986)

The algorithm proposed by Lin and Chen (1986) is intended to overcome the much computational effort required to synthesize an adaptive controller with desired pole-zero placement. It does not require the solution of the Diophantine equation, although it involves factorization of the model estimates of A(d) and B(d).

Consider the system described by

$$A(z)y(z) = B(z)u(z) \tag{2.15}$$

where A(z) and B(z) are relatively prime polynomials given by

$$A(z) = z^{na} + a_1 z^{na-1} + \ldots\ldots + a_{na}$$

$$B(z) = b_0 z^{nb} + b_1 z^{nb-1} + \ldots\ldots + b_{nb}$$

The control law is to be of the form

$$F(z)u(z) = G(z)e(z) \tag{2.16}$$

where e(z) is the tracking error defined by

$$e(z) = r(z) - y(z) \tag{2.17}$$

r(z) is the reference signal, assumed to have the representation

$$r(z) = \frac{N(z)}{M(z)}$$

$$= \frac{N(z)}{M_s(z)M_u(z)} \tag{2.18}$$

where M(z) and N(z) are relatively prime polynomials, and the polynomials $M_s(z)$ and $M_u(z)$ represent factors of M(z) having their zeros in $|z| < 1$ and $|z| \geq 1$, respectively.

Let the sensitivity function be defined as

$$S(z) = (1 + P(z)C(z))^{-1} \tag{2.19}$$

where P(z) and C(z) are the plant and the controller, respectively. Thus the tracking error is given by

$$e(z) = S(z)r(z) \tag{2.20}$$

To track high-order reference signal r(z), S(z)r(z) must not have any pole in $|z| \geq 1$, i.e. S(z) must have a sufficient number of zeros to cancel the poles of r(z) in $|z| \geq 1$. Thus, the design objective is to synthesis a controller C(z) such that S(z) has all desired poles and some desired zeros for the purpose of high-order reference signal tracking.

From (2.18) and (2.20), $S(z)$ must be of the form

$$S(z) = \frac{W(z)M_u(z)}{g(z)} \tag{2.21}$$

where $g(z)$ is a polynomial containing zeros in $|z| < 1$, and $W(z)$ is a monic polynomial to be determined to satisfy the following internal stability constraint.

Definition: The sensitivity function $S(z)$ is said to be internally stable (or realizable) if the resulting closed-loop system is asymptotically stable for some choices of the controller $C(z)$, i.e. no pole-zero cancellation between $C(z)$ and $P(z)$ in $|z| \geq 1$.

Lemma 2.1

The sensitivity function $S(z) \neq 0$ is internally stable if, and only if, all the following conditions hold:

(a) $S(z)$ is analytic in $|z| \geq 1$

(b) every zero of $A(z)$ in $|z| \geq 1$ is a zero of $S(z)$ of at least the same multiplicity

(c) every zero of $B(z)$ in $|z| \geq 1$ is a zero of $1-S(z)$ of at least the same multiplicity.

The polynomials $A(z)$ and $B(z)$ are factorized as follows:

$$A(z) = A_u(z)A_s(z) \tag{2.22}$$

$$B(z) = B_u(z)B_s(z) \tag{2.23}$$

where $A_u(z)$ and $B_u(z)$ have all their zeros in $|z| \geq 1$ while $A_s(z)$ and $B_s(z)$ have all their zeros in $|z| < 1$. Denote

$$B_u(z) = \prod_{i=1}^{n} (z - q_i)^{m_i} \tag{2.24}$$

where $n$ is the number of distinct zeros $q_i$ of $B(z)$ in $|z| \geq 1$, and $m_i$ is the multiplicity of $q_i$. In order to let the desired $S(z)$ satisfy Lemma 2.1-(b), it must be of the form

$$S(z) = \frac{l(z)A_u(z)M_u(z)}{g(z)} \tag{2.25}$$

where $g(z)$ and $M_u(z)$ contain the desired poles and zeros, respectively, and $l(z)$ is a polynomial to be determined by Lemma 2.1-(c).

From (2.25),

$$1 - S(z) = \frac{l(z)A_u(z)M_u(z)}{g(z)} \tag{2.26}$$

To satisfy Lemma 2.1-(c), the numerator of (2.26) must be of the form

$$\begin{aligned} h(z) &= g(z) - l(z)A_u(z)M_u(z) \\ &= B_u(z)E(z) \end{aligned} \tag{2.27}$$

i.e.

$$\begin{aligned} h(q_i) &= g(q_i) - l(q_i)A_u(q_i)M_u(q_i) \\ &= 0 \qquad \text{for } i = 1,2,..,n \end{aligned} \tag{2.28}$$

where n is the number of distinct zeros $q_i$ of B(z) in $|z| \geq 1$, and

$$l(q_i) = \frac{g(q_i)}{A_u(q_i)M_u(q_i)} \tag{2.29}$$

Let

$$l(z) = z^n + l_1 z^{n-1} + \ldots.. + l_n \tag{2.30}$$

(2.29) can be expressed in the following matrix form

$$\begin{bmatrix} q_1^{n-1} & q_1 & 1 \\ q_2^{n-1} & q_2 & 1 \\ | & | & | \\ q_n^{n-1} & q_n & 1 \end{bmatrix} \begin{bmatrix} l_1 \\ l_2 \\ | \\ l_n \end{bmatrix} = \begin{bmatrix} \dfrac{g(q_1)}{A_u(q_1)M_u(q_1)} - q_1^n \\ \dfrac{g(q_2)}{A_u(q_2)M_u(q_2)} - q_2^n \\ | \\ \dfrac{g(q_n)}{A_u(q_n)M_u(q_n)} - q_n^n \end{bmatrix} \tag{2.31}$$

By solving the n equations in (2.31), l(z) can be determined. The term E(z) should be determined once l(z) is determined.

The corresponding controller is now given as

$$\begin{aligned} C(z) &= \frac{1 - S(z)}{P(z)S(z)} \\ &= \frac{A_s(z)E(z)}{B_s(z)l(z)M_u(z)} \end{aligned} \tag{2.32}$$

and the control law (2.16) now becomes

$$B_s(z)l(z)M_u(z)u(z) = A_s(z)E(z)e(z) \tag{2.33}$$

Remarks

As compared to the pole placement of Section 2.2.1 which needs to solve 2p - 1 equations for 2p pole placements, this algorithm only needs to solve n equations, where n is the number of zeros of the plant in the unstable region.

As in the pole placement case, it is required that the system to be controlled is stabilizable, else there will be unstable pole-zero cancellations between C(z) and P(z).

A MIMO version of the algorithm is, at present, unavailable.

## 2.3 Development of Alternative Pole-zero Placement for SISO Systems

In this section, a general pole-zero placement for SISO servo system is developed.

Consider the plant described by

$$A(d)y(t) = B(d)u(t) \tag{2.34}$$

where A(d) and B(d) are polynomials in the unit delay operator d defined as

$$A(d) = 1 + a_1 d + \ldots\ldots + a_{na} d^{na}$$

$$B(d) = b_1 d + \ldots\ldots + b_{nb} d^{nb}$$

where the variation of the deadtime is being catered for by the form assumed by B(d).

The control law is of the form

$$F(d)u(t) = G(d)\left(y^*(t) - y(t)\right) \tag{2.35}$$

where F(d) and G(d) are polynomials which are yet to be determined, and y*(t) is an input reference signal, assumed to be a series of steps.

The closed-loop transfer function relating y to y* is given by

$$T_f(d) = \frac{B(d)G(d)}{A(d)F(d) + B(d)G(d)} \tag{2.36}$$

It is made equal to a prescribed model

$$T_{fp}(d) = \frac{B_m(d)}{A_m(d)} \tag{2.37}$$

where $A_m(d)$ and $B_m(d)$ are assumed relatively prime, and are chosen such that

$$A_m(1) = B_m(1) \tag{2.38}$$

for zero-offset with step $y^*(t)$.

Let $A(d)$ and $B(d)$ be factored as

$$A(d) = A_u(d)A_s(d) \tag{2.39}$$

$$B(d) = B_u(d)B_s(d) \tag{2.40}$$

where all the zeros of $A_s(d)$ (resp. $B_s(d)$) correspond to zeros of $A(d)$ (resp. $B(d)$) outside the unit circle in the complex d-plane(i.e. inside the unit circle in the conventional complex z-plane). As there is no practical objection to cancelling stable polynomials, the controller polynomials can be factored as

$$F(d) = F_1(d)B_s(d) \tag{2.41}$$

$$G(d) = G_1(d)A_s(d) \tag{2.42}$$

where $F_1(d)$, $G_1(d)$ are yet to be determined.

The resulting closed-loop transfer function is given by

$$T_f(d) = \frac{B_u(d)G_1(d)}{A_u(d)F_1(d) + B_u(d)G_1(d)} \tag{2.43}$$

Requiring (2.43) to be equivalent to (2.37) gives

$$G_1(d)B_u(d) = B_m(d) \tag{2.44}$$

$$F_1(d)A_u(d) + G_1(d)B_u(d) = A_m(d) \tag{2.45}$$

If $A_u(d)$ and $B_u(d)$ are relatively prime, then (2.45) has solutions $F_0(d)$ and $G_0(d)$ with

$$\text{degree of } F_0 = \text{degree of } B_u - 1$$
$$\text{degree of } G_0 = \text{degree of } A_m - \text{degree of } B_u$$

or

$$\text{degree of } F_0 = \text{degree of } A_m - \text{degree of } A_u$$
$$\text{degree of } G_0 = \text{degree of } A_u - 1$$

However, (2.45) has general solutions given by

$$F_1(d) = F_0(d) - X(d)B_u(d) \tag{2.46}$$

$$G_1(d) = G_0(d) + X(d)A_u(d) \tag{2.47}$$

where $X(d)$ is any polynomial such that $F_1(1) = 0$ for zero-offset, i.e., such that

$$X(1) = \frac{F_0(1)}{B_u(1)} \tag{2.48}$$

unless $A_u(1) = 0$, i.e. plant already has an integrator. From (2.44) and (2.47)

$$B_m(d) = [G_0(d) + X(d)A_u(d)]B_u(d) \tag{2.49}$$

Hence, $B_m(d)$ cannot, in general, be specified arbitrarily in view of the constraint (2.48) on $X(d)$.

In the following, some special cases are addressed where the design calculations can be simplified.

### 2.3.1 Special case: All process poles cancelled

If all process poles are to be cancelled, then set

$$A_s(d) = A(d) \tag{2.50}$$

$$A_u(d) = 1 \tag{2.51}$$

(2.45) now collapses to

$$F_1(d) + G_1(d)B_u(d) = A_m(d) \tag{2.52}$$

Using (2.44)

$$F_1(d) = A_m(d) - B_m(d) \tag{2.53}$$

A solution thus exists for any $A_m(d)$ and $B_m(d)$. Clearly, $F_1(d)$ and $B_u(d)$ are relatively prime because $A_m(d)$ and $B_u(d)$ are assumed to be relatively prime. Thus, as long as $A(d)$ is stable, then there will be no unstable pole-zero cancellations, though $F_1(d)$ may be unstable.

In view of the restriction (2.38) for achieving zero-offset with step $y^*(t)$, $F_1(1) = 0$. That is, the controller incorporates integral action.

To avoid the decomposition problem, another special case is addressed in the following.

2.3.2 Special case: No process zeros cancelled

If no process zeros are to be cancelled, set

$$B_u(d) = B(d) \tag{2.54}$$

$$B_s(d) = 1 \tag{2.55}$$

(2.43) now becomes

$$T_f(d) = \frac{B(d)G_1(d)}{A_m(d)} \tag{2.56}$$

Without loss of generality, $G_1(d)$ may be normalized such that

$$G_1(1) = 1 \tag{2.57}$$

To satisfy (2.38), $A_m(d)$ in (2.56) must be of the form

$$A_m(d) = P(d)\frac{B(1)}{P(1)} \tag{2.58}$$

where $P(d)$ is monic. Thus using (2.53) and (2.55) gives

$$F_1(d) = P(d)\frac{B(1)}{P(1)} - G_1(d)B(d)$$

$$= F(d) \tag{2.59}$$

The corresponding control law is now given as

$$[P(d)K - G_1(d)B(d)]u(t) = G_1(d)A(d)[y^*(t) - y(t)] \tag{2.60}$$

where

$$K = \frac{B(1)}{P(1)} \tag{2.61}$$

The closed-loop transfer function is now given by

$$T_f(d) = \frac{B(d)G_1(d)}{P(d)K} \tag{2.62}$$

and the corresponding control signal is

$$u(t) = \frac{G_1(d)A(d)}{P(d)K}\, y^*(t) \tag{2.63}$$

Remarks

Here, polynomials $P(d)$ and $G_1(d)$ are introduced which are placed in the denominator and numerator respectively of the transfer function relating $y$ to $y^*$. The zeros of $P(d)$ and $G_1(d)$ will become poles and zeros respectively of the closed-loop system.

The simplified algorithm considered in Astrom and Wittenmark (1980) is characterized by the cancellation of process zeros. Here, the simplified algorithm is characterized by the cancellation of process poles.

As compared to the previous algorithms, this algorithm does not involve the solution of a set of equations and/or spectral factorization, and so is computationally more efficient. Moreover, the solvability question associated with the Diophantine equation in classical pole placement does not arise. In chapter 3, the algorithm will be extended to cover the case when A(d) and B(d) have common roots on the unit circle (i.e., in the presence of deterministic disturbances). The rest of the thesis will concentrate on the control law given by (2.60).

## 2.4 Choices of Pole-zero Polynomials

In this section the choice of the pole and zero placement polynomials P(d) and $G_1(d)$ is discussed.

### 2.4.1 Deadbeat controllers

The choice

$$P(d) = G_1(d) = 1 \tag{2.64}$$

yields the so called DB($\nu$) deadbeat controller Schumann (1979). A practical problem with the DB($\nu$) controller is that it may call for unrealistically large control signals. The DB($\nu$+1) controller Isermann (1981) provides some means of alleviating this problem by trading an increase in the settling time for a reduction in the initial control signal magnitude. The DB($\nu$+1) controller corresponds to the case

$$P(d) = 1 \;, \qquad G_1(d) = \frac{1 + q_1 d}{1 + q_1} \tag{2.65}$$

where the coefficient $q_1$ is chosen as discussed below.

Substituting (2.65) into (2.63) gives

$$u(t) = \frac{(1 + q_1 d)(1 + a_1 d + \ldots + a_{na} d^{na})}{B(1)(1 + q_1)} \; y^*(t) \tag{2.66}$$

From (2.66), for a constant y*(t) the first two values of u(t) are

$$u(0) = \frac{1}{B(1)(1 + q_1)} \; y^*(t) \tag{2.67}$$

$$u(1) = \frac{q_1 + a_1}{B(1)(1 + q_1)} \, y^*(t) \tag{2.68}$$

(2.67) shows that an appropriate choice of $q_1$ reduces the value of $u(0)$. However, this may be accompanied by an increase in the value of $u(t)$ at other sampling instants. A kind of optimum occurs when

$$q_1 = 1 - a_1 \tag{2.69}$$

because then

$$u(0) = u(1)$$

$$= \frac{1}{B(1)(2 - a_1)} \, y^*(t) \tag{2.70}$$

2.4.2 Nondeadbeat response

Choosing

$$G_1(d) = 1, \quad P(d) = 1 + p_1 d + \ldots + p_{np} d^{np} \tag{2.71}$$

gives an " np-th degree " transient response corresponding to a system with poles specified by the polynomial $P(d)$. However, in practice np would be limited to 1 or 2.

For the case np=1, and following an analysis similar to that in Section 2.4.1, the first two values of $u(t)$ would be given by

$$u(0) = \frac{1 + p_1}{B(1)} \tag{2.72}$$

$$u(1) = \frac{1 + p_1}{B(1)(a_1 - p_1)} \tag{2.73}$$

(2.72) shows that an appropriate choice of $p_1$ reduces the value of $u(0)$. This benefit is achieved at the sacrifice of deadbeat response. The more 'negative' $p_1$ is the smaller $u(0)$ will be (up to a point) but also the more sluggish will be the system response. Also, (2.73) shows that this could be accompanied by an increase in the value of $u(t)$ at other sampling instants, with the condition $u(1) = u(0)$ occuring when

$$p_1 = a_1 - 1 \tag{2.74}$$

providing (2.74) corresponds to a stable polynomial $P(d)$.

Further tailoring of the transient response may be achieved by choosing $G_1(d) \neq 1$ on the lines discussed above. For example, choosing

$$G_1(d) = \frac{1 + q_1 d}{1 + q_1} \tag{2.75}$$

$$P(d) = 1 + p_1 d \tag{2.76}$$

gives

$$u(0) = \frac{1 + p_1}{B(1)(1 + q_1)} \tag{2.77}$$

$$u(1) = \frac{1 + p_1}{B(1)(1 + q_1)}(a_1 + q_1 - p_1) \tag{2.78}$$

The condition $u(1) = u(0)$ occurs when

$$q_1 = 1 + p_1 - a_1 \tag{2.79}$$

## 2.5 Review of Multivariable Pole Placement (Prager and Wellstead 1981)

The multivariable pole placement algorithm introduced by Prager and Wellstead (1981) is intended to solve the regulator problem. Here, the reference tracking problem is considered.

Let the plant have a left matrix-fraction description given by

$$A(d)y(t) = B(d)u(t) \tag{2.80}$$

where $u(t)$ and $y(t)$ denote the px1 input vector and px1 output vector, respectively. $A(d)$ and $B(d)$ are pxp polynomial matrices in the unit delay operator $d$ having the forms

$$A(d) = I + A_1 d + \ldots\ldots + A_n d^n$$

$$B(d) = B_1 d + \ldots\ldots + B_m d^m$$

Let the controller be described by

$$u(t) = G(d)F(d)^{-1}[y^*(t) - y(t)] \tag{2.81}$$

where $F(d)$ and $G(d)$ are pxp polynomial matrices given by

$$F(d) = F_0 + F_1 d + \ldots\ldots + F_{nf} d^{nf}$$

$$G(d) = G_0 + G_1 d + \ldots\ldots + G_{ng} d^{ng}$$

and $y^*(t)$ is a p-vector reference signal.

(2.80) and (2.81) imply that the closed-loop system is

$$\begin{aligned} y(t) &= [I + A(d)^{-1}B(d)G(d)F(d)^{-1}]^{-1}A(d)^{-1}B(d)G(d)F(d)^{-1}y^*(t) \\ &= F(d)[A(d)F(d) + B(d)G(d)]^{-1}B(d)G(d)F(d)^{-1}y^*(t) \end{aligned} \tag{2.82}$$

The desired pole placement is achieved by setting

$$A(d)F(d) + B(d)G(d) = P(d) \tag{2.83}$$

where P(d) is an arbitrary stable monic polynomial matrix.

The solution to (2.83) requires the solution of the following set of equations

$$\begin{bmatrix} I & & & & & \\ A_1 & \ddots & & B_1 & & \\ \vdots & & I & \vdots & \ddots & B_1 \\ A_n & & \vdots & B_m & & \vdots \\ & \ddots & & & \ddots & \\ & & A_n & & & B_m \end{bmatrix} \begin{bmatrix} F_0 \\ \vdots \\ F_{nf} \\ G_0 \\ \vdots \\ G_{ng} \end{bmatrix} = \begin{bmatrix} I \\ P_1 \\ \vdots \\ \vdots \\ P_{np} \end{bmatrix} \qquad (2.84)$$

normally with

$$nf = m - 1$$
$$ng = n - 1$$
$$np = n + m - 1$$

For the solution to exist, A and B must be relatively left prime such that the matrix on the left-hand side is nonsingular.

The control law may be implemented in two ways:

(1) The control signal u(t) is determined from

$$|F(d)|\, Iu(t) = G(d) \text{ adjoint } F(d)[y^*(t) - y(t)]$$

(2) Transforming the controller from a right to a left matrix fraction description to obtain

$$u(t) = \tilde{F}(d)^{-1}\, \tilde{G}(d)[y^*(t) - y(t)]$$

## 2.6 Development of Multivariable Pole-zero Placement Algorithm

In this section, the MIMO version of the pole-zero placement technique of Section 2.3 is developed.

Consider the control law of the form

$$F(d)u(t) = G(d)[y^*(t) - y(t)] \qquad (2.85)$$

Let $A_s(d)$ (resp. $B_s(d)$) be any pxp right divisor of A(d) (resp. B(d)) whose zeros correspond to any or all the zeros of A(d) (resp. B(d)) outside the unit circle in the complex d-plane. Further define

$$A_u(d) = A(d)A_s(d)^{-1} \qquad (2.86)$$

$$B_u(d) = B(d)B_s(d)^{-1} \qquad (2.87)$$

such that

$$A(d) = A_u(d)A_s(d) \tag{2.88}$$

$$B(d) = B_u(d)B_s(d) \tag{2.89}$$

represent factorizations of A(d) and B(d) into their unstable and stable parts.

Also, let $F_1(d)$ (resp. $G_1(d)$) be any pxp right divisor of F(d) (resp. G(d)). As there is no practical objection to cancellation of stable parts, define

$$F_1(d) = F(d)B_s(d)^{-1} \tag{2.90}$$

$$G_1(d) = G(d)A_s(d)^{-1} \tag{2.91}$$

so that

$$F(d) = F_1(d)B_s(d) \tag{2.92}$$

$$G(d) = G_1(d)A_s(d) \tag{2.93}$$

By manipulating (2.80) and (2.85), the resulting closed-loop system obtained is given by

$$y(t) = A_s^{-1}(A_u + B_uF_1^{-1}G_1)^{-1}B_uF_1^{-1}G_1A_s\ y^*(t) \tag{2.94}$$

where the argument d has been omitted for brevity sake.

Now let

$$\tilde{G}_1\tilde{F}_1^{-1} = F_1^{-1}G_1 \tag{2.95}$$

represent any relatively right prime factorization of $F_1^{-1}G_1$, a relation which (nonuniquely) defines $\tilde{F}_1$ and $\tilde{G}_1$. Thus, (2.94) becomes

$$\begin{aligned} y(t) &= A_s^{-1}(A_u + B_u\tilde{G}_1\tilde{F}_1^{-1})^{-1}B_u\tilde{G}_1\tilde{F}_1^{-1}A_s\ y^*(t) \\ &= A_s^{-1}\tilde{F}_1(A_u\tilde{F}_1 + B_u\tilde{G}_1)^{-1}B_u\tilde{G}_1\tilde{F}_1^{-1}A_s\ y^*(t) \end{aligned} \tag{2.96}$$

The desired pole placement is achieved by setting

$$A_u\tilde{F}_1 + B_u\tilde{G}_1 = P \tag{2.97}$$

The solution to (2.97) exists if $A_u(d)$ and $B_u(d)$ are relatively left prime.

To simplify the design calculations, the following cases are considered.

### 2.6.1 Special case: All process poles cancelled

Since all stable parts are to be cancelled, set

$$A_s = A \tag{2.98}$$

$$A_u = I \tag{2.99}$$

Using (2.98) and (2.99), (2.94) now becomes

$$\begin{aligned} y(t) &= A^{-1}(I + B_u F_1{}^{-1} G_1)^{-1} B_u F_1{}^{-1} G_1 A\ y^*(t) \\ &= A^{-1} B_u (F_1 + G_1 B_u)^{-1} G_1 A\ y^*(t) \end{aligned} \tag{2.100}$$

The desired pole-placement is achieved by setting

$$F_1 + G_1 B_u = PK \tag{2.101}$$

where K is a constant pxp matrix chosen such that the steady state error is zero for constant $y^*(t)$. With (2.101) and under steady state conditions

$$y(t) = A(1)^{-1} B_u(1) [P(1)K]^{-1} G_1(1) A(1)\ y^*(t) \tag{2.102}$$

Inspection of (2.102) shows that there will be no steady state error if

$$K = P(1)^{-1} G_1(1) B_u(1) \tag{2.103}$$

The zero placement is achieved by appropriate choice of the polynomial matrix $G_1$. Without loss of generality, $G_1$ may be normalized such that

$$G_1(1) = I \tag{2.104}$$

The controller polynomial matrices are now given as

$$G = G_1 A \tag{2.105}$$

$$F_1 = PK - G_1 B_u \tag{2.106}$$

To avoid the decomposition problem, the following case is considered.

### 2.6.2 Special case: No process zeros cancelled

Since no process zeros are to be cancelled, set

$$B_s = I \tag{2.107}$$

$$B_u = B \tag{2.108}$$

The controller polynomial matrices now become

$$G = G_1 A \tag{2.109}$$

$$F = PK - G_1 B \tag{2.110}$$

With (2.101) and (2.108), the closed-loop equation (2.100) now becomes

$$y(t) = A^{-1}B[PK]^{-1}G_1A\ y^*(t) \tag{2.111}$$

and the corresponding control signal is

$$u(t) = [P\ K]^{-1}G_1A\ y^*(t) \tag{2.112}$$

Remark

Special cases of the controller have appeared in the literature before, e.g. Matko and Schumann (1984), Sirisena and Teng (1986).

## 2.7 Alternative Formulation

In the preceding section, it can be seen from (2.112) that the inputs can only be obtained by solving a set of equations. To avoid this additional computation step the following formulation is proposed. Instead of using (2.101), the pole placement can be achieved by setting

$$F + G_1B = P \tag{2.113}$$

where $G_1$ is chosen such that $F(1) = 0$, i.e., such that

$$G_1(1) = Q(1)P(1)B(1)^{-1} \tag{2.114}$$

with

$$Q(1) = I \tag{2.115}$$

This stems from the requirement for zero steady state error with step $y^*(t)$. Here, Q can be treated as the zero placement polynomial matrix.

With (2.113) and (2.114), the control signal is given as

$$u(t) = P^{-1}QP(1)B(1)^{-1}A\ y^*(t) \tag{2.116}$$

Thus, if P is diagonal ( which is usually the case ) each input can be calculated without solving a set of equations, although the finding of the inverse of B(1) is involved.

Remark

In contrast to SISO systems, any one output of a MIMO system can be affected by more than one of its inputs. This feature plays an important role in the design of MIMO systems. A formulation of the decoupling problem is given in the Appendix.

## 2.8 Conclusion

In this chapter, a computationally efficient pole-zero placement algorithm for both SISO and MIMO servo systems has been developed. The pole polynomial P(d) determines the transient response while the zero polynomial $G_1(d)$ specifies additional closed-loop zeros which may modify the control action.

The algorithm has several attractive features: it has an inherent integral action which ensures zero steady-state errors under constant inputs and disturbancess; the controller design step is trivial; it is applicable to nonminimum phase systems, which in the discrete-time context, are common; and it can handle unknown and varying but bounded time delays. However, the algorithm is applicable only to stable plants, although this is not a severe limitation as most practical processes are inherently stable.

# CHAPTER 3

# ADAPTIVE CONTROL OF DETERMINISTIC SYSTEMS

## 3.1 Introduction

In chapter 2, the main concern has been on the control of known systems. In this chapter, the control of systems whose parameters are unknown is considered. By combining a parameter estimation algorithm with the control design algorithm of chapter 2, an adaptive controller is obtained.

Adaptive controllers are generally viewed in a stochastic frame -work. However, the emphasis here is on the deterministic adaptive control problem. The stochastic adaptive control problem is addressed in chapter 4.

There are two desirable properties of an adaptive controller: namely, stability and convergence. By stability, it is meant that bounded system inputs lead to bounded system outputs. By convergence, this is usually taken to mean that the adaptive controller tends asymptotically to the corresponding controller designed on the basis of known plant parameters.

In this chapter, the convergence properties of the pole-zero placement adaptive control algorithm applied to linear, deterministic, time-invariant systems are studied. The systems may have purely deterministic disturbances.

The organization of this chapter is as follows. Section 3.2 presents the fixed pole-zero placement strategy for systems having purely deterministic disturbances. Section 3.3 presents two parameter estimation algorithms and their basic properties. Section 3.4 addresses the convergence properties of the pole-zero placement strategy in an adaptive scheme. Section 3.5 gives some examples illustrating the results of Section 3.4. Section 3.6 presents examples demonstrating the robustness of the adaptive scheme. Section 3.7 addresses the convergence properties of the MIMO version of the pole-zero placement adaptive controller. Section 3.8 presents some simulated examples to illustrate the performance of the MIMO adaptive controller.

## 3.2 Pole-zero Placement Strategy for Known Plants

In this section, the pole-zero placement control of systems having purely deterministic disturbances is addressed.

### 3.2.1 Deterministic disturbances

This section considers a class of deterministic disturbances that can be modelled by a linear finite dimensional state-space model.

A sinusoidal disturbance given by

$$d(t) = A\sin(\omega t + \phi) \tag{3.1}$$

can be modelled by

$$\begin{aligned} d_1(t+1) &= d_2(t) \;;\; d_1(0) = d_1 \\ d_2(t+1) &= (2\cos\omega)d_2(t) - d_1(t) \;;\; d_2(0) = d_2 \\ d(t) &= d_1(t) \end{aligned} \tag{3.2}$$

An appropriate state-space model for the disturbance d(t) is

$$x(t+1) = \begin{bmatrix} 2\cos\omega & -1 \\ 1 & 0 \end{bmatrix} x(t) \tag{3.3}$$

$$d(t) = [\,0 \quad 1\,]\, x(t) \tag{3.4}$$

It is obvious that the model (3.3), (3.4) is uncontrollable, having 2 uncontrollable roots on the unit circle at $\cos\omega \pm j\sin\omega$. But a simple calculation shows that the model is completely observable. Then there exists a similarity transformation that converts the observable model into an observer canonical form, having the following structure:

$$\bar{x}(t+1) = \begin{bmatrix} 2\cos\omega & 1 \\ -1 & 0 \end{bmatrix} \bar{x}(t) \tag{3.5}$$

$$d(t) = [\,1 \quad 0\,]\, \bar{x}(t) \tag{3.6}$$

Using (3.6) in (3.5) gives the following ARMA (auto regressive moving average) model:

$$\left(1 - (2\cos\omega)d + d^2\right)d(t) = 0 \tag{3.7}$$

Example

Consider the following system

$$y(t) = bu(t-1) + A\sin(\omega t + \phi) \tag{3.8}$$

An appropriate state-space model is

$$x(t+1) = \begin{bmatrix} 2\cos\omega & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u(t) \tag{3.9}$$

$$y(t) = [\,1 \quad 0 \quad b\,]\, x(t) \tag{3.10}$$

The model (3.9), (3.10) is uncontrollable but completely observable. Thus the model can be transformed into an observer canonical form:

$$\bar{x}(t+1) = \begin{bmatrix} 2\cos\omega & 1 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \bar{x}(t) + \begin{bmatrix} b \\ -2b\cos\omega \\ b \end{bmatrix} u(t) \tag{3.11}$$

$$y(t) = [\,1 \quad 0 \quad 0\,]\, \bar{x}(t) \tag{3.12}$$

Using (3.12) in (3.11) gives the following ARMA model:

$$\left(1 - (2\cos\omega)d + d^2\right)y(t) = b\left(1 - (2\cos\omega)d + d^2\right)u(t-1) \tag{3.13}$$

Thus, a sinusoidal disturbance discussed above always gives rise to an ARMA model

$$A(d)y(t) = B(d)u(t) \tag{3.14}$$

where

$$A(d) = 1 + a_1 d + \ldots\ldots + a_{na} d^{na}$$

$$B(d) = b_1 d + \ldots\ldots\ldots + b_{nb} d^{nb}$$

in which $A(d)$ and $B(d)$ have common roots on the unit circle.

In general, a purely deterministic disturbance can be modelled as a finite sum of sinusoids, e.g.

$$d(t) = \sum_{i=1}^{l} A_i \sin(\omega_i t + \phi_i) \tag{3.15}$$

(3.15) can be modelled by an observable state space model having $2l$ uncontrollable roots on the unit circle at $\cos\omega_i \pm j\sin\omega_i$. The corresponding ARMA model is given by

$$D(d)d(t) = 0 \tag{3.16}$$

where

$$D(d) = \prod_{i=1}^{l} \left[1 - (2\cos\omega_i)d + d^2\right] \tag{3.17}$$

As seen from the above example, a linear system given by

$$\bar{A}(d)y(t) = \bar{B}(d)u(t) + d(t) \tag{3.18}$$

where d(t) is of the form in (3.15) can be described by an observable but uncontrollable state-space model:

$$x(t+1) = Ax(t) + Bu(t) \tag{3.19}$$

$$y(t) = Cx(t) \tag{3.20}$$

The model (3.19), (3.20) can be transformed into an observer canonical form having a structure of the form

$$\bar{x}(t+1) = \begin{bmatrix} -\alpha_1 & 1 & & 0 \\ \vdots & & \ddots & \\ & & & 1 \\ -\alpha_n & 0 & \cdots & 0 \end{bmatrix} \bar{x}(t) + \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} u(t) \tag{3.21}$$

$$y(t) = [\ 1\ 0\ \ldots\ldots\ 0\ ]\ \bar{x}(t) \tag{3.22}$$

Using (3.22) in (3.21) gives the following ARMA model:

$$A(d)y(t) = B(d)u(t) \tag{3.23}$$

where

$$\begin{aligned} A(d) &= \bar{A}(d)D(d) \\ B(d) &= \bar{B}(d)D(d) \end{aligned} \tag{3.24}$$

Thus, A(d) and B(d) cannot be assumed to be relatively prime if purely deterministic disturbances are present.

### 3.2.2 Pole-zero placement

Consider the system described by

$$A(d)y(t) = B(d)u(t) \tag{3.25}$$

where $A(d)$ and $B(d)$ are not assumed relatively prime and are defined by

$$A(d) = 1 + a_1 d + \dots\dots + a_{na} d^{na}$$

$$B(d) = b_1 d + \dots\dots + b_{nb} d^{nb}.$$

The following result relates to the pole-zero placement control of systems described by (3.25).

Theorem 3.1

Consider the system described by (3.25) and the control law of the form

$$[P(d)K - G_1(d)B(d)]u(t) = G_1(d)A(d)[y^*(t) - y(t)] \tag{3.26}$$

where $y^*(t)$ is the reference signal, and

$$K = \frac{B(1)}{P(1)}$$

(a) The resulting closed-loop system is given by

$$A(d)P(d)Ky(t) = G_1(d)A(d)B(d)y^*(t) \tag{3.27}$$

$$P(d)Ku(t) = G_1(d)A(d)y^*(t) \tag{3.28}$$

(b) The resulting closed-loop system has bounded inputs and bounded outputs if the following conditions are satisfied:

(i) All modes of (3.27) (i.e., the zeros of $A(z^{-1})P(z^{-1})$) lie inside or on the unit circle,

(ii) All controllable modes of (3.27) ( i.e., the poles of the transfer function $\dfrac{G_1(z^{-1})A(z^{-1})B(z^{-1})}{A(z^{-1})P(z^{-1})}$ ) should be inside the unit circle,

(iii) Any modes of (3.27) on the unit circle must have a Jordan block size of 1.

If there are no roots on the unit circle, $A(z^{-1})P(z^{-1})$ should have all roots inside the unit circle.

Proof

(a) Straightforward.

(b) The proof is as follows. It is assumed that the $P(z^{-1})$

has roots inside the unit circle. If purely deterministic disturbances are present, the model (3.25) will have uncontrollable modes on the unit circle of Jordan block size 1, otherwise it has only controllable modes inside the unit circle.

The following Lemma (Appendix B.3.3 of Goodwin and Sin 1984) is required.

Lemma 3.1

Consider the system (of order n, with r inputs and m outputs.)

$$x(t+1) = Ax(t) + Bu(t); \qquad x(0) = x_0$$

$$y(t) = Cx(t) + Du(t)$$

Provided that the following conditions are satisfed:

(i) $|\lambda_i(A)| \leq 1$ ; $i = 1,\dots,n$

(ii) All controllable modes of (A,B) are inside the unit circle

(iii) Any eigenvalues of A on the unit circle have a Jordan block of size 1.

Then

(a) There exists constants $K_1$ and $K_2$ ( $0 < K_1 < \infty$, $0 \leq K_2 < \infty$) which are independent of N such that

$$\sum_{t=1}^{N} ||y(t)||^2 \leq K_1 \sum_{t=0}^{N} ||u(t)||^2 + K_2 \qquad \text{for all } N \geq 0$$

(b) There exists constants $0 \leq m_1 < \infty$, $0 < m_2 < \infty$ which are independent of t such that

$$|y_i(t)| \leq m_1 + m_2 \max_{1 \leq \tau \leq N} ||u(\tau)|| \qquad \text{for all } 1 \leq t \leq N$$
$$i = 1,\dots,m$$

Note that (3.27) and (3.28) are equivalent to observable state-space models. The result then follows from Lemma 3.1 and a bounded sequence $\{y^*(t)\}$.

Remark

The disturbances can be removed from the system output.

## 3.3 Parameter Estimation Algorithms

In this section, the classical least squares parameter estimation algorithm is first described, followed by a modified form of the algorithm. The discussion is based on the following SISO systems of the form

$$A(d)y(t) = B(d)u(t) \tag{3.33}$$

where

$$A(d) = 1 + a_1 d + \ldots\ldots + a_{na} d^{na}$$

$$B(d) = b_1 d + \ldots\ldots\ldots + b_{nb} d^{nb}$$

The system (3.33) to be identified can be written as

$$y(t) = \phi(t-1)^T \theta_0 \tag{3.34}$$

where

$$\phi(t-1) = [\ y(t-1),\ldots,y(t-na),u(t-1),\ldots,u(t-nb)\ ]^T$$

$$\theta_0 = [\ -a_1,\ldots,-a_{na},b_1,\ldots,b_{nb}\ ]^T$$

### 3.3.1 Least squares

The least squares parameter estimation algorithm is described as follows:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{P(t-2)\phi(t-1)}{1 + \phi(t-1)^T P(t-2)\phi(t-1)}\left(y(t) - \phi(t-1)^T \hat{\theta}(t-1)\right) \tag{3.35}$$

$$P(t-1) = P(t-2) - \frac{P(t-2)\phi(t-1)\phi(t-1)^T P(t-2)}{1 + \phi(t-1)^T P(t-2)\phi(t-1)} \tag{3.36}$$

with $\hat{\theta}(0)$ given and $P(-1)$ is any positive definite matrix $P_0$.

Let

$$e(t) = y(t) - \phi(t-1)^T \hat{\theta}(t-1) \tag{3.37}$$

The basic convergence properties of the above algorithm are summarized in the following lemma.

Lemma 3.2

For the algorithm (3.35) and (3.36) and subject to (3.34), it follows that

(1) $||\hat{\theta}(t) - \theta_0||^2 \leq \kappa_1 ||\hat{\theta}(0) - \theta_0||^2$ ; $t \geq 1$

where

$\kappa_1$ = condition number of $(P(-1)^{-1})$

(2) $$\lim_{t \to \infty} \frac{e(t)^2}{1 + \kappa_2 \phi(t-1)^T \phi(t-1)} = 0$$

where

$\kappa_2 = \lambda_{max}(P(-1))$

$\lambda_{max}$ denotes the maximum eigenvalue.

(3) $$\lim_{t \to \infty} ||\hat{\theta}(t) - \hat{\theta}(t-k)|| = 0 \quad \text{for any finite } k$$

Proof

See Goodwin and Sin (1984).

3.3.2 Constrained least squares (p. 92, Goodwin and Sin 1984)

In this algorithm the parameter estimates are constrained to lie within a closed convex region in parameter space. The algorithm is described as follows:

$$\hat{\theta}'(t) = \hat{\theta}(t-1) + \frac{P(t-2)\phi(t-1)}{1 + \phi(t-1)^T P(t-2)\phi(t-1)} [y(t) - \phi(t-1)^T \hat{\theta}(t-1)] \quad (3.38)$$

$$P(t-1) = P(t-2) - \frac{P(t-2)\phi(t-1)\phi(t-1)^T P(t-2)}{1 + \phi(t-1)^T P(t-2)\phi(t-1)} \quad (3.39)$$

The estimated parameter $\hat{\theta}'(t)$ is modified according to the following projection facility:

$$\hat{\theta}(t) = \begin{cases} \hat{\theta}'(t), & \text{if } \hat{\theta}'(t) \in C \\ \\ \hat{\theta}^*(t), & \text{if } \hat{\theta}'(t) \notin C \end{cases} \quad (3.40)$$

where C is the defined closed-convex region.

The computation of $\hat{\theta}^*(t)$ is given as follows.

Since $P(t-1)^{-1}$ is a symmetric, positive definite matrix there exists at least a $P(t-1)^{-1/2}$ such that

$$P(t-1)^{-1} = P(t-1)^{-T/2}P(t-1)^{-1/2} \tag{3.41}$$

and denote the image of C under the linear transformation $P(t-1)^{-1/2}$ by C*. Then C* is also a closed convex region. Under $P(t-1)^{-1/2}$, the image of $\hat{\theta}'(t)$ is given by

$$\hat{\rho}'(t) = P(t-1)^{-1/2}\,\hat{\theta}'(t) \tag{3.42}$$

Also the image of $\theta_0$ under such transformation is

$$\rho_0 = P(t-1)^{-1/2}\theta_0 \tag{3.43}$$

Now $\hat{\theta}^*(t)$ can be found by projecting $\hat{\rho}'(t)$ orthogonally onto the surface of C*. Define

$$\hat{\theta}^*(t) = P(t-1)^{1/2}\,\hat{\rho}^*(t) \tag{3.44}$$

The computation of $\hat{\rho}^*(t)$ is particularly easy in special cases, e.g., when the constrained region is defined by hyperplanes. Then, the projection algorithm of Goodwin and Sin (1984) can be applied.

The above discussion is illustrated by a simple example given below.

Example

Consider a first order system given by

$$A(d)y(t) = B(d)u(t) \tag{3.45}$$

where

$$A(d) = 1 + ad$$

$$B(d) = b_1d + b_2d^2$$

For some reason, the estimated model has to remain stable. Thus the closed convex set in parameter space is given by

$$C = \{\ \theta:\ \theta_1 \leqq \alpha\ ,\ |\alpha| < 1\ \} \tag{3.46}$$

The constrained boundary in the original space is

$$v^T\theta = \alpha \tag{3.47}$$

where

$$v = [\ 1 \quad 0 \quad 0\ ]^T$$

The corresponding constrained boundary in the transformed space is

$$v^T P(t-1)^{1/2} \rho = \alpha \tag{3.48}$$

For a given $\hat{\theta}'(t)$, the corresponding $\hat{\rho}'(t)$ is projected orthogonally onto the boundary of $C^*$ to yield $\hat{\rho}^*(t)$.

The projection algorithm is a consequence of the following optimization problem. Given $\hat{\rho}'(t)$ and $\alpha$, find $\hat{\rho}^*(t)$ such that

$$I = \frac{1}{2}||\hat{\rho}^*(t) - \hat{\rho}'(t)||^2 \tag{3.49}$$

is minimized subject to

$$\alpha = v^T P(t-1)^{1/2}\, \hat{\rho}^*(t) \tag{3.50}$$

Using a Lagrange multiplier for (3.50)

$$I = \frac{1}{2}||\hat{\rho}^*(t) - \hat{\rho}'(t)||^2 + \lambda[\alpha - v^T P(t-1)^{1/2}\, \hat{\rho}^*(t)] \tag{3.51}$$

Hence the necessary conditions for a minimum are

$$\frac{\partial I}{\partial \hat{\rho}^*(t)} = 0 \tag{3.52}$$

$$\frac{\partial I}{\partial \lambda} = 0 \tag{3.53}$$

These give

$$\hat{\rho}^*(t) - \hat{\rho}'(t) - \lambda P(t-1)^{T/2} v = 0 \tag{3.54}$$

$$\alpha - v^T P(t-1)^{1/2}\, \hat{\rho}^*(t) = 0 \tag{3.55}$$

Substituting (3.55) into (3.54) and rearranging gives

$$\lambda = \frac{\alpha - v^T P(t-1)^{1/2}\, \hat{\rho}^*(t)}{v^T P(t-1) v} \tag{3.56}$$

Substituting (3.56) into (3.54) gives

$$\hat{\rho}^*(t) = \hat{\rho}'(t) + \frac{P(t-1)^{1/2} v}{v^T P(t-1) v}[\alpha - v^T P(t-1)^{1/2}\, \hat{\rho}'(t)] \tag{3.57}$$

Thus using (3.42) and (3.44) lead to

$$\hat{\theta}^*(t) = \hat{\theta}'(t) + \frac{P(t-1) v}{v^T P(t-1) v}[\alpha - v^T\, \hat{\theta}'(t)] \tag{3.58}$$

The key step in extending the convergence proof of the least squares estimation algorithm to the constrained form is to note that when projection is used

$$\begin{aligned} V(t) &= (\hat{\theta}'(t) - \theta_0)^T P(t-1)^{-1} (\hat{\theta}'(t) - \theta_0) \\ &= (\hat{\rho}'(t) - \rho_0)^T (\hat{\rho}'(t) - \rho_0) \\ &\geqq (\hat{\rho}^*(t) - \rho_0)^T (\hat{\rho}^*(t) - \rho_0) \end{aligned} \tag{3.59}$$

Hence,

$$(\hat{\theta}^*(t) - \theta_0)^T P(t-1)^{-1} (\hat{\theta}^*(t) - \theta_0) \leqq (\hat{\theta}'(t) - \theta_0)^T P(t-1)^{-1} (\hat{\theta}'(t) - \theta_0) \tag{3.60}$$

By the usual Lyapunov-type argument, all the properties of the least squares estimation algorithm are retained as in Lemma 3.2.

## 3.4 Pole-zero Placement Adaptive Control

In this section, the properties of the pole-zero placement adaptive control algorithm are studied.

Consider the adaptive control of a linear, time-invariant SISO system described by

$$A(d)y(t) = B(d)u(t) \tag{3.61}$$

where

$$A(d) = 1 + a_1 d + \ldots\ldots + a_{na} d^{na}$$

$$B(d) = b_1 d + \ldots\ldots\ldots + b_{nb} d^{nb}$$

(3.61) can be written as

$$y(t) = \phi(t-1)^T \theta_0 \tag{3.62}$$

where

$$\phi(t-1) = [\ y(t-1), \ldots, y(t-n), u(t-1), \ldots, u(t-m)\ ]^T$$

$$\theta_0 = [\ -a_1, \ldots\ldots, -a_{na}, 0, \ldots, 0, b_1, \ldots, b_{nb}, 0, \ldots, 0\ ]^T$$

where

$$na \leqq n\ , \quad nb \leqq m$$

The situation is that the coefficients in $A(d)$ and $B(d)$ are unknown and only the input $u(t)$ and output $y(t)$ can be directly measured. The problem is to determine a control law such that $u(t)$ and $y(t)$ remain bounded for all time and that the desired closed-loop polynomial approaches $P(d)$, for a given setpoint sequence $\{y^*(t)\}$.

The following assumptions about the system are made:

Assumption 3A

(1) Upper bounds of na and nb are known

(2) $B(1) \neq 0$

(3) (i) All modes of (3.61) (i.e., the zeros of $A(z^{-1})$) lie inside or on the unit circle,

(ii) All controllable modes of (3.61) (i.e. the poles of the transfer function $B(z^{-1})/A(z^{-1})$) lie inside the unit circle,

(iii) Any modes of (3.61) on the unit circle have a Jordan block size of 1.

The adaptation algorithm is described as follows:

$$\hat{\theta}'(t) = \hat{\theta}(t-1) + \frac{P(t-2)\phi(t-1)}{1 + \phi(t-1)^T P(t-2)\phi(t-1)}\left[y(t) - \phi(t-1)^T\,\hat{\theta}(t-1)\right] \qquad (3.63)$$

$$P(t-1) = P(t-2) - \frac{P(t-2)\phi(t-1)\phi(t-1)^T P(t-2)}{1 + \phi(t-1)^T P(t-2)\phi(t-1)} \qquad (3.64)$$

with $P(-1)$ any positive definite matrix. The symbols in (3.63) have the following meaning:

$$\hat{\theta}(t) = [\,-\hat{a}_1(t),\ldots,-\hat{a}_n(t),\hat{b}_1(t),\ldots,\hat{b}_m(t)\,]^T$$

$$\hat{A}(t,d) = 1 + \hat{a}_1(t)d + \ldots\ldots + \hat{a}_n(t)d^n$$

$$\hat{B}(t,d) = \hat{b}_1(t)d + \ldots\ldots + \hat{b}_m(t)d^m$$

The estimated parameter $\hat{\theta}'(t)$ is modified according to the following projection facility (see section 3.3):

$$\hat{\theta}(t) = \begin{cases} \hat{\theta}'(t), & \text{if } \hat{\theta}'(t) \in C \\ \\ \hat{\theta}^*(t), & \text{if } \hat{\theta}'(t) \notin C \end{cases} \qquad (3.65)$$

where C is a closed convex set in parameter space satisfying:

(1) $\theta_0 \varepsilon C$

(2) $C \subset \{ \theta(t): \hat{\rho}_i(t) = 1 - \rho < 1, \quad i = 1,..,n$ (3.66)

$\hat{\rho}_i(t)$ are the roots of $\hat{A}(t,d) \}$

The input $\{u(t)\}$ is determined from the control law

$$\hat{F}(t,d)u(t) = \hat{G}(t,d)[y^*(t) - y(t)] \tag{3.67}$$

where

$$\hat{G}(t,d) = G_1(d)\hat{A}(t,d)$$

$$\hat{F}(t,d) = P(d)\hat{K}(t) - G_1(d)\hat{B}(t,d)$$

$$\hat{K}(t) = \begin{cases} \dfrac{\hat{B}(t,1)}{P(1)} & \text{if } |\hat{B}(t,1)| > \varepsilon \\ \\ \dfrac{\varepsilon}{P(1)} & \text{otherwise} \end{cases} \tag{3.68}$$

$\varepsilon$ is an arbitrary positive value.

Remark

The purpose of the projection facility in (3.65) is to ensure the stability of the sequence of estimated models.

(3.68) is a simple scheme to prevent division by zero in finding the control signal u(t).

The following additional assumptions are made:

Assumption 3B

(1) P(d) is an arbitrary stable polynomial

(2) $|y^*(t)| < M_1 < \infty$

The convergence properties of the algorithm (3.63) - (3.68) are summarized in the following theorem.

Theorem 3.2

Subject to Assumptions 3A - 3B, the algorithm (3.63) - (3.68) leads to

(1) $\{u(t)\}$ and $\{y(t)\}$ are bounded sequences

(2) The closed-loop characteristic polynomial tends to $\hat{A}(\infty,d)P(d)$ in the sense that

$$\lim_{t \to \infty} \left[\hat{A}(\infty,d)P(d)\hat{K}(\infty)y(t) - \hat{G}(\infty,d)\hat{B}(\infty,d)y^*(t)\right] = 0$$

(3) $$\lim_{t \to \infty} \left[P(d)\hat{K}(\infty)y(t) - G_1(d)\hat{B}(\infty,d)y^*(t)\right] = 0$$

since $\hat{A}(\infty,d)$ is stable.

Proof

The proof follows closely the standard proof paradigm in Goodwin and Sin (1984). The modifications needed stem from the fact that unlike conventional algorithms, the desired closed-loop polynomial is time-varying.

First, the notation. Given time-varying polynomial operators $\hat{A}(t,d)$, $\hat{B}(t,d)$ define the following:

$$\hat{A}\hat{B} = \sum_{i}\sum_{j} \hat{a}_i(t)\hat{b}_j(t)d^{i+j} \tag{3.69}$$

$$\hat{A}*\hat{B} = \sum_{i}\sum_{j} \hat{a}_i(t)\hat{b}_j(t-i)d^{i+j} \tag{3.70}$$

Note that $\hat{A}\hat{B} = \hat{A}*\hat{B} = \hat{B}*\hat{A}$ when A and B are time-invariant. Also define

$$\hat{B}'' = \hat{B}(t-1,d) \tag{3.71}$$

The key equations required are

$$Ay(t) = Bu(t) \qquad [\text{from } (3.61)] \tag{3.72}$$

$$\hat{F}u(t) = \hat{G}y^*(t) - \hat{G}y(t) \qquad [\text{from } (3.67)] \tag{3.73}$$

$$\hat{F} + G_1\hat{B} = P\hat{K}(t) \qquad [\text{from } (3.67)] \tag{3.74}$$

$$\begin{aligned} e(t) &= y(t) - \phi(t-1)^T\,\hat{\theta}(t-1) \\ &= y(t) - \left[(1 - \hat{A}'')y(t) + \hat{B}''u(t)\right] \\ &= \hat{A}''y(t) - \hat{B}''u(t) \end{aligned} \tag{3.75}$$

Now define

$$\hat{A}^*\hat{G}y^*(t) = \hat{A}^*\hat{F}u(t) + \hat{A}^*\hat{G}y(t)$$

$$= \hat{A}\hat{F}u(t) + [\hat{A}^*\hat{F} - \hat{A}\hat{F}]u(t) + \hat{A}\hat{G}y(t) + [\hat{A}^*\hat{G} - \hat{A}\hat{G}]y(t)$$

$$= \hat{A}\hat{F}u(t) + \hat{G}\hat{B}u(t) + \hat{G}e(t) + [\hat{A}^*\hat{F} - \hat{A}\hat{F}]u(t)$$

$$+ [\hat{A}^*\hat{G} - \hat{A}\hat{G}]y(t) + [\hat{G}^*\hat{B}'' - \hat{G}\hat{B}]u(t) - [\hat{G}^*\hat{A}'' - \hat{G}\hat{A}]y(t)$$

using (3.75)

$$= \hat{A}\hat{P}Ku(t) + \hat{G}e(t) + \left([\hat{A}^*\hat{F} - \hat{A}\hat{F}] + [\hat{G}^*\hat{B}'' - \hat{G}\hat{B}]\right)u(t)$$

$$+ \left([\hat{A}^*\hat{G} - \hat{A}\hat{G}] - [\hat{G}^*\hat{A}'' - \hat{G}\hat{A}]\right)y(t) \quad \text{using (3.74)} \tag{3.76}$$

Similarly, arguing as above

$$\hat{B}^*\hat{G}y^*(t) = \hat{A}\hat{P}Ky(t) - \hat{F}e(t) + \left([\hat{B}^*\hat{F} - \hat{B}\hat{F}] - [\hat{F}^*\hat{B}'' - \hat{F}\hat{B}]\right)u(t)$$

$$+ \left([\hat{B}^*\hat{G} - \hat{B}\hat{G}] + [\hat{F}^*\hat{A}'' - \hat{F}\hat{A}]\right)y(t) \tag{3.77}$$

Define

$$\hat{V}1 = [\hat{A}^*\hat{F} - \hat{A}\hat{F}] + [\hat{G}^*\hat{B}'' - \hat{G}\hat{B}]$$

$$\hat{V}2 = [\hat{A}^*\hat{G} - \hat{A}\hat{G}] - [\hat{G}^*\hat{A}'' - \hat{G}\hat{A}]$$

$$\hat{V}3 = [\hat{B}^*\hat{F} - \hat{B}\hat{F}] - [\hat{F}^*\hat{B}'' - \hat{F}\hat{B}]$$

$$\hat{V}4 = [\hat{B}^*\hat{G} - \hat{B}\hat{G}] + [\hat{F}^*\hat{A}'' - \hat{F}\hat{A}]$$

Combining (3.76) and (3.77) yields

$$\begin{bmatrix} \hat{A}\hat{P}K + \hat{V}1 & \hat{V}2 \\ \hat{V}3 & \hat{A}\hat{P}K + \hat{V}4 \end{bmatrix} \begin{bmatrix} u(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} \hat{A}^*\hat{G} \\ \hat{B}^*\hat{G} \end{bmatrix} y^*(t) + \begin{bmatrix} -\hat{G} \\ \hat{F} \end{bmatrix} e(t) \tag{3.78}$$

(3.78) can be considered as a linear time-varying dynamical system with inputs filtered versions of $\{y^*(t)\}$ and $\{e(t)\}$, and outputs $\{u(t)\}$ and $\{y(t)\}$. Now it follows from Lemma 3.2 that $\hat{A}(t,d)$ and $\hat{B}(t,d)$ have bounded coefficients and converge. Hence, for a sufficiently large but finite t, the system of (3.78) is arbitrarily close to an asymptotically exponentially stable system having characteristic polynomial

$$[\hat{A}(t,z)P(z)]^2$$

It also follows from (3.78) that $u(t)$ and $y(t)$ (i.e. $||\phi(t)||$) will not grow faster than linearily with respect to $e(t)$.

The following key technical lemma due to Goodwin and Sin 1984 is needed.

If

$$\lim_{t \to \infty} \frac{e(t)^2}{c_1 + c_2\ \phi(t)^T \phi(t)} = 0$$

where $0 < c_1 < \infty$, $0 < c_2 < \infty$, and $\{e(t)\}$ is a real scalar sequence and $\{\phi(t)\}$ is a p-vector sequence; then subject to

$$||\phi(t)|| \leq C_1 + C_2 \max_{0 \leq \tau \leq t} |e(\tau)|$$

where $0 \leq C_1 < \infty$, $0 \leq C_2 < \infty$, it follows that

$$\lim_{t \to \infty} e(t) = 0$$

and $\{||\phi(t)||\}$ is bounded.

Thus, applying the above lemma it can be concluded from Lemma 3.2-(2) that $e(t) \to 0$ as $t \to \infty$ and $||\phi(t)||$ is bounded. Hence, from the definition of $\phi(t)$, it follows that $\{y(t)\}$ and $\{u(t)\}$ are bounded. This establishes part (1) of the theorem.

Part (2) of the theorem is immediate from (3.77).

Finally, part (3) of the theorem follows from part (2).

Remark

It can be seen from Theorem 3.1-(3) that, for a constant $y^*(t)$, there will be no steady state errors in the output sequence if $\hat{K}(t)$, defined by (3.68), converges to $\hat{B}(\infty,1)/P(1)$.

When considering the adaptive control of systems having no disturbances the only change required is to reduce the dimension of the parameter vector.

## 3.5 Examples

This section presents simulated examples showing the performances of the adaptive schemes treated in Section 3.4.

### Example 1

Consider the deterministic system originally discussed in Clark (1984). It is described by

$$(1 - 0.7d)y(t) = (1 + 2d)u(t-1) \tag{3.79}$$

Note that this is a nonminimum phase system with an unstable zero at -2.0.

The following specifications were used:

$$P(d) = 1 - 0.2d, \quad G_1(d) = 1$$

The true value of the parameter vector, $\theta_0$, and the initial parameter estimates, $\hat{\theta}(0)$ were as follows:

| parameters | $a_1$ | $b_1$ | $b_2$ |
|---|---|---|---|
| true value | -0.7 | 1.0 | 2.0 |
| initial value | -0.5 | 0.8 | 1.6 |

The reference input was taken to be a square wave of amplitude 10 and period of 100 samples.

Fig. 3.1 shows the performance of the ordinary least squares adaptive control algorithm. The output and reference input, input, and parameter estimates are shown in Fig. 3.1(a), (b), and (c), respectively.

### Example 2

Fig. 3.2 shows the output and input of the plant of Example 1 under least squares adaptive control with the following specifications:

$$P(d) = 1 - 0.2d,$$

$$G_1(d) = \frac{1 + (1 + p_1 - \hat{a}_1)d}{1 + (1 + p_1 - \hat{a}_1)}$$

It is clear that the zero-polynomial $G_1(d)$ modifies the control signal. In comparison with the control signal magnitude in Fig. 3.1(b) the one in Fig. 3.2(b) is smaller.

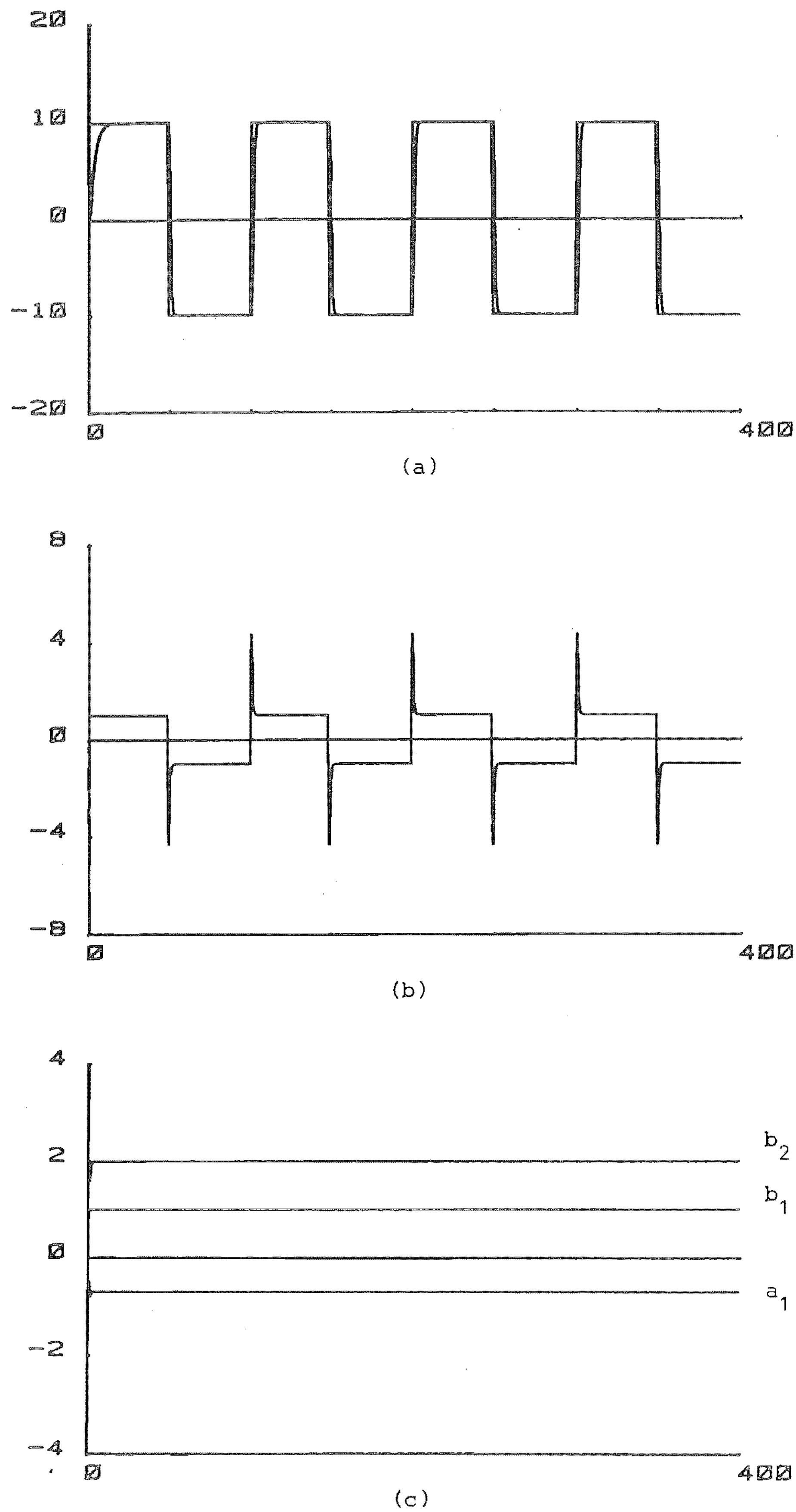



Fig. 3.1 Example 1 (a) output; (b) input; (c) estimated parameters.

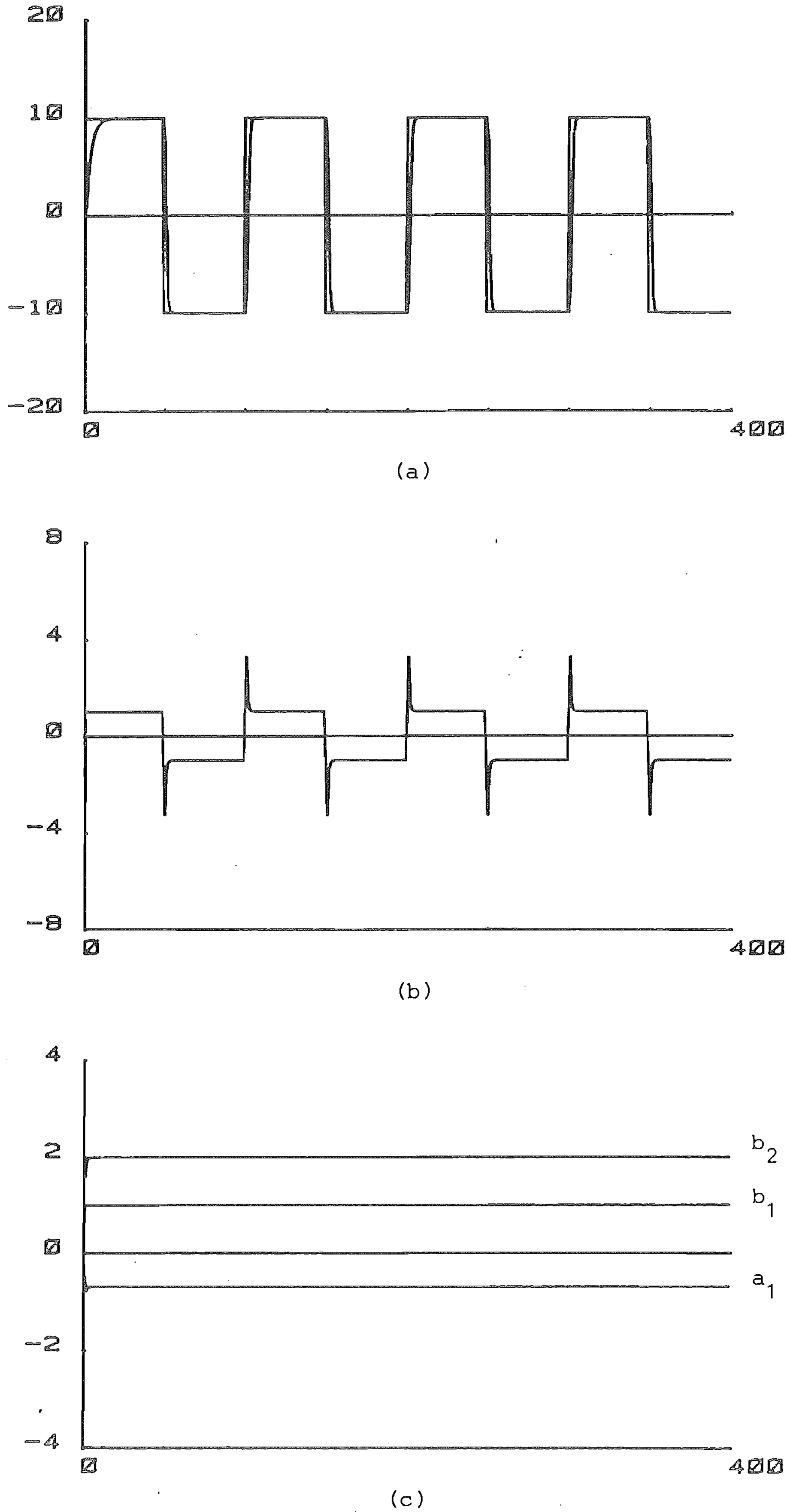



Fig. 3.2 Example 2 (a) output; (b) input; (c) estimated parameters

Example 3

Supposing the output of (3.79) is corrupted by a purely deterministic disturbance given by

$$d(t) = \sin(0.5t + \phi) \tag{3.80}$$

Then the autoregressive moving average model for the system is

$$(1 - 1.7d + 1.7d^2 - 0.7d^3)y(t) = (1 + d - d^2 + 2d^3)u(t-1) \tag{3.81}$$

The true value of the parameter vector, $\theta_0$, and the initial parameter estimates, $\hat{\theta}(0)$ were as follows:

| parameters | $a_1$ | $a_2$ | $a_3$ | $b_1$ | $b_2$ | $b_3$ | $b_4$ |
|---|---|---|---|---|---|---|---|
| true value | -1.7 | 1.7 | -0.7 | 1.0 | 1.0 | -1.0 | 2.0 |
| initial value | -1.5 | 1.5 | -0.8 | 1.1 | 1.0 | -1.2 | 2.0 |

Fig. 3.3(a) and (b) show the output and input respectively of the system (3.81) under least squares adaptive control with the same specifications as in Example 1. The parameter estimates are shown in Fig. 3.3(c). Note that in spite of the disturbance, excellent tracking was obtained. The uncontrollable modes were rendered unobservable at the output by feedback control.

## 3.6 Robustness: Examples

In this section, the robustness of the adaptive control scheme are illustrated by simulated examples.

Example 1

In this example, the robustness of the adaptive scheme to a mild violation of the system stability assumption is illustrated.

Consider the adaptive control of a 'slighlty' unstable process whose continuous-time model is described by

$$G(s) = \frac{-0.5\ e^{-0.1s}}{s - 0.25} \tag{3.82}$$

Using an input zero-order hold with a sampling period of 0.2 sec the discrete-time system obtained is

$$G(d) = \frac{-0.05d - 0.052d^2}{1 - 1.05d} \tag{3.83}$$

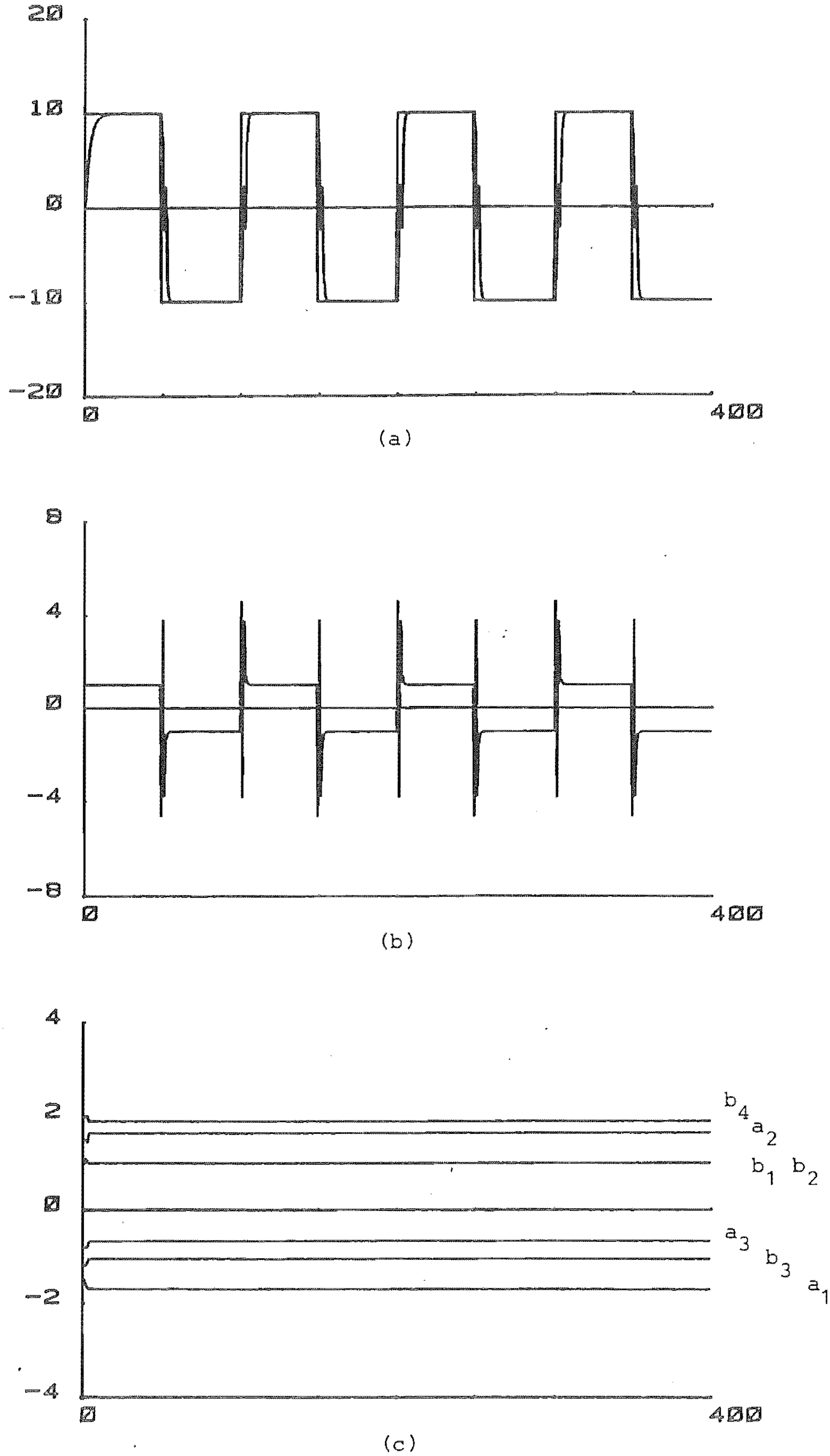



Fig. 3.3 Example 3 (a) output; (b) input; (c) estimated parameters

The reference input was taken to be a square wave of amplitude 10 and period of 200 samples.

The following specifications were used:

$$P(d) = 1 - 0.7d, \quad G_1(d) = 1$$

The performance of the constrained least squares adaptive control algorithm with $|\hat{a}_1| \leq 0.95$ is shown in Fig. 3.4. Fig. 3.4(a), (b), and (c) show the output and reference input, input, and parameter estimates, respectively. Despite the violation of system stability assumption, the adaptive controller remained stable. However, when the sampling period was increased, simulations showed that the closed-loop system became unstable. It can also be seen that the robustness was achieved at the sacrifice of smaller input magnitude.

Example 2

In this example, the robustness of the adaptive scheme with respect to unmodelled dynamics is illustrated.

Consider the system originally discussed in Clark (1984) which is described by

$$(1 - 1.7d + 0.72d^2)y(t) = (0.1 + 0.2d)u(t-1) \tag{3.84}$$

Note that this is a nonminimum phase plant with an unstable zero at -2.0. The system has stable poles at 0.8 and 0.9.

A first order model was used in the adaptive control algorithm and it is described by

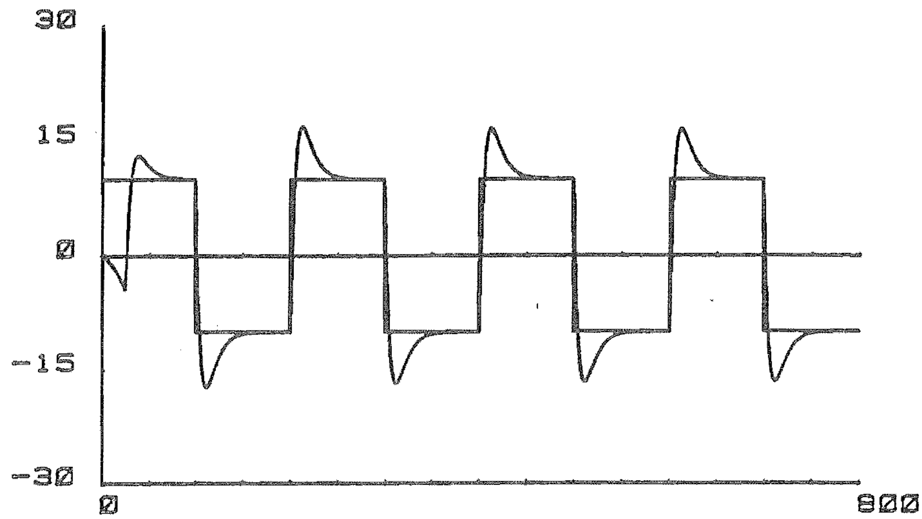
$$(1 + ad)y(t) = (b_1 + b_2d^2)u(t-1) \tag{3.85}$$

The reference input sequence was taken to be a square wave of amplitude 10 and period of 200 samples.

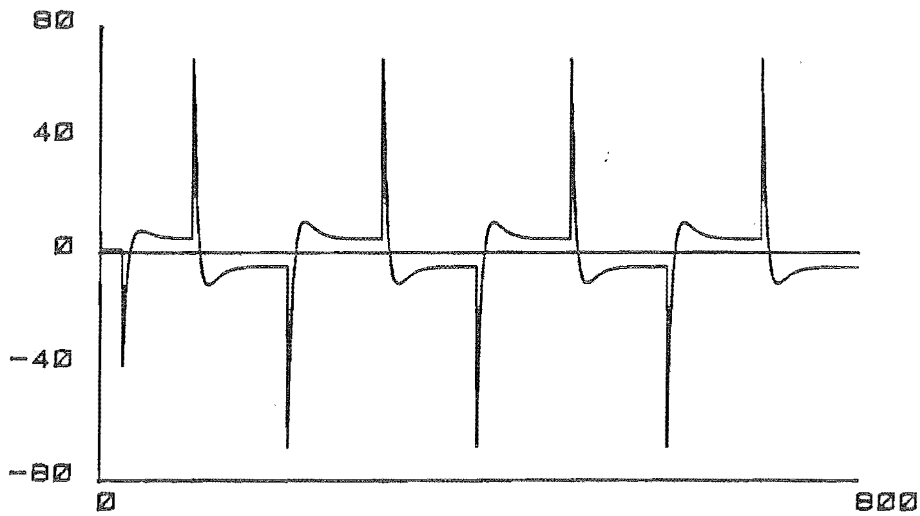
The following specifications were used:

$$P(d) = 1 - 0.7d, \quad G_1(d) = 1$$

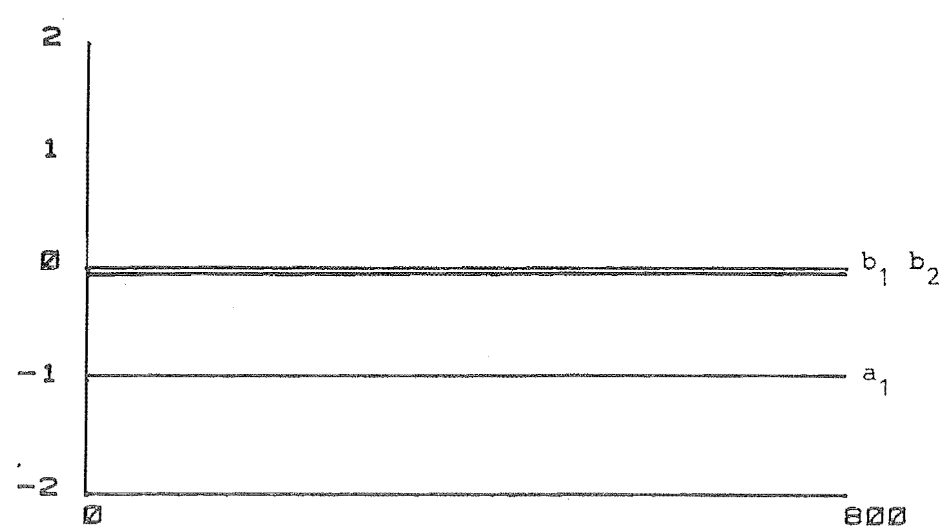
Fig. 3.5 shows the performance of the constrained least squares adaptive control algorithm using $|\hat{a}_1| \leq 0.95$. The output and reference input, input, and parameter estimates are shown in Fig. 3.5-(a) and (b), respectively. Clearly, the adaptive controller is robustly stable. In contrast, Clark (1984) has shown that the classical pole placement adaptive controller is not robust when the plant (3.84) is modelled by the first order model (3.85).




(a)

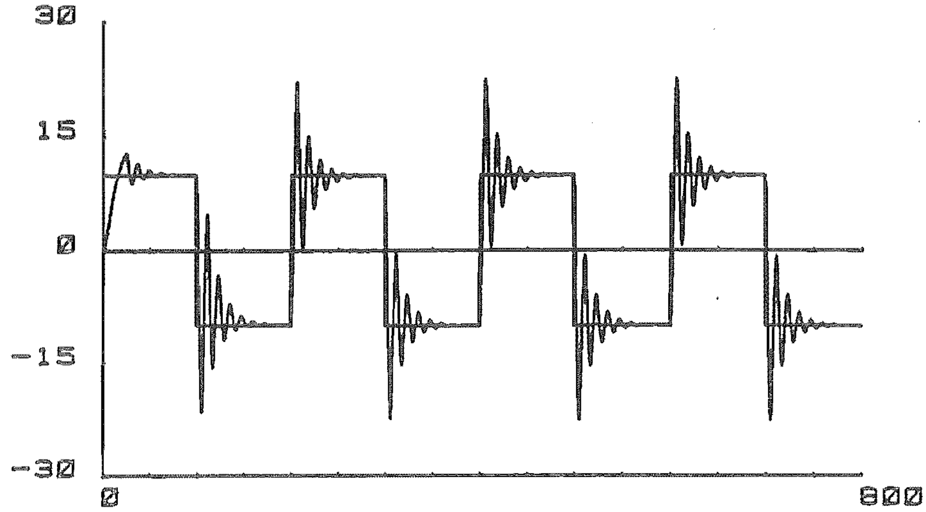



(b)

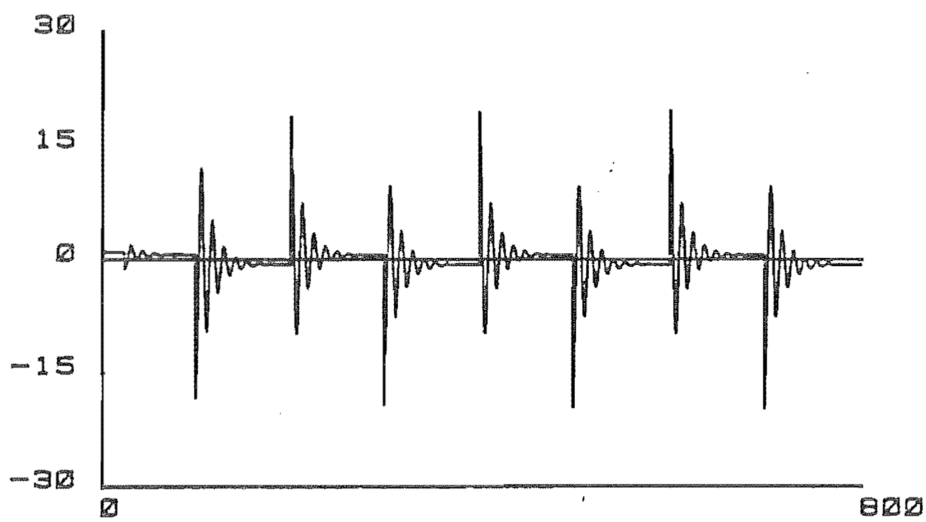



(c)

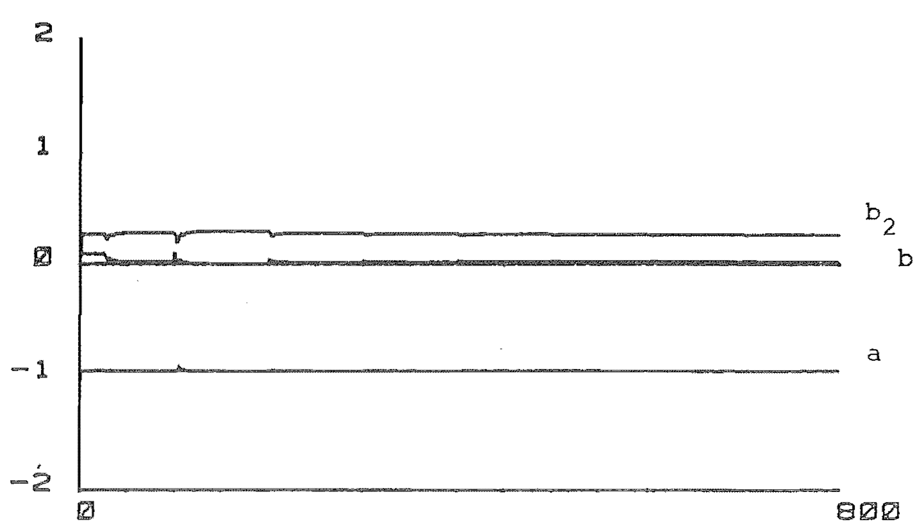
Fig. 3.4 Example 1 (a) output; (b) input;
(c) estimated parameters




(a)




(b)




(c)

Fig. 3.5 Example 2 (a) output; (b) input;
(c) estimated parameters

## 3.7 Multi-input Multi-output Systems

The results of Section 3.4 for SISO systems can be easily extended to MIMO systems. In this section, convergence properties for the MIMO adaptive pole-zero placement algorithm are studied.

### 3.7.1 Problem statement

Consider the adaptive control of a square MIMO process described by

$$A(d)y(t) = B(d)u(t) \tag{3.86}$$

where the p-vectors $y(t)$ and $u(t)$ are the output and input, respectively. $A(d)$ and $B(d)$ are the pxp polynomial matrices in the unit delay operator d having the forms

$$A(d) = I + A_1 d + \ldots\ldots + A_{na} d^{na}$$

$$B(d) = B_1 d + \ldots\ldots + B_{nb} d^{nb}$$

It is assumed that pure time delays are accounted for by the leading zero coefficients in $B(d)$.

The situation is that the coefficient matrices in $A(d)$ and $B(d)$ are unknown and that only $u(t)$ and $y(t)$ can be directly measured. The problem is to determine a feedback control law which causes $u(t)$ and $y(t)$ to remain bounded for all time, and that the tracking error asymptotically converges to zero for a given setpoint sequence $\{y^*(t)\}$.

The following system assumptions are made:

Assumption 3C

(1) $\det[A(z^{-1})] \neq 0$ for $|z| \geq 1$

(2) $\det[\sum_{i=1}^{m} B_i(1)] \neq 0$

(3) upper bounds of na and nb are known

A more general condition including uncontrollable roots on the unit circle follows from Assumption 3A.

### 3.7.2 The Adaptation Algorithm

The system (3.86) to be identified can be written as

$$y(t) = \theta_0^T \phi(t-1) \tag{3.87}$$

where the (m+n)x1 column vector $\phi(t)$ is defined as

$$\phi(t-1) = [\ y(t-1)^T,\ldots,y(t-n)^T,u(t-1)^T,\ldots,u(t-m)^T\ ]^T$$

$\theta_0$ is an (m+n)xp matrix consisting of the coefficients of A(d) and B(d)

where

$$na \leqq n, \quad nb \leqq m$$

Estimates at time t of $\theta_0$, A(d), and B(d) are denoted by $\hat{\theta}(t)$, $\hat{A}(t,d)$, and $\hat{B}(t,d)$, respectively.

The adaptation algorithm is described as follows:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{a(t-1)P(t-2)\phi(t-1)}{1 + a(t-1)\phi(t-1)^T P(t-2)\phi(t-1)}[y(t)^T - \phi(t-1)^T\ \hat{\theta}(t-1)] \tag{3.88}$$

$$P(t-1) = P(t-2) - \frac{a(t-1)P(t-2)\phi(t-1)\phi(t-1)^T P(t-2)}{1 + a(t-1)\phi(t-1)^T P(t-2)\phi(t-1)} \tag{3.89}$$

with P(-1) any positive definite matrix, and

$$0 \leqq a(t) \leqq 2$$

The control signal is determined from

$$[P(d)\hat{K}(t) - G_1(d)\hat{B}(t,d)]u(t) = G_1(d)\hat{A}(t,d)[y^*(t) - y(t)] \tag{3.90}$$

where

$$\hat{K}(t) = P(1)^{-1}\ \hat{B}(t,1) \tag{3.91}$$

To satisfy the following conditions:

$$\text{C1} \quad \det[\sum_{i=0}^{n} \hat{A}_i(t,d)] \neq 0 \ \text{ for } t \to \infty \ ; \quad \hat{A}_0(t,d) = I$$

$$\text{C2} \quad \det[\sum_{i=1}^{m} \hat{B}_i(t,1)] \neq 0 \ \text{ for all } t$$

appropriate choice of a(t) can be chosen. One possibility due to Goodwin et al (1980) is to choose a(t) such that

$$a(t) = \begin{cases} 1 & \text{if C1 and C2 are met} \\ \gamma & \text{otherwise, where } 2 > \gamma > 1, \text{ or } 1 > \gamma > 0 \end{cases}$$

Let

$$b(t) = y(t) - \hat{\theta}(t)^T \phi(t-1) \tag{3.92}$$

The identification algorithm (3.88) to (3.89) has the following elementary properties:

Lemma 3.3

For the algorithm (3.88) to (3.89) and subject to (3.87), it follows that

(1) $b(t)$ is bounded, i.e.,

$$\lim_{t \to \infty} b(t) = 0$$

(2) $$\lim_{t \to \infty} ||\hat{\theta}(t) - \hat{\theta}(t-1)|| = 0$$

where

$$||M||^2 = \text{trace } (M^T M)$$

(3) $\hat{\theta}(t)$ is bounded and converges to constant values, i.e.

$$\lim_{t \to \infty} \hat{\theta}(t) = \theta_\infty$$

Proof

The proof for the monovariable case can be found in Matko and Schumann (1984). It is easily extended to the multivariable case if the scalars are replaced by the corresponding vectors.

The following additional assumptions are required:

Assumption 3D

(1) $\det[P(z^{-1})] \neq 0$ for $|z| \geqq 1$

(2) $|y^*_i(t)| < M_2 < \infty, \quad 1 \leqq i \leqq p$

The convergence properties of the algorithm are summarized in the following theorem.

Theorem 3.2

Subject to Assumptions 3C-3D, the algorithm (3.88)-(3.90) leads to

(1) $\{u(t)\}$ and $\{y(t)\}$ are bounded

(2) $\lim_{t \to \infty} (y^*(t) - y(t)) = 0$ for a constant $y^*(t)$

Proof

(3.90) can be rewritten as (with $G_1(d) = I$)

$$P(d)\hat{K}(t)u(t) = \hat{A}(t,d)y^*(t) + \hat{B}(t,d)u(t) - \hat{A}(t,d)y(t)$$

$$= \hat{A}(t,d)y^*(t) + \{-y(t) + \hat{\theta}(t)^T\phi(t-1)\}$$

$$= \hat{A}(t,d)y^*(t) - b(t) \quad \text{using (3.92)} \qquad (3.93)$$

In view of Lemma 3.3 (1)-(2), Assumption 3C and C2, part (1) follows immediately from (3.93). Part (2) follows from (3.90) where

$$\lim_{t \to \infty} (y^*(t) - y(t))$$

$$= [\hat{A}(\infty,1)]^{-1}[P(1)\hat{K}(\infty) - \hat{B}(\infty,1)]u(\infty)$$

$$= 0 \quad \text{for a constant } y^*(t) \qquad (3.94)$$

since C1 holds.

Remark

A special case of the adaptive controller has appeared in the literature before, for example, the deadbeat controller of Matko and Schumann (1984).

## 3.8 Examples

In this section, simulated examples are presented to illustrate the performance of the MIMO adaptive controller, with emphasis on the decoupling problem.

Consider the following system originally introduced by Prager and Wellstead (1981) in a stochastic setting:

$$\begin{bmatrix} 1 - 1.4d + 0.48d^2 & -0.2d + 0.1d^2 \\ -0.1d & 1 - 0.9d + 0.2d^2 \end{bmatrix} y = \begin{bmatrix} d + 1.5d^2 & d^2 \\ 0 & d^2 \end{bmatrix} u \qquad (3.95)$$

A series of runs of 400 samples with different pole-zero specifications were executed using a standard least squares estimator. During the first 50 samples the adaptive controller was run in the commissioning mode. A series of steps of 60 samples in the range [0,10] were used as reference signals.

The following specifications were used:

(1) $$P(d) = \begin{bmatrix} 1 - 0.5d & 0 \\ 0 & 1 - 0.5d \end{bmatrix}, \quad G_1(d) = I$$

(2) $$P(d) = \begin{bmatrix} 1 - 0.5d & 0.1d \\ 0 & 1 - 0.5d \end{bmatrix}, \quad G_1(d) = I$$

(3) $$P(d) = \begin{bmatrix} 1 - 0.75d & 0 \\ 0 & 1 - 0.75d \end{bmatrix}, \quad G_1(d) = I$$

(4) $$P(d) = \begin{bmatrix} 1 - 0.75d & 0.05d \\ 0 & 1 - 0.75d \end{bmatrix}, \quad G_1(d) = I$$

The corresponding results for Example (1), (2), (3), and (4) are shown in Fig. 3.6, 3.7, 3.8, and 3.9 respectively. It is seen that channel 1 is more sensitive to the setpoint changes in channel 2 with a diagonal structure for P(d) ( Fig. 3.6 and 3.8). However, this interaction was reduced by using a nondiagonal structure for P(d) (Fig. 3.7 and 3.9).

Example 5

To illustrate the role of the zero-polynomial $G_1(d)$, consider the following specifications:

$$P(d) = \begin{bmatrix} 1 - 0.5d & 0 \\ 0 & 1 - 0.5d \end{bmatrix},$$

$$G_1(d) = [I + (I + P_1 - \hat{A}_1)]^{-1}(I + (I + P_1 - \hat{A}_1)d)$$

where $P_1$ and $\hat{A}_1$ are the first coefficients of P(d) and the estimated $\hat{A}(d)$, respectively.

The results obtained using the above specifications are shown in Fig. 3.10. It is seen that for this low $y_1$-$u_2$ coupling system, the effect of the introduction of $G_1(d)$ on interactions is insignificant.

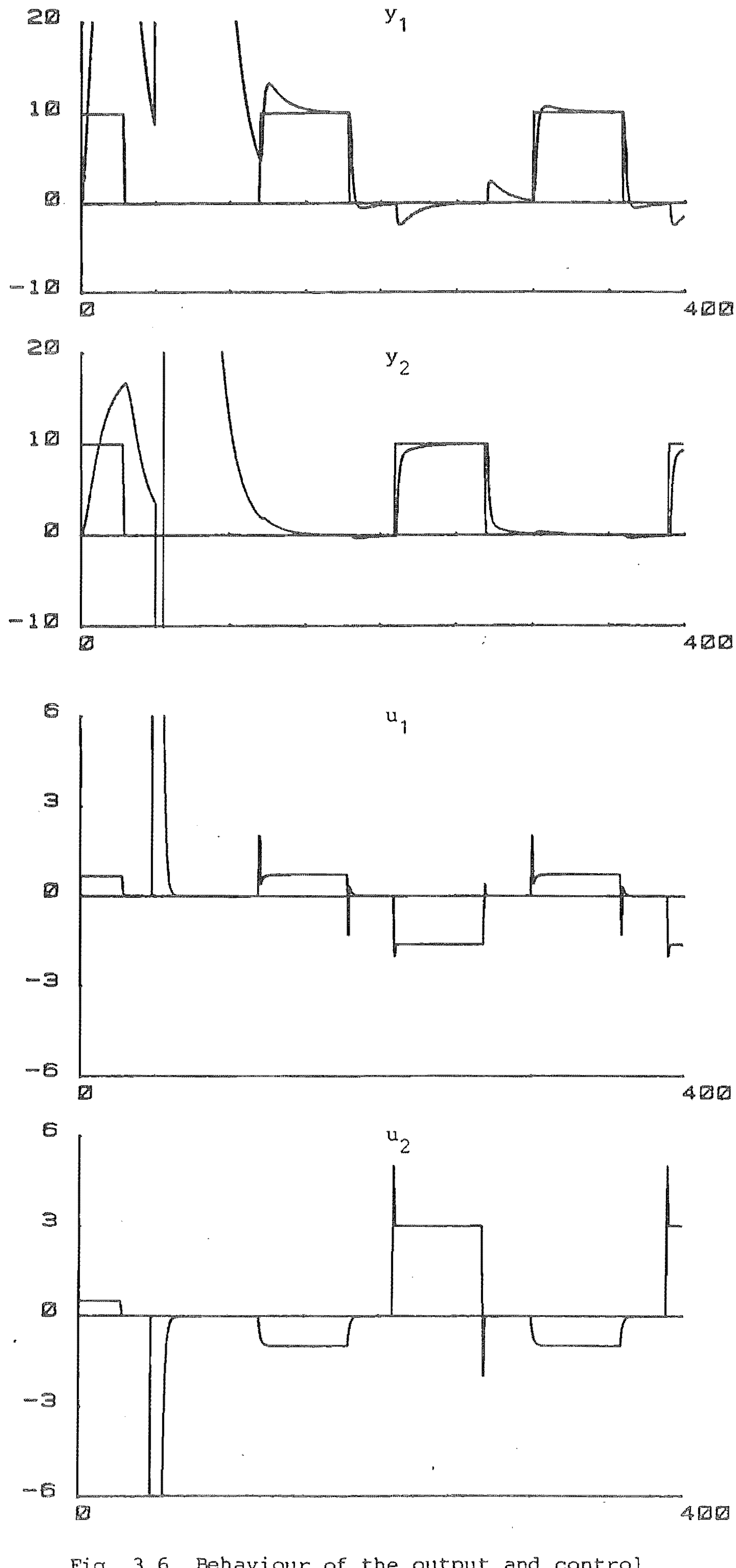



Fig. 3.6 Behaviour of the output and control variables of Example 1

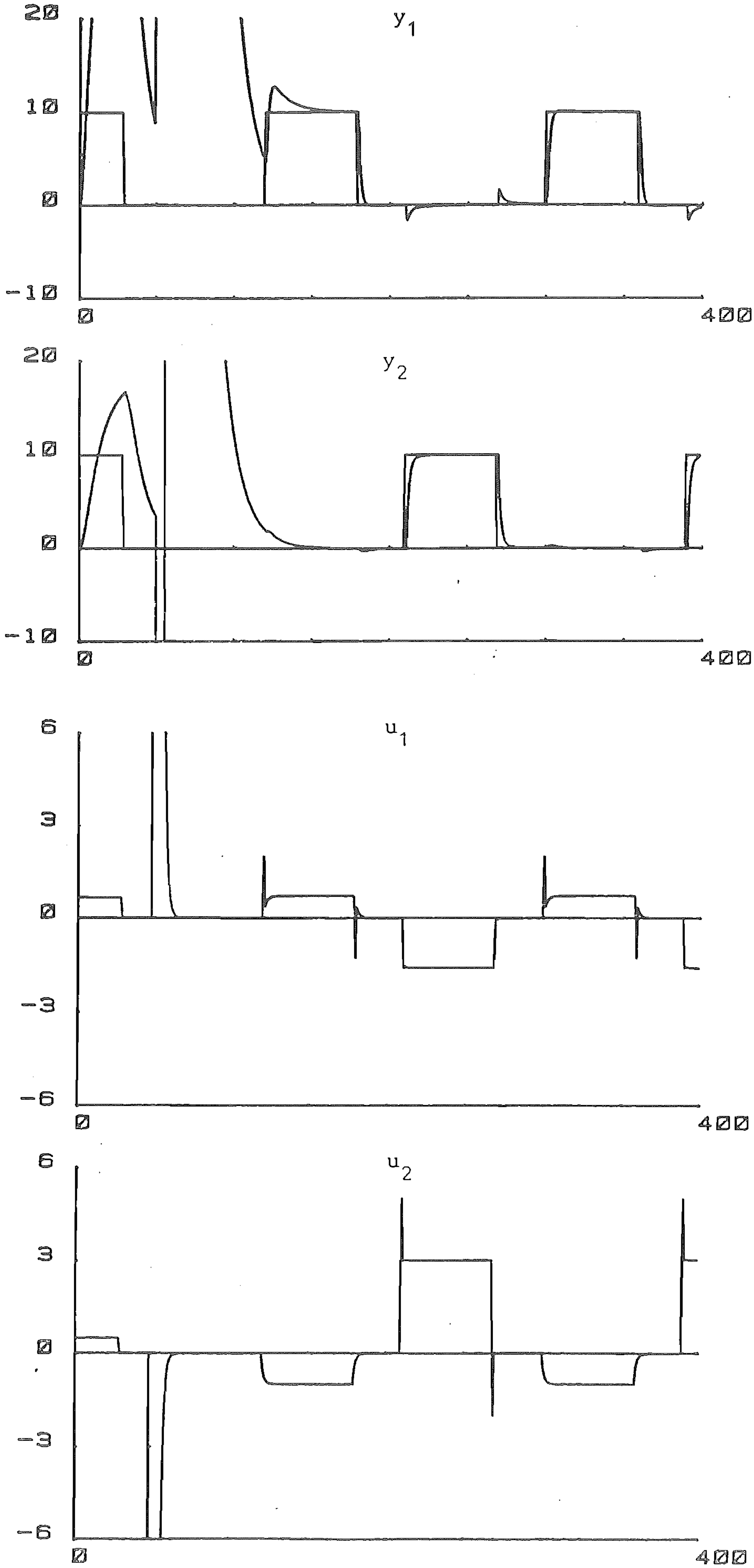



Fig. 3.7 Behaviour of the output and control variables of Example 2

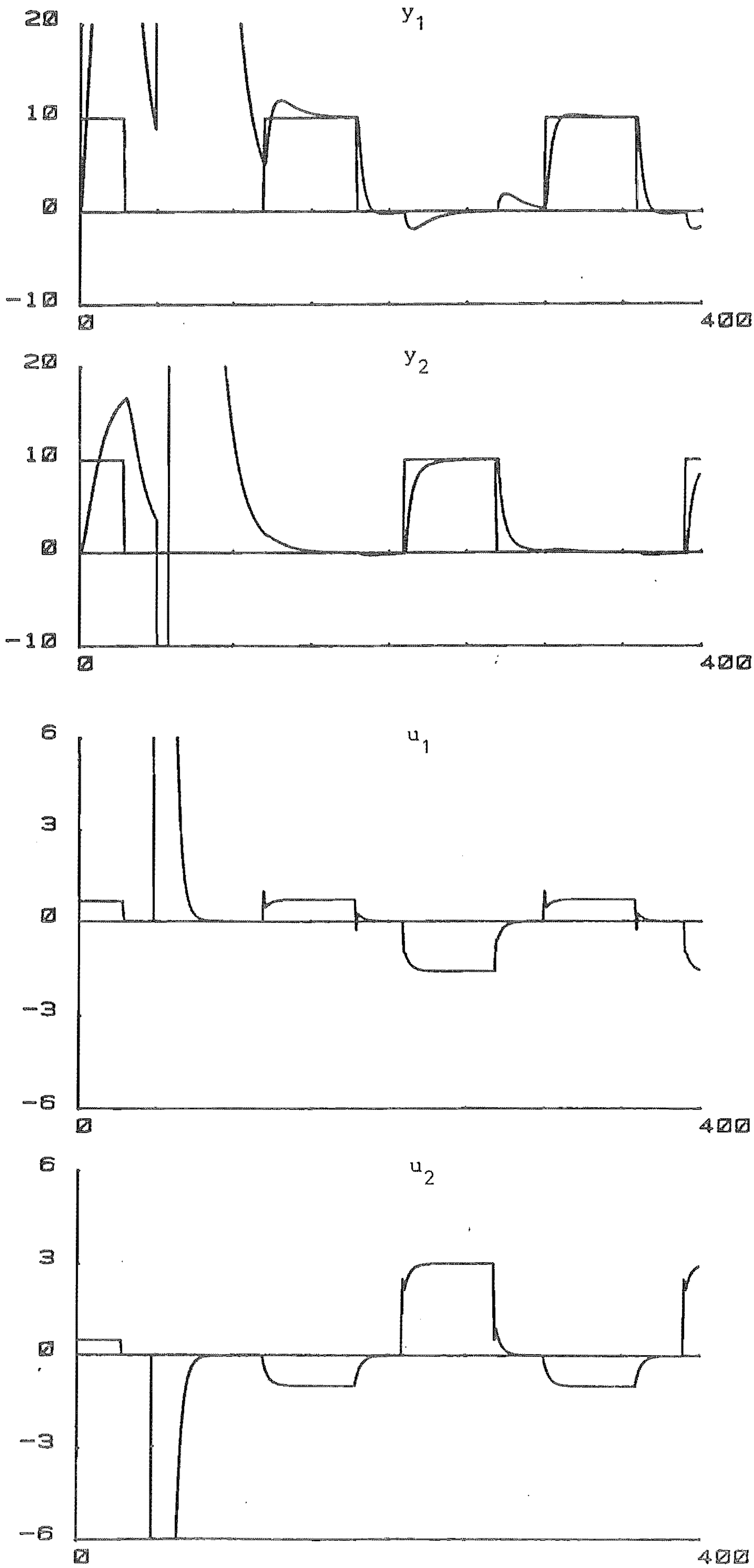



Fig. 3.8 Behaviour of the output and control variables of Example 3

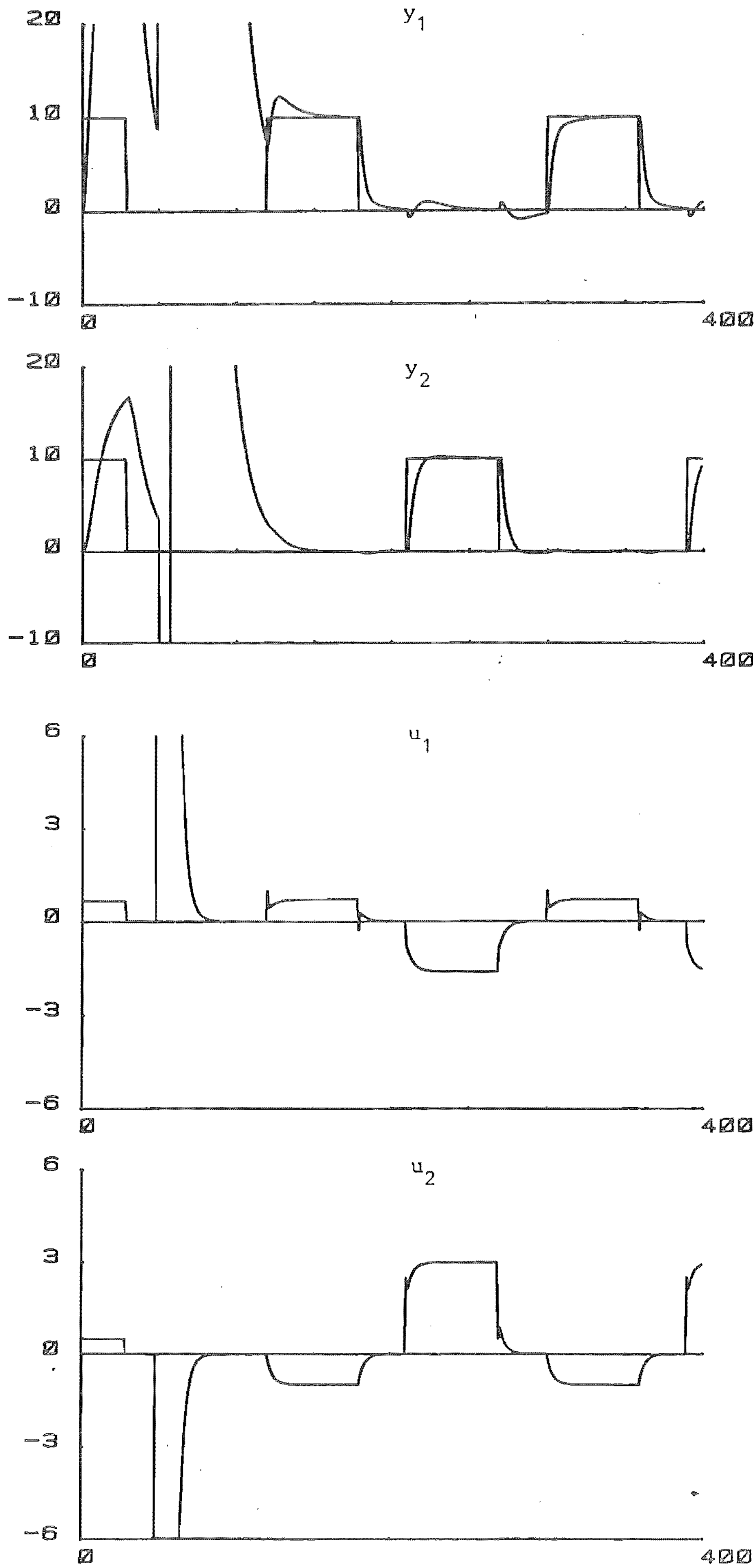



Fig. 3.9 Behaviour of the output and control variables of Example 4

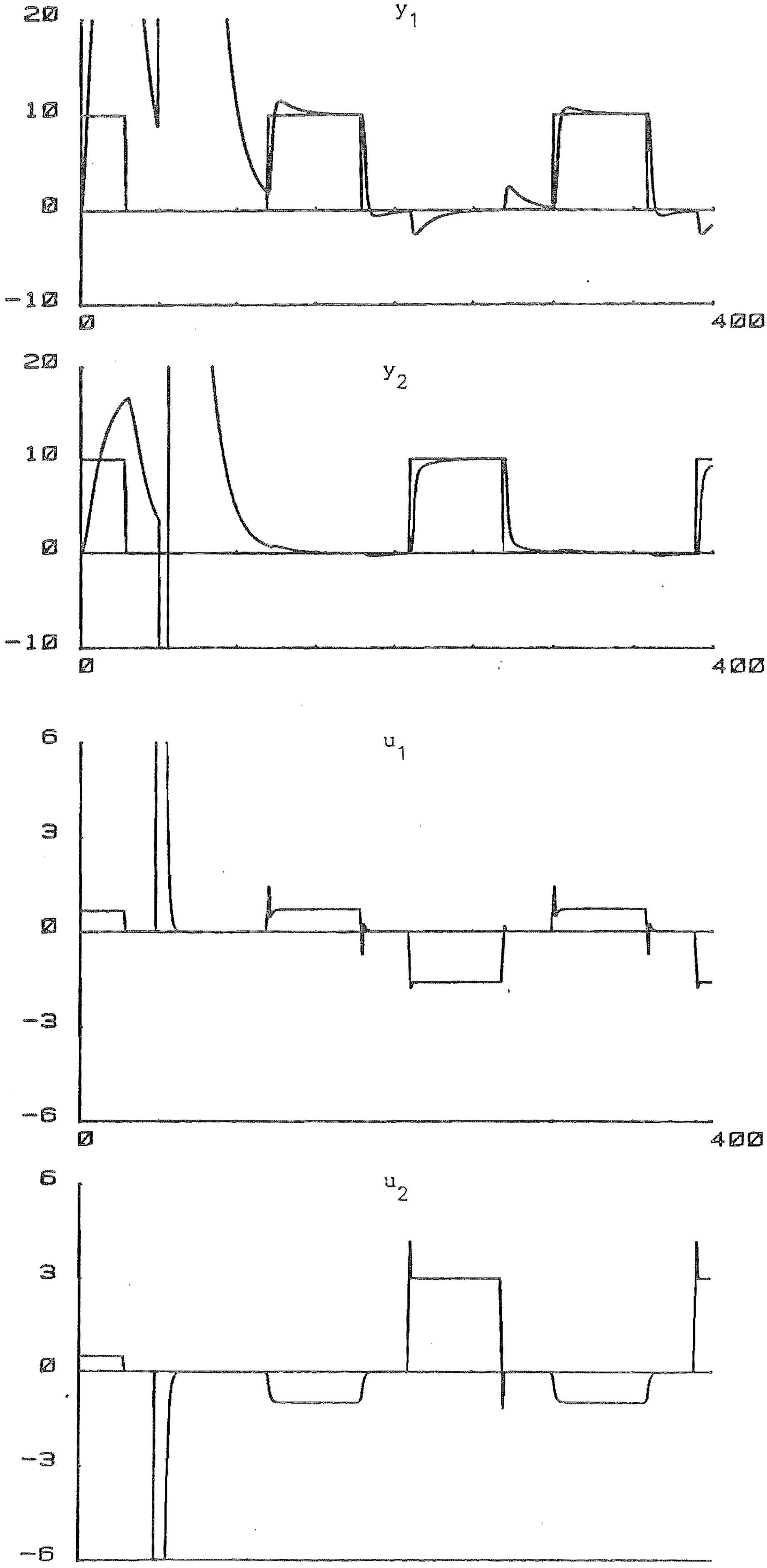



Fig. 3.10 Behaviour of the output and control variables of Example 5

Remark

The simulated examples do indicate that by using appropriate choice of P(d) and $G_1(d)$, interactions may be reduced. However, much is needed in the way of theory to develop a systematic procedure for doing this. An interesting approach to the decoupling problem has been presented for a stochastic pole placement self-tuning controller by McDermott and Mellichamp (1986).

## 3.9 Conclusion

In this chapter, stability and convergence properties of the pole-zero placement adaptive control algorithm applied to linear deterministic, time-invariant systems have been studied. The systems, although stable, need not be minimum phase and may have purely deterministic disturbances. Also, simulated examples illustrating the performance and robustness of the algorithm have been given. Both single-input single-output systems and multi-input multi-output systems have been treated.

# CHAPTER 4

# ADAPTIVE CONTROL OF STOCHASTIC SYSTEMS

## 4.1 Introduction

The adaptive control of deterministic systems in which the disturbances are predictable has been addressed in chapter 3.

In this chapter, the ideas of Chapter 3 are extended to systems having random disturbances. Specifically, this chapter considers the stochastic adaptive control of a class of nonminimum phase systems.

The stochastic adaptive control of minimum phase systems has been well studied, and complete stability and convergence results have been obtained (see, for example, Goodwin et al 1984, Kumar 1985).

However, few results have been obtained for stochastic nonminimum phase systems. The main difficulty in the analysis, as in the deterministic case (see, for example, M'saad et al 1985) has been that the estimated parameters may have common unstable roots in the estimated model. In Fuchs (1980), under a stabilizability condition on the estimated model, sufficient conditions for the estimator and the control law are established such that the closed-loop adaptive system is stable. Also, by imposing a more stringent requirement than that in Fuchs (1980), a convergent stochastic adaptive controller has been established in Hersh and Zarrop (1986). In both cases, assumptions are imposed on the asymptotic behaviour of the parameter estimates.

This chapter is organized as follows. In section 4.2, the pole -zero placement (fixed) control of systems with both deterministic and stochastic disturbances is addressed. In Section 4.3, the pole-zero placement adaptive control of systems having white noise is analyzed. In Section 4.4, the adaptive control of systems having coloured noise is also analyzed.

## 4.2 Pole-zero Placement Strategy for Known Plant

In this section, the pole-zero placement control of systems having purely deterministic and stochastic disturbances is addressed.

### 4.2.1 Deterministic and stochastic disturbances

Consider the following system:

$$y(t) = bu(t-1) + d(t) + v(t) \tag{4.1}$$

where

$$d(t) = A\sin(\rho t + \phi)$$

$$v(t) = \text{white noise with variance } \sigma^2$$

$$u(t) = \text{a known input}$$

The corresponding (stochastic) state-space model is given by

$$x(t+1) = \begin{bmatrix} 2\cos\rho & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u(t) \tag{4.2}$$

$$y(t) = [\, 1 \quad 0 \quad b \,]\, x(t) + v(t) \tag{4.3}$$

(4.2), (4.3) is an observable but uncontrollable model, and hence it can be expressed in observer form having the following structure:

$$\bar{x}(t+1) = \begin{bmatrix} 2\cos\rho & 1 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \bar{x}(t) + b \begin{bmatrix} 1 \\ -2\cos\rho \\ 1 \end{bmatrix} u(t) \tag{4.4}$$

$$y(t) = [\, 1 \quad 0 \quad 0 \,]\, \bar{x}(t) + v(t) \tag{4.5}$$

Using Kalman filtering ideas ( Anderson and Moore 1979 ), the associated innovations model for (4.4), (4.5) is

$$\hat{x}(t+1) = A\hat{x}(t) + Bu(t) + K(t)\omega(t) \tag{4.6}$$

$$y(t) = C\hat{x}(t) + \omega(t) \tag{4.7}$$

where

$$A = \begin{bmatrix} 2\cos\rho & 1 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}; \quad B = b\begin{bmatrix} 1 \\ -2\cos\rho \\ 1 \end{bmatrix}; \quad K(t) = \begin{bmatrix} k_1(t) \\ k_2(t) \\ k_3(t) \end{bmatrix}; \quad C^T = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

K(t) is obtained from the solution of the matrix Riccati equation:

$$\Sigma(t+1) = A\Sigma(t)A^T - A\Sigma(t)C^T\left(C\Sigma(t)C^T + \sigma^2\right)^{-1}C\Sigma(t)A^T;$$

$$\Sigma(0) = P_0 \tag{4.8}$$

$$K(t) = A\Sigma(t)C^T\left(C\Sigma(t)C^T + \sigma^2\right)^{-1} \tag{4.9}$$

Using (4.7) in (4.6) gives the following time-varying ARMAX (Autoregressive moving-average with auxillary input) model:

$$A(d)y(t) = B(d)u(t) + C(t,d)\omega(t) \tag{4.10}$$

where

$$A(d) = 1 - (2\cos\rho)d + d^2$$

$$B(d) = b[d - (2\cos\rho)d^2 + d^3]$$

$$C(t,d) = 1 + (k_1(t-1) - 2\cos\rho)d + (1 - k_2(t-2))d^2 + k_3(t-3)d^3$$

In many cases(see Theorem 4.1 below), the error covariance $\Sigma(t)$ and hence the Kalman gain $K(t)$ converge to steady state values as $t\to\infty$. Thus, if $\Sigma(t)$ converges as $t\to\infty$, the limiting solution $\Sigma$ will satisfy the ARE(algebraic Riccati equation):

$$\Sigma - A\Sigma A^T + A\Sigma C^T[C\Sigma C^T + \sigma^2]^{-1}C\Sigma A^T = 0 \tag{4.11}$$

$$K = A\Sigma C^T[C\Sigma C^T + \sigma^2]^{-1} \tag{4.12}$$

The properties of the solution of the ARE are of central importance since they give conditions for the stability of the filter. The necessary key properties are given below.

Definition: A real symmetric positive semidefinite solution of the ARE is said to be a strong solution if the corresponding filter state transition matrix

$$\bar{A} = A - KC$$

has all its eigenvalues inside or on the unit circle.

Lemma 4.1 (p. 253, Goodwin and Sin 1984)

Provided $(C,A)$ is detectable, the strong solution of the ARE exists and is unique.

Theorem 4.1 (p. 253, Goodwin and Sin 1984)

Subject to:

(i) $(C,A)$ is observable,

(ii) $(\Sigma_0 - \Sigma_s) > 0$ ,

then

$$\lim_{t\to\infty} \Sigma(t) = \Sigma_s$$

$$\lim_{t\to\infty} K(t) = K_s$$

$$\lim_{t \to \infty} \bar{A}(t) = \bar{A}_s$$

where $\Sigma_s$ is the unique strong solution of the ARE and $K_s$ and $\bar{A}_s$ are the corresponding steady state filter gain and state transition matrix.

The original system is observable but has two uncontrollable modes on the unit circle. Thus the strong solution to the ARE exists and is unique, and gives a steady state filter with two uncontrollable roots on the unit circle. The strong solution to the ARE above is $\Sigma_s = 0$, giving $K_s = 0$. Also, $\Sigma(t)$ and $K(t)$ converge to $\Sigma_s$ and $K_s$ for any $\Sigma(0) > 0$ (Theorem 4.1). Thus, the resulting steady state filter is given by

$$\hat{x}(t+1) = A\hat{x}(t) + Bu(t) \tag{4.13}$$

$$y(t) = C\hat{x}(t) + \omega(t) \tag{4.14}$$

The corresponding steady state ARMAX model is

$$A(d)y(t) = B(d)u(t) + C(d)\omega(t) \tag{4.15}$$

where

$$A(d) = 1 - (2\cos\rho)d + d^2$$

$$B(d) = b\left[d - (2\cos\rho)d^2 + d^3\right]$$

$$C(d) = 1 - (2\cos\rho)d + d^2$$

4.2.2 Pole-zero placement

Consider the autoregressive moving-average model with both deterministic and stochastic disturbances:

$$\bar{A}(d)D(d)y(t) = \bar{B}(d)D(d)u(t) + \bar{C}(d)D(d)\omega(t) \tag{4.16}$$

where

$$\bar{A}(d) = 1 + a_1 d + \ldots\ldots + a_{na}d^{na}$$

$$\bar{B}(d) = b_1 d + \ldots\ldots + b_{nb}d^{nb}$$

$$\bar{C}(d) = 1 + c_1 d + \ldots\ldots + c_{nc}d^{nc}$$

$D(d)$ denotes the uncontrollable modes arising from the purely deterministic disturbances.

and $\{y(t)\}$, $\{u(t)\}$ and $\{\omega(t)\}$ denote the output, input and noise sequence, respectively.

The following theorem concerns the pole-zero placement control of systems described by (4.16).

Theorem 4.2

Consider the system described by (4.16) and the control law

$$[P(d)K - G_1(d)B(d)]u(t) = G_1(d)A(d)[y^*(t) - y(t)] \quad (4.17)$$

(a) The control law (4.17) gives the closed-loop system:

$$A(d)P(d)Ky(t) = G_1(d)A(d)B(d)y^*(t) + C(d)F(d)\omega(t) \quad (4.18)$$

$$P(d)Ku(t) = G_1(d)A(d)y^*(t) - G_1(d)C(d)\omega(t) \quad (4.19)$$

where

$$A(d) = \bar{A}(d)D(d)$$

$$B(d) = \bar{B}(d)D(d)$$

$$C(d) = \bar{C}(d)D(d)$$

$$F(d) = P(d)K - G_1(d)B(d)$$

(b) $\{y(t)\}$ and $\{u(t)\}$ will be sample mean-square bounded w.p.1 ( with probability one ) provided $\{\omega(t)\}$ is sample mean-square bounded w.p.1 and if the following conditions are met:

(i) All zeros of $A(z^{-1})P(z^{-1})$ lie on or inside the unit circle,

(ii) All poles of the transfer function

$$\frac{1}{A(z^{-1})P(z^{-1})}\left[\, G_1(z^{-1})A(z^{-1})B(z^{-1}) \;\middle|\; C(z^{-1})F(z^{-1}) \,\right]$$

lie strictly inside the unit circle,

(iii) Any zeros of $A(z^{-1})P(z^{-1})$ on the unit circle have Jordan block of size 1.

Proof

(a) Straightforward.

(b) Follows from (4.18), (4.19), Lemma 3.1 of Chapter 3, and a bounded $\{y^*(t)\}$.

Remark

Theorem 4.2 is the stochastic generalization of Theorem 3.1.

## 4.3 Adaptive Control: White Noise Case

Attention is now turned to the adaptive control problem. The first case to be considered is where deterministic disturbances are absent and the stochastic disturbance is of the white noise type.

### 4.3.1 Problem statement

Consider the adaptive control of a linear time-invariant single -input single-output system admitting a representation of the form:

$$A(d)y(t) = B(d)u(t) + \omega(t) \tag{4.20}$$

where

$$A(d) = 1 + a_1 d + \ldots\ldots + a_{na} d^{na}$$

$$B(d) = b_1 d + \ldots\ldots + b_{nb} d^{nb}$$

The sequences $\{y(t)\}$, $\{u(t)\}$ and $\{\omega(t)\}$ are outputs, inputs and disturbances, respectively. The sequence $\{\omega(t)\}$ is assumed to be a real stochastic process defined on a probability space $(\Omega, F, P)$ adapted to the sequence of increasing sub-sigma algebras $(F(t), t \in N)$ where $F(t)$ is generated by the observations up to time t, and such that $\{\omega(t)\}$ satisfies with w.p.1:

Assumption 4A

(1) $E(\omega(t)/F(t-1)) = 0$

(2) $E(\omega(t)^2/F(t-1)) = \sigma^2$

(3) $\lim_{N} \frac{1}{N} \sum_{t=1}^{N} \omega(t)^2 = \sigma^2$

(4.20) can be written as

$$y(t) = \phi(t-1)^T \theta_0 + \omega(t) \tag{4.21}$$

where

$$\phi(t-1) = [\ y(t-1),\ldots,y(t-n),u(t-1),\ldots,u(t-m)\ ]^T$$

$$\theta_0 = [\ -a_1,\ldots,-a_{na},0,..0,b_1,\ldots,b_{nb},0,..,0\ ]^T$$

where

$$na \leq n, \quad nb \leq m$$

The following assumptions about the system are made:

Assumption 4B

(1) upper bounds of na and nb are known

(2) $A(z^{-1})$ has all zeros inside the unit circle

(3) $B(1) \neq 0$

The situation is that the coefficients in A(d) and B(d), and $\sigma^2$ are unknown, and that only $\{y(t)\}$ and $\{u(t)\}$ are directly available. The problem is to find a feedback control law that leads to a closed-loop system stable in some satisfactory stochastic sense.

4.3.2 The adaptation algorithm

The algorithm is described as follows:

$$\hat{\theta}'(t) = \hat{\theta}(t-1) + \frac{\phi(t-1)}{r(t-1)}[y(t) - \phi(t-1)^T \hat{\theta}(t-1)] \tag{4.22}$$

$$r(t-1) = r(t-2) + \phi(t-1)^T\phi(t-1) \ ; \quad r(-1) = 1 \tag{4.23}$$

The symbols in (4.22) have the following meaning:

$$\hat{\theta}(t) = [\ -\hat{a}_1(t),\ldots,-\hat{a}_n(t),\hat{b}_1(t),\ldots\ldots,\hat{b}_m(t)\ ]^T$$

$$\hat{A}(t,d) = 1 + \hat{a}_1(t)d + \ldots\ldots + \hat{a}_n(t)d^n$$

$$\hat{B}(t,d) = \hat{b}_1(t)d + \ldots\ldots + \hat{b}_m(t)d^m$$

The estimated parameter $\hat{\theta}'(t)$ is modified by the following projection scheme:

$$\hat{\theta}(t) = \begin{cases} \hat{\theta}'(t), & \text{if } \hat{\theta}'(t) \in C \\ \\ \hat{\theta}^*(t), & \text{if } \hat{\theta}'(t) \notin C \end{cases} \tag{4.24}$$

where C is a closed-convex set satisfying:

(1) $\theta_0 \in C$

(2) $C \subset \{\ \theta(t)\colon\ \hat{\rho}_i(t) = 1 - \rho < 1,\ i = 1,\ldots,n$ (4.25)

$\hat{\rho}_i(t)$ are the roots of $\hat{A}(t,d)\ \}$

If the algorithm gives rise to a $\hat{\theta}'(t)$ outside C, $\hat{\theta}'(t)$ is projected orthogonally onto the surface of C before continuing. Here, $\hat{\theta}^*(t)$ can be computed using a similar procedure as the one in Chapter 3 (the computation is simpler as it does not involve a linear transformation).

The control signal at each time t is determined from the control law

$$\hat{F}(t,d)u(t) = \hat{G}(t,d)[y^*(t) - y(t)] \qquad (4.26)$$

where

$$\hat{G}(t,d) = G_1(d)\hat{A}(t,d)$$

$$\hat{F}(t,d) = P(d)\hat{K}(t) - G_1(d)\hat{B}(t,d)$$

$$\hat{K}(t) = \begin{cases} \dfrac{\hat{B}(t,1)}{P(1)} & \text{if } |\hat{B}(t,1)| > \varepsilon \\ \\ \dfrac{\varepsilon}{P(1)} & \text{otherwise} \end{cases} \qquad (4.27)$$

$\varepsilon$ is an arbitrary positive scalar.

The following additional assumptions are made:

Assumption 4C

(1) $|y^*(t)| < M_1 < \infty$

(2) $P(z^{-1})$ has all roots inside the unit circle.

Let

$$\tilde{\theta}(t) = \hat{\theta}(t) - \theta_0 \qquad (4.28)$$

$$\beta(t) = -\phi(t)^T \tilde{\theta}(t) \qquad (4.29)$$

The basic properties of the constrained stochastic gradient algorithm (4.22)-(4.24) are summarized in the following lemma.

Lemma 4.2

Subject to Assumption 4A, the algorithm (4.22) - (4.24) ensures that, w.p.1:

(1) $||\hat{\theta}(t)|| < M_2$

(2) $||\hat{\theta}(t) - \hat{\theta}(t-1)|| \longrightarrow 0$ as $t \longrightarrow \infty$

(2') $\sum_{t=1}^{\infty} ||\hat{\theta}(t) - \hat{\theta}(t-k)||^2 < \infty$ for any finite k

(3) $\sum_{t=1}^{\infty} \dfrac{\beta(t)^2}{r(t)} < \infty$

Proof

The key step in extending the proof of the unconstrained stochastic gradient algorithm (Fuchs 1980) to the constrained one is to note that when projection is used

$$\tilde{\theta}(t)^T \tilde{\theta}(t) \le \tilde{\theta}'(t)^T \tilde{\theta}'(t)$$

Hence, by the usual stochastic Lyapunov argument, the basic properties (1), (2), and (3) of the unconstrained algorithm are retained.

(2') Follows immediately from

$$\sum_{t=1}^{\infty} ||\hat{\theta}(t) - \hat{\theta}(t-k)||^2 \le k \sum_{t=1}^{\infty} \sum_{j=0}^{k-1} ||\hat{\theta}(t-j) - \hat{\theta}(t-j-1)||^2$$

4.3.3 Stability Analysis

In this section, the stability of the algorithm (4.22) - (4.27) is studied. The key result is summarized in the theorem below.

Theorem 4.3

With Assumptions 4A - 4C, the algorithm (4.22) - (4.27) ensures that, w.p.1:

(1) $\sup_N \frac{1}{N} \sum_{t=1}^{N} ||\phi(t)||^2 < \infty$

(2) $\lim_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} E((y(t) - \phi(t-1)^T \hat{\theta}(t-1))^2/F(t-1)) = \sigma^2$

Proof

To make the principal idea more transparent, the simplified pole placement case (i.e. $G_1(d) = 1$) will be considered first. The extension to the general case will then be quite straightforward proceeding along the lines of the pole placement case.

Using (4.21) and (4.28), y(t) can be written as

$$y(t) = \phi(t-1)^T \hat{\theta}(t-1) - \phi(t-1)^T \tilde{\theta}(t-1) + \omega(t) \qquad (4.30)$$

The control law (4.26) (with $G_1(d) = 1$) can be written as

$$\begin{aligned} P(d)\hat{K}(t)u(t) &= \hat{A}(t,d)y^*(t) + \hat{B}(t,d)u(t) - \hat{A}(t,d)y(t) \\ &= w^*(t) + \phi(t-1)^T \hat{\theta}(t) - y(t) \qquad (4.31) \end{aligned}$$

where

$$w^*(t) = \hat{A}(t,d)y^*(t) \qquad (4.32)$$

Notice that $\{w^*(t)\}$ is uniformly bounded, since $\{y^*(t)\}$ and $\{\hat{a}_i(t)\}$ are uniformly bounded by Assumption 4C and Lemma 4.2-(1).

Let

$$P(d) = 1 + p_1 d + \ldots\ldots + p_{np} d^{np}; \quad np \le m \tag{4.33}$$

Rewriting (4.31) using (4.30) and rearranging gives

$$\begin{aligned} u(t) = \hat{f}(t) w^*(t) &+ [\psi^T + \hat{f}(t)(\hat{\theta}(t) - \hat{\theta}(t-1))^T]\phi(t-1) \\ &- \hat{f}(t)(-\phi(t-1)^T \tilde{\theta}(t-1) + \omega(t)) \end{aligned} \tag{4.34}$$

where

$$\hat{f}(t) = \hat{K}(t)^{-1} \tag{4.35}$$

$$\psi = [\ 0 \ \ldots \ 0 \ | -p_1 \ldots\ldots -p_{np} \ \ 0 .. 0\ ]^T$$

Combining (4.30) and (4.34) yields

$$\phi(t) = F(t-1)\phi(t-1) + B(t)(\beta(t-1) + \omega(t)) + D(t) w^*(t) \tag{4.36}$$

where

$$F(t-1) = \begin{bmatrix} \longleftarrow & \hat{\theta}(t-1)^T & \longrightarrow \\ I_{n-1} & & 0 \\ \longleftarrow & \psi^T + \hat{f}(t)(\hat{\theta}(t) - \hat{\theta}(t-1))^T & \longrightarrow \\ 0 & & I_{m-1} \end{bmatrix}$$

$$B(t) = [\ 1 \ 0 \ \ldots \ 0 \ | -\hat{f}(t) \ 0 \ \ldots \ 0\ ]^T$$

$$D(t) = [\ 0 \ \ldots\ldots \ 0 \ | \ \hat{f}(t) \ 0 \ \ldots \ 0\ ]^T$$

To proceed with the analysis, the following lemma due to Fuchs (1980) is required.

Lemma 4.3

Consider the l-dimensional time-varying linear system

$$X(t+1) = F(t)X(t) + Bu(t)$$

with $\{u(t)\}$ a scalar input sequence, and $X(t)$ the lx1 state vector. Assume that:

(1) $||F(t)|| < M_3 \quad \forall\ t > 0$

(2) $||F(t) - F(t-1)|| \to 0$ as $t \to \infty$

(3) $\rho(F(t)) \leq \delta < 1 \quad \forall\ t > N'$ [ $\rho$ = spectral radius of $F(t)$ ]

Then, there exist $C_1$ and $C_2$ , which are independent of N such that

$$\sum_{t=1}^{N} ||X(t)||^2 \leq C_1 \sum_{t=1}^{N} |u(t-1)|^2 + C_2$$

Now, Lemma 4.2-(1) and (4.27) imply the uniform boundedness of $||F(t)||$ and $\hat{f}(t)$; Lemma 4.2-(2) implies that $||F(t) - F(t-1)|| \to 0$, $|\hat{f}(t) - \hat{f}(t-1)| \to 0$ as $t \to \infty$ w.p.1. Also, for a sufficiently large, but finite t, the eigenvalues of $F(t)$ are arbitrarily close to the zeros of $\hat{A}(t,z^{-1})P(z^{-1})$, which are inside the unit circle due to (4.25) and Assumption 4CH(2). Hence, assumptions (1) - (3) of Lemma 4.3 hold.

Using the superposition and Lemma 4.3 gives

$$\sum_{t=1}^{N} ||\phi(t)||^2 \leq C_1 + C_2 \sum_{t=1}^{N} \left(\beta(t-1)^2 + \omega(t)^2 + w^*(t)^2\right) \tag{4.37}$$

Using Assumption 4A-(3), and the boundedness of $\{w^*(t)\}$, there exists an $N_1$ such that, w.p.1

$$\frac{1}{N} \sum_{t=1}^{N} ||\phi(t)||^2 \leq C_3 + \frac{C_2}{N} \sum_{t=1}^{N} \beta(t-1)^2 \quad , N \geq N_1 \tag{4.38}$$

(4.38) implies

$$\frac{r(N)}{N} \leq C_4 + \frac{C_2}{N} \sum_{t=1}^{N} \beta(t-1)^2 \quad , N \geq N_1 \tag{4.39}$$

Now, if $r(t) < L_1 < \infty$, then Lemma 4.2-(3) implies that

$$\lim_{N \to \infty} \frac{1}{L_1} \sum_{t=1}^{N} \beta(t)^2 < \infty \quad \text{w.p.1} \tag{4.40}$$

and hence

$$\lim_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} \beta(t)^2 = 0 \quad \text{w.p.1} \tag{4.41}$$

Alternatively, if $r(t)$ is unbounded, then since the sum in Lemma 4.2-(3) is nondecreasing, using the following lemma

Kronecker Lemma (Appendix D, Goodwin and Sin 1984)

$$\left.\begin{array}{ll}\text{(a)} & \sum_{k=1}^{n} x_k \quad \text{converges} \\ \text{(b)} & \{b_n\} \text{ nondecreasing} \\ \text{(c)} & \lim_{n \to \infty} b_n = \infty \end{array}\right\} \text{implies that} \lim_{n \to \infty} \frac{1}{b_n} \sum_{k=1}^{n} b_k x_k = 0$$

gives

$$\lim_{N \to \infty} \frac{N}{r(N)} \frac{1}{N} \sum_{t=1}^{N} \beta(t)^2 = 0 \qquad \text{w.p.1} \qquad (4.42)$$

Substituting (4.39) into (4.42) gives

$$\lim_{N \to \infty} \frac{\frac{1}{N} \sum_{t=1}^{N} \beta(t)^2}{C_4 + \frac{C_2}{N} \sum_{t=1}^{N} \beta(t-1)^2} = 0 \qquad \text{w.p.1} \qquad (4.43)$$

(4.41) follows immediately. Using (4.41) in (4.39) yields

$$\lim_{N \to \infty} \sup \frac{r(N)}{N} < L_2 < \infty \qquad \text{w.p.1} \qquad (4.44)$$

which establishes Theorem 4.3-(1).

For the second part of Theorem 4.3, the proof is as follows. From (4.30) and taking squares yields

$$(y(t) - \phi(t-1)^T \hat{\theta}(t-1))^2 = (\beta(t-1) + \omega(t))^2$$

$$= (\beta(t-1)^2 + 2\beta(t-1)\omega(t) + \omega(t)^2) \qquad (4.45)$$

Taking the conditional expectation of both sides and using Assumption 4A-(1) and that $\beta(t-1)$ is $F(t-1)$ measurable (i.e., $E(\beta(t-1)/F(t-1)) = \beta(t-1)$) gives

$$E((y(t) - \phi(t-1)^T \hat{\theta}(t-1))^2/F(t-1)) = \beta(t-1)^2 + E(\omega(t)^2/F(t-1)) \qquad (4.46)$$

Theorem 4.3-(2) now follows from Assumption 4A-(2), (4.41) and (4.46).

The above proof can be extended to the general case as follows.

Let

$$G_1(d) = \alpha_0 + \alpha_1 d + \ldots\ldots + \alpha_r d^r \tag{4.47}$$

The control law (4.26) can be written as

$$P(d)\hat{K}(t)u(t) = G_1(d)\hat{A}(t,d)y^*(t) + G_1(d)\hat{B}(t,d)u(t) - G_1(d)\hat{A}(t,d)y(t)$$

$$= v^*(t) + \sum_{i=0}^{r} \alpha_i \phi(t-i-1)^T \hat{\theta}(t) - \sum_{i=0}^{r} \alpha_i y(t-i) \tag{4.48}$$

where

$$v^*(t) = G_1\hat{A}(t,d)y^*(t) \tag{4.49}$$

Rewriting (4.48) using (4.30) and rearranging gives

$$u(t) = \hat{f}(t)v^*(t) + \psi^T\phi(t-1) + \hat{f}(t)\sum_{i=0}^{r} \alpha_i(\hat{\theta}(t) - \hat{\theta}(t-i-1))^T\phi(t-i-1)$$

$$- \hat{f}(t)\sum_{i=0}^{r} \alpha_i(-\phi(t-i-1)^T\tilde{\theta}(t-i-1) + \omega(t-i)) \tag{4.50}$$

(4.50) can be written in the form

$$u(t) = \hat{f}(t)v^*(t) + [\bar{\psi}^T + \hat{f}(t)\sum_{i=0}^{r} \alpha_i(\bar{\theta}(t) - \bar{\theta}(t-i-1))^T]X(t-1)$$

$$- \hat{f}(t)\sum_{i=0}^{r} \alpha_i[-\phi(t-i-1)^T\tilde{\theta}(t-i-1) + \omega(t-i)] \tag{4.51}$$

where

$$X(t-1) = [\ y(t-1),\ldots,y(t-n-r),u(t-1),\ldots,u(t-m-r)\ ]^T$$

and $\bar{\psi}$ and $\bar{\theta}(t)$ are constructed from $\psi$ and $\hat{\theta}(t)$, respectively.

Similarly, (4.30) can be written as

$$y(t) = X(t-1)^T\bar{\theta}(t-1) - \phi(t-1)^T\tilde{\theta}(t-1) + \omega(t) \tag{4.52}$$

Combining (4.51) and (4.52) yields

$$X(t) = F(t-1)X(t-1) + B_0(t)(\beta(t-1) + \omega(t)) + \ldots.$$

$$+ B_r(t)(\beta(t-r-1) + \omega(t-r)) + D(t)v^*(t) \tag{4.53}$$

where

$$F(t-1) = \left[\begin{array}{cc} \longleftarrow\!\!-\!\!-\!\!- \; \bar{\theta}(t-1)^T & -\!\!-\!\!-\!\!\longrightarrow \\ I_{n+r-1} & 0 \\ \longleftarrow\!\!-\; \bar{\psi}^T + \hat{f}(t)\displaystyle\sum_{i=0}^{r} \alpha_i(\bar{\theta}(t) - & \bar{\theta}(t-i-1))^T \;-\!\!\longrightarrow \\ 0 & I_{m+r-1} \end{array}\right]$$

$$B_0(t) = [\; \underbrace{1 \;\; 0 \; \ldots \; 0}_{n+r} \;|\; \underbrace{-\alpha_0\hat{f}(t) \;\; 0 \; \ldots \; 0}_{m+r} \;]^T$$

$$B_j(t) = [\; 0 \; \ldots\ldots \; 0 \;|\; -\alpha_j\hat{f}(t) \;\; 0 \; \ldots \; 0 \;]^T; \quad j \neq 0$$

$$D(t) = [\; 0 \; \ldots\ldots \; 0 \;|\; \hat{f}(t) \;\; 0 \; \ldots \; 0 \;]^T$$

Using an argument similar to the pole placement case, it can be concluded that

$$\sum_{t=1}^{N} ||X(t)||^2 \leq K_1 + K_2 \sum_{t=1}^{N} [(\sum_{i=0}^{r} \beta(t-i-1)^2 + \omega(t-i)^2) + v^*(t)^2] \tag{4.54}$$

Using Assumption 4A-(3), and the boundedness of $\{v^*(t)\}$, there exists an $N_2$ such that, w.p.1

$$\frac{1}{N}\sum_{t=1}^{N} ||X(t)||^2 \leq K_3 + \frac{K_2}{N} \sum_{t=1}^{N} (\sum_{i=0}^{r} \beta(t-i-1)^2) \quad , N \geq N_2 \tag{4.55}$$

By definition,

$$||\phi(t)|| \leq ||X(t)||$$

and thus

$$\frac{1}{N}\sum_{t=1}^{N} ||\phi(t)||^2 \leq K_3 + \frac{K_2}{N} \sum_{t=1}^{N} (\sum_{i=0}^{r} \beta(t-i-1)^2) \quad , N \geq N_2 \tag{4.56}$$

(4.56) implies

$$\frac{r(N)}{N} \leq K_4 + \frac{K_2}{N} \sum_{t=1}^{N} (\sum_{i=0}^{r} \beta(t-i-1)^2) \quad , N \geq N_2 \tag{4.57}$$

The rest of the proof is exactly as in the pole placement case.

4.3.4 Convergence Analysis

In Section 4.3.3, it has been shown that the pole-zero placement adaptive control algorithm leads to stability in the sense of Theorem 4.3-(1).

In this section it is shown that the same algorithm leads to convergence in the sense that, w.p.1:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} [\hat{A}(t)P\hat{K}(t)y(t) - \hat{B}(t)\hat{G}(t)y^*(t) - \hat{F}(t)\omega(t)]^2 = 0$$

Proof

In the following, the manipulation of time-varying polynomial operators will be required. This is facilitated by the following notation. Given time-varying polynomials $\hat{A}(t,d)$, $\hat{B}(t,d)$, define

$$\hat{A}\hat{B} = \sum_{ij}\sum \hat{a}_i(t)\hat{b}_j(t)d^{i+j} \qquad (4.58)$$

$$\hat{A}*\hat{B} = \sum_{ij}\sum \hat{a}_i(t)\hat{b}_j(t-i)d^{i+j} \qquad (4.59)$$

Also define

$$\hat{B}'' = \hat{B}(t-1,d) \qquad (4.60)$$

The key equations required are

$$\hat{F}u(t) = \hat{G}y^*(t) - \hat{G}y(t) \quad \text{[from (4.17)]} \qquad (4.61)$$

$$\hat{F} + G_1\hat{B} = P\hat{K}(t) \quad \text{[from (4.18)]} \qquad (4.62)$$

$$\begin{aligned} e(t) &= y(t) - \phi(t-1)^T \hat{\theta}(t-1) \\ &= y(t) - [(1 - \hat{A}'')y(t) + \hat{B}''u(t)] \\ &= \hat{A}''y(t) - \hat{B}''u(t) \end{aligned} \qquad (4.63)$$

Now define

$$\begin{aligned} \hat{B}*\hat{G}y^*(t) &= \hat{B}*\hat{F}u(t) + \hat{B}*\hat{G}y(t) \\ &= \hat{B}\hat{G}y(t) + [\hat{B}*\hat{G} - \hat{B}\hat{G}]y(t) + \hat{B}\hat{F}u(t) + [\hat{B}*\hat{F} - \hat{B}\hat{F}]u(t) \\ &= \hat{B}\hat{G}y(t) + \hat{A}\hat{F}y(t) - \hat{F}e(t) + [\hat{B}*\hat{F} - \hat{B}\hat{F}]u(t) \\ &\quad + [\hat{B}*\hat{G} - \hat{B}\hat{G}]y(t) - [\hat{F}*\hat{B}'' - \hat{F}\hat{B}]u(t) + [\hat{F}*\hat{A}'' - \hat{F}\hat{A}]y(t) \end{aligned}$$

using (4.63)

$$= \hat{A}\hat{P}Ky(t) - \hat{F}e(t) + [\hat{B}^*\hat{F} - \hat{B}\hat{F}]u(t) + [\hat{B}^*\hat{G} - \hat{B}\hat{G}]y(t)$$

$$- [\hat{F}^*\hat{B}'' - \hat{F}\hat{B}]u(t) + [\hat{F}^*\hat{A}'' - \hat{F}\hat{A}]y(t) \quad \text{using (4.62)} \qquad (4.64)$$

From (4.63)

$$e(t) = y(t) - \phi(t-1)^T \hat{\theta}(t-1)$$

$$= \phi(t-1)^T\theta_0 + \omega(t) - \phi(t-1)^T \hat{\theta}(t-1) \quad \text{using (4.21)}$$

$$= -\phi(t-1)^T \tilde{\theta}(t-1) + \omega(t) \quad \text{using (4.28)}$$

$$= \beta(t-1) + \omega(t) \quad \text{using (4.29)} \qquad (4.65)$$

(4.64) could be rearranged to give

$$\hat{A}\hat{P}Ky(t) - \hat{B}\hat{G}y^*(t) - \hat{F}\omega(t) = \lambda(t-1) + \gamma(t-1) + \zeta(t-1) \qquad (4.66)$$

where

$$\lambda(t-1) = \hat{F}\beta(t-1) \qquad (4.67)$$

$$\gamma(t-1) = -\Big[ \big([\hat{B}^*\hat{F} - \hat{B}\hat{F}] - [\hat{F}^*\hat{B}'' - \hat{F}\hat{B}]\big)u(t)$$

$$+ \big([\hat{B}^*G - \hat{B}\hat{G}] + [\hat{F}^*\hat{A}'' - \hat{F}\hat{A}]\big)y(t) \Big] \qquad (4.68)$$

$$\zeta(t-1) = [\hat{B}^*\hat{G} - \hat{B}\hat{G}]y^*(t) \qquad (4.69)$$

Hence using the Schwarz inequality leads to

$$[\hat{A}\hat{P}Ky(t) - \hat{B}\hat{G}y^*(t) - \hat{F}\omega(t)]^2 \leq 3\,[\lambda(t-1)^2 + \gamma(t-1)^2 + \zeta(t-1)^2] \qquad (4.70)$$

Summing from 1 to N and dividing by N gives

$$\frac{1}{N}\sum_{t=1}^{N} [\hat{A}\hat{P}Ky(t) - \hat{B}\hat{G}y^*(t) - \hat{F}\omega(t)]^2$$

$$\leq \frac{3}{N}\sum_{t=1}^{N} [\lambda(t-1)^2 + \gamma(t-1)^2 + \zeta(t-1)^2] \qquad (4.71)$$

From (4.67), taking squares and using the Schwarz inequality gives

$$\lambda(t-1)^2 = \Big(\sum_{i=0}^{m+r} \hat{f}_i(t)\beta(t-i-1)\Big)^2$$

$$\leq \Big(\sum_{i=0}^{m+r} \hat{f}_i(t)^2\Big)\Big(\sum_{i=0}^{m+r} \beta(t-i-1)^2\Big) \qquad (4.72)$$

So summing from 1 to N and dividing by N yield

$$\frac{1}{N}\sum_{t=1}^{N}\lambda(t-1)^2 \leq \frac{1}{N}\Big[\sum_{t=1}^{N}\big(\sum_{i=0}^{m+r}\hat{f}_i(t)^2\big)\Big]\Big[\sum_{t=1}^{N}\big(\sum_{i=0}^{m+r}\beta(t-i-1)^2\big)\Big] \tag{4.73}$$

Now, (4.68) can be written as

$$\begin{aligned}\gamma(t-1) = \Big[ &(\hat{B}\hat{G}y(t) + \hat{B}\hat{F}u(t)) - (\hat{B}^*\hat{G}y(t) + \hat{B}^*\hat{F}u(t)) \\ &+ (\hat{F}\hat{A}y(t) - \hat{F}\hat{B}u(t)) - (\hat{F}^*\hat{A}''y(t) - \hat{F}^*\hat{B}''u(t)) \Big] \\ = &\sum_{i=1}^{m}\hat{b}_i(t)(\psi(t) - \psi(t-i))^T\phi_{ex}(t-i) \\ &+ \sum_{i=0}^{m+r}\hat{f}_i(t)(\hat{\theta}(t-i-1) - \hat{\theta}(t))^T\phi(t-i-1)\end{aligned} \tag{4.74}$$

where

$$\psi(t) = [\ \hat{g}_0,\ldots,\hat{g}_{n+r}(t),\hat{f}_0(t),\ldots,\hat{f}_{m+r}(t)\ ]^T$$

$$\phi_{ex}(t) = [\ y(t),\ldots,y(t-n-r),u(t),\ldots,u(t-m-r)\ ]^T$$

Note that

$$||\phi_{ex}(t-1)|| \leq \sum_{i=0}^{r+1}||\phi(t-i-1)||$$

and so using Theorem 4.3-(1) gives

$$\begin{aligned}\sup_N \frac{1}{N}\sum_{t=1}^{N}||\phi_{ex}(t-1)||^2 &\leq \sup_N \frac{r+2}{N}\sum_{t=1}^{N}\big(\sum_{i=0}^{r+1}||\phi(t-i-1)||^2\big) \\ &< K\end{aligned} \tag{4.75}$$

Apply the Schwarz inequality to (4.74) to obtain

$$\begin{aligned}\gamma(t-1)^2 \leq 2\ \Big\{ &\big(\sum_{i=1}^{m}\hat{b}_i(t)(\psi(t) - \psi(t-i))^T\phi_{ex}(t-i)\big)^2 \\ &+ \big(\sum_{i=0}^{m+r}\hat{f}_i(t)(\hat{\theta}(t-i-1) - \hat{\theta}(t))^T\phi(t-i-1)\big)^2 \Big\} \\ \leq 2\ \Big\{ &m\ \big(\sum_{i=1}^{m}\hat{b}_i(t)^2||\psi(t) - \psi(t-i)||^2||\phi_{ex}(t-i)||^2\big) \\ &+ (m+r+1)\ \big(\sum_{i=0}^{m+r}\hat{f}_i(t)^2||\hat{\theta}(t-i-1) - \hat{\theta}(t)||^2||\phi(t-i-1)||^2\big)\Big\}\end{aligned} \tag{4.76}$$

So summing from 1 to N and dividing by N gives

$$\frac{1}{N}\sum_{t=1}^{N}\gamma(t-1)^2 \leq 2\ \{\ \frac{m}{N}\sum_{t=1}^{N}\ (\sum_{i=1}^{m}\hat{b}_i(t)^2||\psi(t)\ -\ \psi(t-i)||^2||\phi_{ex}(t-i)||^2).$$

$$+\ \frac{(m+r+1)}{N}\sum_{t=1}^{N}\ (\sum_{i=0}^{m+r}\hat{f}_i(t)^2||\hat{\theta}(t-i-1)\ -\ \hat{\theta}(t)||^2||\phi(t-i-1)||^2)\ \}$$

(4.77)

Finally from (4.69), taking squares, summing from 1 to N and dividing by N gives

$$\frac{1}{N}\sum_{t=1}^{N}\ \zeta(t-1)^2 = \frac{1}{N}\sum_{t=1}^{N}\left[\sum_{i=i}^{m}\ \sum_{j=0}^{n+r}\hat{b}_i(t)(\hat{g}_j(t-i)\ -\ \hat{g}_j(t))y^*(t-i-j)\right]^2$$

(4.78)

Hence using the boundedness of the coefficients of $\hat{A}$, $\hat{B}$, and the conver-gence of $||\hat{\theta}(t) - \hat{\theta}(t-k)||$ to zero, (4.41) and (4.73), plus (4.75), (4.77) and (4.78), it can be concluded from (4.71) that for a given $\delta$ arbitrarily small there exists an $N_3$ such that, for $N \geq N_3$

$$\frac{1}{N}\sum_{t=1}^{N}\ [\hat{A}\hat{P}Ky(t)\ -\ \hat{B}\hat{G}y^*(t)\ -\ \hat{F}\omega(t)]^2 < \delta \qquad \text{w.p.1} \qquad (4.79)$$

Again, using the convergence of $||\hat{\theta}(t) - \hat{\theta}(t-k)||$ to zero and (4.41) lead to

$$\lim_{N\to\infty}\ \frac{1}{N}\sum_{t=1}^{N}\ [\hat{A}\hat{P}Ky(t)\ -\ \hat{B}\hat{G}y^*(t)\ -\ \hat{F}\omega(t)]^2 = 0 \qquad \text{w.p.1} \qquad (4.80)$$

which is the announced property.

## 4.4 Adaptive Control: Coloured-Noise Case

In this section, a more general noise structure is considered for the pole-zero placement adaptive control scheme. A key difference between this case and the white noise case is that a modified version of the previous parameter estimation algorithm is used.

### 4.4.1 Problem statement

Consider the adaptive control of a linear time-invariant SISO system admitting an autoregressive moving-average representation of the form

$$A(d)y(t) = B(d)u(t) + C(d)\omega(t) \qquad (4.81)$$

where

$$A(d) = 1 + a_1 d + \dots\dots + a_{na} d^{na}$$

$$B(d) = b_1 d + \dots\dots\dots + b_{nb} d^{nb}$$

$$C(d) = 1 + c_1 d + \dots\dots\dots\dots + c_{nc} d^{nc}$$

with the following usual assumption on the noise (see Assumption 4A):

Assumption 4D

(1) $E(\omega(t)/F(t-1)) = 0$ w.p.1

(2) $E(\omega(t)^2/F(t-1)) = \sigma^2$ w.p.1

(3) $\lim_{N} \frac{1}{N} \sum_{t=1}^{N} \omega(t)^2 = \sigma^2$ w.p.1

(4.81) can be written as

$$y(t) = \phi(t-1)^T \theta_0 + \omega(t) \tag{4.82}$$

where

$$\phi(t-1) = [\ y(t-1),..,y(t-n),u(t-1),..,u(t-m),\omega(t-1),..,\omega(t-p)\ ]^T$$

$$\theta_0 = [\ -a_1,\dots,-a_{na},0,..,0,b_1,\dots,b_{nb},0,..,0,c_1,..,c_{nc}]^T$$

where

$$na \leqq n, \quad nb \leqq m, \quad nc = p$$

The following assumptions about the system are made:

Assumption 4E

(1) upper bounds of na, nb and nc are known
(2) $C(z^{-1})$ has all zeros inside the unit circle.
(3) $(C(z) - 1/2)$ is strictly positive real
(4) $A(z^{-1})$ has all zeros inside the unit circle.
(5) $B(1) \neq 0$

The situation is that the coefficients in $A(d)$, $B(d)$ and $C(d)$, and $\sigma^2$ are unknown, and that only $\{y(t)\}$ and $\{u(t)\}$ are directly available. The problem is to find a feedback control law that leads to a closed-loop system stable in some satisfactory stochastic sense.

### 4.4.2 The adaptation algorithm

The algorithm is described as follows:

$$\hat{\theta}'(t) = \hat{\theta}(t-1) + \frac{\hat{\phi}(t-1)}{r(t-1)}\left[y(t) - \hat{\phi}(t-1)^T \hat{\theta}(t-1)\right] \tag{4.83}$$

$$r(t-1) = r(t-2) + \hat{\phi}(t-1)^T \hat{\phi}(t-1); \quad r(-1) = 1 \tag{4.84}$$

The symbols in (4.83) have the following meaning:

$$\hat{\phi}(t-1) = [\ y(t-1),..,y(t-n),u(t-1),..,u(t-m),e(t-1),..,e(t-p)\ ]^T$$

$$\hat{\theta}(t) = [\ -\hat{a}_1(t),..,-\hat{a}_n(t),\hat{b}_1(t),..,\hat{b}_m(t),\hat{c}_1(t),..,\hat{c}_p(t)\ ]^T$$

$$e(t) = y(t) - \hat{\phi}(t-1)^T \hat{\theta}(t-1) \tag{4.85}$$

$$\hat{A}(t,d) = 1 + \hat{a}_1(t)d + ...... + \hat{a}_n(t)d^n$$

$$\hat{B}(t,d) = \hat{b}_1(t)d + ........ + \hat{b}_m(t)d^m$$

$$\hat{C}(t,d) = 1 + \hat{c}_1(t)d + .......... + \hat{c}_p(t)d^p$$

The estimated parameter $\hat{\theta}'(t)$ is modified by the following projection scheme:

$$\hat{\theta}(t) = \begin{cases} \hat{\theta}'(t), & \text{if } \hat{\theta}'(t) \in C \\ \\ \hat{\theta}^*(t), & \text{if } \hat{\theta}'(t) \notin C \end{cases} \tag{4.86}$$

where C is a closed-convex set satisfying:

(1) $\theta_0 \in C$

(2) $C \subset \{\ \theta(t):\ \hat{\rho}_i(t) = 1 - \rho < 1\ ,i = 1,...,n$ (4.87)

$\hat{\rho}_i(t)$ are the roots of $\hat{A}(t,d)\ \}$

If the algorithm gives rise to a $\hat{\theta}'(t)$ outside C, $\hat{\theta}'(t)$ is projected orthogonally onto the surface of C before continuing.

The control signal is determined as in (4.26). (4.88)

Remark

Although the coefficients of $\hat{C}(t,d)$ are available, they are, however, not used in the calculation of the controller parameters.

The following addtional assumptions are made:

Assumption 4E

(1) $|y^*(t)| < M_3 < \infty$

(2) $P(z^{-1})$ has all roots inside the unit circle.

Let

$$\xi(t-1) = e(t) - \omega(t)$$

$$= \phi(t-1)^T\theta_0 - \hat{\phi}(t-1)^T \hat{\theta}(t-1) \qquad (4.89)$$

$$\tilde{\theta}(t) = \hat{\theta}(t) - \theta_0 \qquad (4.90)$$

The elementary properties of the algorithm (4.83) - (4.86) are summarized in the following lemma.

Lemma 4.4

Subject to Assumption 4D, the algorithm (4.83) - (4.86) ensures that, w.p.1:

(1) $||\hat{\theta}(t)|| < M_4$

(2) $||\hat{\theta}(t) - \hat{\theta}(t-1)|| \longrightarrow 0$ as $t \longrightarrow \infty$

(2') $\sum_{t=1}^{\infty} ||\hat{\theta}(t) - \hat{\theta}(t-k)||^2 < \infty$ for any finite k

(3) $\sum_{t=1}^{\infty} \frac{\xi(t)^2}{r(t)} < \infty$

Proof

The algorithm (4.83) - (4.86) is a constrained version of the algorithm used by Hersh and Zarrop (1986). The basic properties of the unconstrained version are retained as in Lemma 4.4 due to the projection procedure.

4.4.3 Stability analysis

Here, the stability of the adaptive algorithm (4.83) - (4.88) is analyzed. The key result is summarized in the theorem below.

Theorem 4.4

With Assumptions 4D - 4F, the algorithm (4.83) - (4.88) ensures that, w.p.1:

(1) $\sup_N \frac{1}{N} \sum_{t=1}^{N} ||\phi(t)||^2 < \infty$

(2) $\lim_{N \longrightarrow \infty} \frac{1}{N} \sum_{t=1}^{N} E((y(t) - \hat{\phi}(t-1)^T \hat{\theta}(t-1))^2/F(t-1)) = \sigma^2$

Proof

Let

$$P(d) = 1 + p_1 d + \ldots\ldots + p_{np} d^{np}; \quad np \leq m \tag{4.91}$$

$$G_1(d) = \alpha_0 + \alpha_1 d + \ldots\ldots + \alpha_r d^r \tag{4.92}$$

Using (4.82) and (4.90), y(t) can be written as

$$y(t) = \phi(t-1)^T \hat{\theta}(t-1) - \phi(t-1)^T \tilde{\theta}(t-1) + \omega(t) \tag{4.93}$$

The control law (4.26) can be written as

$$P(d)\hat{K}(t)u(t) = G_1(d)\hat{A}(t,d)y^*(t) + G_1(d)\hat{B}(t,d)u(t) - G_1(d)\hat{A}(t,d)y(t)$$

$$= v^*(t) + \sum_{i=0}^{r} \alpha_i \phi(t-i-1)^T \hat{\theta}(t) - \sum_{i=0}^{r} \alpha_i y(t-i) \tag{4.94}$$

where

$$v^*(t) = G_1\hat{A}(t,d)y^*(t) \tag{4.95}$$

$$\hat{\theta}(t) = [\ -\hat{a}_1(t),\ldots,-\hat{a}_n(t),\hat{b}_1(t),\ldots,\hat{b}_m(t),0,\ldots\ldots,0\ ]^T$$

Notice that $\{v^*(t)\}$ is uniformly bounded, since $\{y^*(t)\}$ and $\{\hat{a}_i(t)\}$ are uniformly bounded by Assumption 4F and Lemma 4.4-(1). Rewriting (4.94) using (4.93) and rearranging gives

$$u(t) = \hat{f}(t)v^*(t) + \psi^T\phi(t-1) + \hat{f}(t)\sum_{i=0}^{r} \alpha_i(\hat{\theta}(t) - \hat{\theta}(t-i-1))^T\phi(t-i-1)$$

$$- \hat{f}(t)\sum_{i=0}^{r} \alpha_i\left(-\phi(t-i-1)^T \tilde{\theta}(t-i-1) + \omega(t-i)\right) \tag{4.96}$$

where

$$\hat{f}(t) = K(t)^{-1} \tag{4.97}$$

$$\psi = [\ 0\ \ldots\ 0\ |-p_1\ \ldots\ldots\ -p_{np}\ \ 0\ \ldots\ 0\ |0\ \ldots\ 0\ ]^T$$

(4.96) can be written in the form

$$u(t) = \hat{f}(t)v^*(t) + \left(\bar{\psi}^T + \hat{f}(t)\sum_{i=0}^{r} \alpha_i(\bar{\theta}(t) - \bar{\theta}(t-i-1))^T\right)X(t-1)$$

$$- \hat{f}(t)\sum_{i=0}^{r} \alpha_i\left(-\phi(t-i-1)^T \tilde{\theta}(t-i-1) + \omega(t-i)\right) \tag{4.98}$$

where

$$X(t-1) = [y(t-1),..,y(t-n-r),u(t-1),..,u(t-m-r),\omega(t-1),..,\omega(t-p-r)]^T$$

and $\bar{\psi}$, $\bar{\Theta}(t)$, and $\bar{\theta}(t)$ are constructed from $\psi$, $\hat{\Theta}(t)$, and $\hat{\theta}(t)$, respectively.

Similarly, (4.93) can be written as

$$y(t) = X(t-1)^T \bar{\theta}(t-1) - \phi(t-1)^T \tilde{\theta}(t-1) + \omega(t) \qquad (4.99)$$

Let

$$\beta(t-1) = -\phi(t-1)^T \tilde{\theta}(t-1) \qquad (4.100)$$

Combining (4.98) and (4.99) yields

$$X(t) = F(t-1)X(t-1) + B_0(t)\beta(t-1) + \ldots + B_r(t)\beta(t-r-1)$$

$$+ D_0(t)\omega(t) + \ldots + D_r(t)\omega(t-r) + D(t)v^*(t) \qquad (4.101)$$

where

$$F(t-1) = \begin{bmatrix} \longleftarrow & \bar{\theta}(t-1)^T & \longrightarrow \\ I_{n+r-1} & & 0 \\ \longleftarrow & \bar{\psi}^T + \hat{f}(t) \sum_{i=0}^{r} \alpha_i(\bar{\Theta}(t) - \bar{\theta}(t-i-1))^T & \longrightarrow \\ 0 & I_{m+r-1} & \\ & & I_{p+r-1} \end{bmatrix}$$

$$B_0(t) = [\underbrace{1\ 0\ ..\ 0}_{n+r} \mid \underbrace{-\alpha_0\hat{f}(t)\ 0\ ..\ 0}_{m+r} \mid \underbrace{0\ \ldots\ 0}_{p+r}]^T$$

$$B_j(t) = [\ 0\ \ldots\ 0 \mid -\alpha_j\hat{f}(t)\ 0\ \ldots\ 0 \mid 0\ \ldots\ 0\ ]^T;\ j \neq 0$$

$$D_0(t) = [\ 1\ 0\ ..\ 0 \mid -\alpha_0\hat{f}(t)\ 0\ ..\ 0 \mid 1\ 0\ \ldots\ 0\ ]^T$$

$$D_j(t) = [\ 0\ \ldots\ 0 \mid -\alpha_j\hat{f}(t)\ 0\ \ldots\ 0 \mid 0\ \ldots\ 0\ ]^T;\ j \neq 0$$

$$D(t) = [\ 0\ \ldots\ 0 \mid \hat{f}(t)\ 0\ ..\ 0 \mid 0\ \ldots\ 0\ ]^T$$

Thus, using a similar argument to the white noise case, it can be concluded, using the superposition and Lemma 4.3, that

$$\sum_{t=1}^{N} ||X(t)||^2 \leq K_1 + K_2 \sum_{t=1}^{N} [(\sum_{i=0}^{r} \beta(t-i-1)^2 + \omega(t-i)^2) + v^*(t)^2] \quad (4.102)$$

Using Assumption 4D-(3), the boundedness of $\{v^*(t)\}$, the definition of $||\phi(t)||$, $||X(t)||$, and $r(t)$, there exists an $N_4$ such that, w.p.1

$$\frac{r(N)}{N} \leq K_3 + \frac{K_2}{N} \sum_{t=1}^{N} (\sum_{i=0}^{r} \beta(t-i-1)^2) \qquad \text{for } N \geq N_4 \quad (4.103)$$

Now, using (4.90) and (4.89), $\beta(t-1)$ can be written as

$$\begin{aligned} \phi(t-1)^T \tilde{\theta}(t-1) &= \phi(t-1)^T \hat{\theta}(t-1) - \phi(t-1)^T\theta_0 \\ &= (\phi(t-1) - \hat{\phi}(t-1))^T \hat{\theta}(t-1) - \phi(t-1)^T\theta_0 + \hat{\phi}(t-1)^T \hat{\theta}(t-1) \\ &= (\phi(t-1) - \hat{\phi}(t-1))^T \hat{\theta}(t-1) - \xi(t-1) \qquad (4.104) \end{aligned}$$

Hence, using the defintion of $\phi(t)$, $\hat{\phi}(t)$ and (4.89), (4.104) becomes

$$\begin{aligned} \phi(t-1)^T \tilde{\theta}(t-1) &= -\xi(t-1) - \sum_{i=1}^{p} \hat{c}_i(t-1)\xi(t-i-1) \\ &= -\hat{C}(t-1,d)\xi(t-1) \qquad (4.105) \end{aligned}$$

Using the Schwarz inequality yields

$$\begin{aligned} (\phi(t-1)^T \tilde{\theta}(t-1))^2 &\leq (p+1)[\xi(t-1)^2 + \sum_{i=1}^{p} (\hat{c}_i(t-1)\xi(t-i-1))^2] \\ &\leq (p+1)K'' [\sum_{i=0}^{p} \xi(t-i-1)^2] \qquad (4.106) \end{aligned}$$

where

$$K'' = \max(1, \hat{c}_1(t-1)^2, \ldots, \hat{c}_p(t-1)^2)$$

Note that $\{\hat{c}_i(t)\}$ is uniformly bounded due to Lemma 4.4-(1). Substituting (4.106) into (4.103) gives

$$\frac{r(N)}{N} \leq K_3 + K_4 \frac{1}{N} \sum_{t=1}^{N} \sum_{i=0}^{p+r} \xi(t-i-1)^2 \qquad \text{for } N \geq N_4 \quad (4.107)$$

From Lemma 4.4-(3) and using Kronecker Lemma

$$\lim_{N \to \infty} \frac{N}{r(N)} \frac{1}{N} \sum_{t=1}^{N} \xi(t)^2 = 0 \qquad \text{w.p.1} \quad (4.108)$$

Thus substituting (4.107) into (4.108) yields

$$\lim_{N \to \infty} \frac{\frac{1}{N}\sum_{t=1}^{N} \xi(t)^2}{K_3 + \frac{K_4}{N}\sum_{t=1}^{N} \left(\sum_{i=0}^{p+r} \xi(t-i-1)^2\right)} = 0 \qquad \text{w.p.1} \qquad (4.109)$$

and hence

$$\lim_{N \to \infty} \frac{1}{N}\sum_{t=1}^{N} \xi(t)^2 = 0 \qquad \text{w.p.1} \qquad (4.110)$$

Using (4.110) in (4.107) gives

$$\lim_{N \to \infty} \sup \frac{r(N)}{N} < L_3 < \infty \qquad \text{w.p.1} \qquad (4.111)$$

which establishes Theorem 4.4-(1).

For the second part of Theorem 4.4, the proof is as follows. Subtract $\hat{\phi}(t-1)^T \hat{\theta}(t-1)$ from both sides of (4.82) and taking squares lead to

$$\begin{aligned}(y(t) - \hat{\phi}(t-1)^T \hat{\theta}(t-1))^2 &= (\phi(t-1)^T\theta_0 - \hat{\phi}(t-1)^T \hat{\theta}(t-1) + \omega(t))^2 \\ &= (\xi(t-1) + \omega(t))^2 \quad \text{using (4.89)} \\ &= (\xi(t-1)^2 + 2\xi(t-1)\omega(t) + \omega(t)^2) \qquad (4.112)\end{aligned}$$

Taking the conditional expectation of both sides and using Assumption 4D-(1), (2) and that $\xi(t-1)$ is $F(t-1)$ measurable gives

$$E((y(t) - \hat{\phi}(t-1)^T \hat{\theta}(t-1))^2/F(t-1)) = \xi(t-1)^2 + \sigma^2 \qquad (4.113)$$

Theorem 4.4-(2) now follows from (4.110) and (4.113).

### 4.4.4 Convergence analysis

In section 4.4.3, stability of the adaptive control algorithm has been established in the sense of Theorem 4.4-(1).

In this section, it is shown that the same adaptive control leads to convergence in the sense that, w.p.1:

$$\lim_{N \to \infty} \frac{1}{N}\sum_{t=1}^{N} [\hat{A}(t)P\hat{K}(t)y(t) - \hat{B}(t)\hat{G}(t)y^*(t) - \hat{F}(t)\hat{C}(t)\omega(t)]^2 = 0$$

The proof is as follows. The key equations required are given below (the notation on time-varying operators can be found in 4.3.4).

$$e(t) = y(t) - \hat{\phi}(t-1)^T \hat{\theta}(t-1) \quad [\text{from } (4.85)]$$

$$= y(t) - \left[(1 - \hat{A}'')y(t) + \hat{B}''u(t) + (\hat{C}'' - 1)e(t)\right]$$

$$= \hat{A}''y(t) - \hat{B}''u(t) - (\hat{C}'' - 1)e(t) \tag{4.114}$$

(4.114) gives

$$\hat{A}''y(t) - \hat{B}''u(t) - \hat{C}''e(t) = 0 \tag{4.115}$$

From (4.88)

$$\hat{F} + G_1\hat{B} = PK(t) \tag{4.116}$$

Now define

$$\hat{B}^*\hat{G}y^*(t) = \hat{B}^*\hat{F}u(t) + \hat{B}^*\hat{G}y(t)$$

$$= \hat{B}\hat{G}y(t) + [\hat{B}^*\hat{G} - \hat{B}\hat{G}]y(t) + \hat{B}\hat{F}u(t) + [\hat{B}^*\hat{F} - \hat{B}\hat{F}]u(t)$$

$$= \hat{B}\hat{G}y(t) + \hat{A}\hat{F}y(t) - \hat{F}^*\hat{C}''e(t) + [\hat{B}^*\hat{F} - \hat{B}\hat{F}]u(t)$$

$$+ [\hat{B}^*\hat{G} - \hat{B}\hat{G}]y(t) - [\hat{F}^*\hat{B}'' - \hat{F}\hat{B}]u(t) + [\hat{F}^*\hat{A}'' - \hat{F}\hat{A}]y(t)$$

$$\text{using } (4.115)$$

$$= \hat{A}P\hat{K}y(t) - \hat{F}^*\hat{C}''e(t) + [\hat{B}^*\hat{F} - \hat{B}\hat{F}]u(t) + [\hat{B}^*\hat{G} - \hat{B}\hat{G}]y(t)$$

$$- [\hat{F}^*\hat{B}'' - \hat{F}\hat{B}]u(t) + [\hat{F}^*\hat{A}'' - \hat{F}\hat{A}]y(t) \text{ using } (4.116) \tag{4.117}$$

(4.117) can be rearranged to give

$$\hat{A}P\hat{K}y(t) - \hat{B}\hat{G}y^*(t) - \hat{F}\hat{C}e(t)$$

$$= -\left([\hat{F}^*\hat{A}'' - \hat{F}\hat{A}] + [\hat{B}^*\hat{G} - \hat{B}\hat{G}]\right)y(t) - \left([\hat{B}^*\hat{F} - \hat{B}\hat{F}] - [\hat{F}^*\hat{B}'' - \hat{F}\hat{B}]\right)u(t)$$

$$+ [\hat{B}^*\hat{G} - \hat{B}\hat{G}]y^*(t) + [\hat{F}^*\hat{C}'' - \hat{F}\hat{C}]e(t) \tag{4.118}$$

Substituting (4.89) into (4.118) gives

$$\hat{A}P\hat{K}y(t) - \hat{B}\hat{G}y^*(t) - \hat{F}\hat{C}\omega(t)$$

$$= -\left([\hat{F}^*\hat{A}'' - \hat{F}\hat{A}] + [\hat{B}^*\hat{G} - \hat{B}\hat{G}]\right)y(t) - \left([\hat{B}^*\hat{F} - \hat{B}\hat{F}] - [\hat{F}^*\hat{B}'' - \hat{F}\hat{B}]\right)u(t)$$

$$+ [\hat{B}^*\hat{G} - \hat{B}\hat{G}]y^*(t) + [\hat{F}^*\hat{C}'' - \hat{F}\hat{C}]\omega(t) + \hat{F}^*\hat{C}''\xi(t-1) \tag{4.119}$$

Thus using a similar argument to the one for the white noise case yields the announced result:

$$\lim_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} [\hat{A}P\hat{K}y(t) - \hat{B}\hat{G}y^*(t) - \hat{F}\hat{C}\omega(t)]^2 = 0 \quad \text{w.p.1} \tag{4.120}$$

## 4.5 Conclusion

In this chapter the stability and convergence properties of the pole-zero placement adaptive control algorithm applied to stochastic, systems have been studied. The systems, although stable, have not been assumed to be minimum phase. A noteworthy feature is that no assumptions on the asymptotic behaviour of the parameter estimates are imposed.

The stability and convergence results have been derived under the assumption that all zeros of $A(z^{-1})$ are within the unit circle. This means that these results do not apply in the presence of purely deterministic disturbances of the type considered in chapter 3.

So far, the adaptive control algorithms have been typically analyzed under ideal conditions. In the remainder of the thesis, attention will be paid to the question of preserving stability when those ideal conditions do not hold.

# CHAPTER 5

# ROBUSTNESS: FIXED STRATEGY

## 5.1 Introduction

In this chapter, the robustness properties of the pole-zero placement design method in the presence of plant perturbations are analyzed. The nature and extend of the perturbations that preserve stability are discussed. Then the problem of robust disturbance attenuation is examined. The same tools as in Gawthrop and Lim (1982) and Lim (1982) are used: namely, the input-output stability approach.

The studies of Gawthrop and Lim (1982) and Lim (1982) were on the robustness of a model reference nonadaptive controller in the presence of certain modelling errors. Lim (1982) also studied the robustness of the pole-zero placement methods( Wellstead et al 1979, Astrom and Wittenmark 1980).

Previously, the robustness of a different pole-zero placement method (which involves the solution of a Diophantine equation) applied to linear systems has been studied by Astrom (1980) in which sufficient conditions are given for stability.

The organization of this chapter is as follows. Section 5.2 presents the terminology and definitions of the input-output approach. In Section 5.3, a review of some classical stability theorems is given. Section 5.4 presents a fairly general system structure that models higher order and nonlinear systems. The remaining of the chapter addresses the issue of robustness of feedback stability in the presence of plant perturbations and the problem of robust disturbance attenuation.

## 5.2 Preliminaries

The notations and definitions presented here are easily available from standard texts on functional analysis, and also from papers by Zames (1963), (1966).

Discrete-time signals can be regarded as sequences of real numbers. Each signal forms a vector (of appropriate dimension) and is an element of a set, known as a linear vector space in which addition and scalar multiplication are defined.

The norm of a vector x in a linear vector space X is a measure of the size of the vector x. It maps each vector in X into a real number. Let $x = \{ \xi_1, \xi_2, .... \}$ be an element in X. The typical norms on X are:

$$||x||_2 = (\sum^{\infty} |\xi_i|^2)^{1/2}$$

$$||x||_\infty = \sup_{i \geq 1} |\xi_i|$$

The corresponding normed spaces are known as the $l_2$ and $l_\infty$, respectively.

To allow for infinite norms, a space $X_e$ is constructed, in which every finite-time truncation of every signal in $X_e$ will be in X. Thus, x belongs to the extended signal space $l_{2e}$, if

$$||x||_T = (\sum^{T} |\xi_i|^2)^{1/2} < \infty$$

for all positive integers of T.

Unless otherwise stated, $||.||$ will refer to $l_2$ norms.

An operator will be referred to as the relationship between an element in one space with another element in the same space. For example, a dynamical input-output representation is a (stable) operator $H: l_2 \longrightarrow l_2$, mapping an $l_2$ input sequence into an $l_2$ output sequence.

The gain of the operator H is defined as the smallest $\gamma_1$ such that

$$||Hx|| \leq \gamma_1 ||x|| + \gamma_2, \quad \text{for all } x \in l_2 \text{ and } \gamma_2$$

An operator H is said to be interior (exterior) conic if there are real constants c and $r \geq 0$ such that

$$||Hx - cx|| \leq r||x|| \quad ( \geq r||x|| ), \quad x \in l_2$$

c is known as the centre and r the radius of the cone.

An operator H is said to be inside (outside) the sector $(\alpha, \beta)$ if the inner product

$$\langle Hx - \alpha x, Hx - \beta x \rangle \leq 0 \quad ( \geq 0 ), \; x \in l_2$$

H is passive (or positive) if

$$\langle x, Hx\rangle \geq 0, \ x \in l_2$$

or if H is in the sector $[0,\infty)$.

## 5.3 Classical Stability Theorems

Unless otherwise stated, all systems will be assumed to be discrete time-invariant and single-input single-output.

### Nyquist Stability Criterion

The Nyquist Criterion is a well-known stability test for continuous-time systems. It can easily be reformulated to handle discrete-time systems. The criterion is particularly useful for the analysis of the stability of closed-loop system when the open-loop system is given.

Consider the standard feedback system shown in Fig. 5.1.

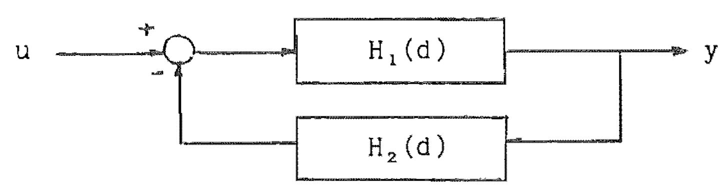



Fig. 5.1 Standard feedback system

Then the system is $l_p$ stable ( i.e., $u \in l_p$ implies $y \in l_p$ ) if the following conditions hold:

(1) $1 + H_1(0)H_2(0) \neq 0$ ,

(2) The Nyquist locus of $H_1(d)H_2(d)$ does not pass through $(-1,0)$,

(3) The Nyquist locus of $H_1(d)H_2(d)$ encircles the $-1+j0$ point p times in the anticlockwise direction where p is the number of poles of $H_1(d)H_2(d)$ in the unstable region,

(4) $H_2(d)$ does not have zeros which cancel the poles of $H_1(d)$ in the unstable region.

Proof

See Desoer and Vidyasagar (1975), where the proof is developed using the z-transform.

The Nyquist Stability Criterion provides the necessary and sufficient conditions for stability. Also, the Nyquist locus provides robustness measures in terms of the well-known gain and phase margins.

The gain margin is the ratio by which the open-loop gain must increase at a phase angle of 180 degrees to cause instability. The phase margin is the additional phase lag at unity open-loop gain required to bring about instability. Proper phase and gain margins ensure that modest variations in system components can be tolerated without causing instability.

Small-gain Theorem

When systems $H_1$ and $H_2$ in Fig. 5.1 are not necessarily linear the small-gain theorem gives a sufficient condition under which a 'bounded input' produces a 'bounded output'. Consider the general feedback system shown in Fig. 5.2.

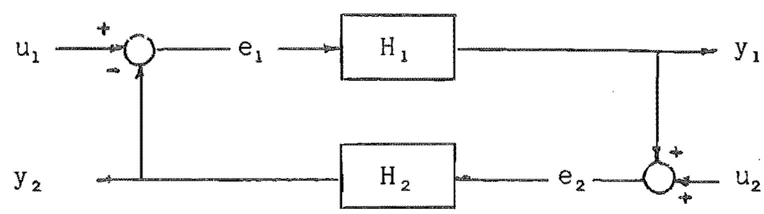



Fig. 5.2 General feedback system

Let all the signals in Fig. 5.2 be elements of the extended space, $l_{2e}$. Also, let $\gamma_1$ and $\gamma_2$ be the gains of the operators $H_1$ and $H_2$, respectively. Suppose there are constants $\beta_1$ and $\beta_2$ such that

$$||H_1e_1||_T \leq \gamma_1||e_1||_T + \beta_1$$

$$||H_2e_2||_T \leq \gamma_2||e_2||_T + \beta_2$$

If $\gamma_1\gamma_2 < 1$, then the feedback system is $l_{2e}$ bounded in the sense that $u_1, u_2 \in l_{2e}$ implies that $e_1, e_2, y_1, y_2 \in l_{2e}$.

Proof

See Desoer and Vidyasagar (1975).

The small-gain theorem is only sufficient. Also, it is only applicable to feedback systems with stable operators.

Of course, there are other classical stability results, for example, the circle theorem for nonlinear systems. They will not be discussed here, but will be referred to when necessary.

## 5.4 System Structure

It is assumed that the system as seen by the controller has a structure shown in Fig. 5.3.

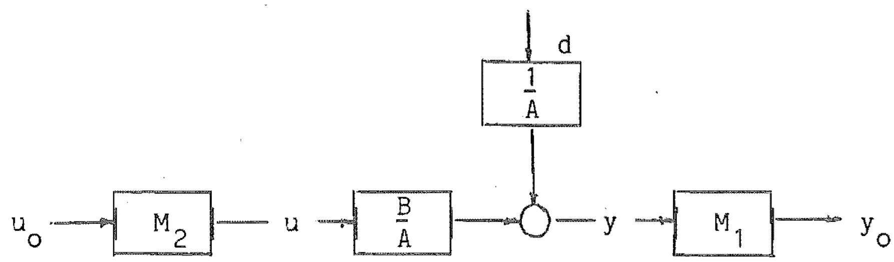



Fig. 5.3 System structure

A(d) and B(d) are polynomials in the backward shift operator d and d(t) is assumed to be a zero-mean bounded sequence (this is commonly taken to be a sequence of zero-mean uncorrelated random variables). It is assumed that y(t) is unavailable but rather $y_0(t)$ is measured. Similarly, u(t) is not directly manipulated but rather $u_0(t)$.

The system of Fig. 5.3 is general enough to include a large class of systems, as shown in the following sections.

### 5.4.1 Higher-order systems

Assume the actual system is described by a higher-order model given by

$$A_0(d)y_0(t) = B_0(d)u_0(t) + d_0(t) \tag{5.1}$$

where

$$\text{degree of } A_0(d) \geq \text{degree of } A(d)$$

$$\text{degree of } B_0(d) \geq \text{degree of } B(d)$$

The system (5.1) fits the system structure of Fig. 5.3 if

$$M_2 = 1 , \qquad M_1 = \frac{A\ B_0}{B\ A_0} , \qquad d(t) = \frac{B}{B_0}\ d_0(t) \tag{5.2}$$

where the argument d has been dropped for brevity. Also, the system (5.1) fits the structure of Fig. 5.3 if

$$M_1 = 1 , \qquad M_2 = \frac{A\ B_0}{B\ A_0} , \qquad d(t) = \frac{A}{A_0}\ d_0(t) \tag{5.3}$$

In either case, B/A represents the modelled dynamics and M contains the unmodelled dynamics.

If $d_0(t) \in l_2$, and if either $A_0$ or $B_0$ is stable then $d(t) \in l_2$. However, if $A_0$ and $B_0$ are unstable, it is assumed that $d_0(t)$ can be written as

$$d_0(t) = Qd'(t) \tag{5.4}$$

where $d'(t) \in l_2$, and Q contains the unstable roots of $A_0$ or $B_0$. The motivation for (5.4) is that in the subsequent analysis d(t) is required to be a bounded sequence.

5.4.2 Nonlinearities

When considering nonlinear systems, M may be taken to be a memoryless nonlinearity:

$$M_i = N_i \tag{5.5}$$

For example, by suitably rescaling B, $N_1$ may represent the actual output nonlinearity premultipled by a real number and $N_2$ the actual input nonlinearity postmultiplied by a real number.

## 5.5 Robustness of the Fixed Strategy

The problem of preserving stability in the presence of plant perturbations is examined is this section. It is assumed that the parameters of the nominal plant (i.e. A and B) are given and that the magnitude of any plant perturbations known. The essence of the problem is now clear: How to adjust the pole-zero polynomials to improve robust -ness in the presence of plant perturbations ?

For the analysis to be feasible, the system structure of Fig. 5.3 is considered.

It is assumed that the nominal plant is described by

$$Ay(t) = Bu(t) + d(t) \tag{5.6}$$

It is assumed that y(t) is not directly available but rather $y_0(t)$ is measured:

$$y_0(t) = M_1 y(t) \tag{5.7}$$

Alternatively, it is assumed that u(t) is not manipulated but rather $u_0(t)$, where

$$u(t) = M_2 u_0(t) \tag{5.8}$$

is applied to the plant.

The input and output errors are defined as:

$$\tilde{y}(t) = y_0(t) - y(t)$$
$$= (M_1 - 1)y(t) \tag{5.9}$$

$$\tilde{u}(t) = u_0(t) - u(t)$$
$$= (M_2^{-1} - 1)u(t) \tag{5.10}$$

Also, the actual plant is assumed to be given by a higher-order model:

$$A_0 y_0(t) = B_0 u_0(t) + d_0(t) \tag{5.11}$$

It then follows that

$$M = \frac{B_0 A}{A_0 B} \tag{5.12}$$

The following assumptions are made:

Assumption 5A

(1) $A_0(d)$ and $A(d)$ are stable polynomials

(2) $P(d)$ is a stable polynomial

In the following, the two cases of $\tilde{y}(t) = 0$ and $\tilde{u}(t) = 0$ are are considered.

5.5.1 Output error

Assume that $\tilde{u}(t) = 0$. Then the control law is given by

$$Fu(t) = G_1 A(y^*(t) - y_0(t)) \tag{5.13}$$

where

$$F = PK - G_1 B$$

$$K = \frac{B(1)}{P(1)}$$

(5.6), (5.9), (5.13) imply

$$y(t) = (APK)^{-1}(G_1 ABy^*(t) - G_1 AB\tilde{y}(t) + Fd(t)) \tag{5.14}$$

Then (5.9) and (5.14) yield the output error feedback system as shown in Fig. 5.4.

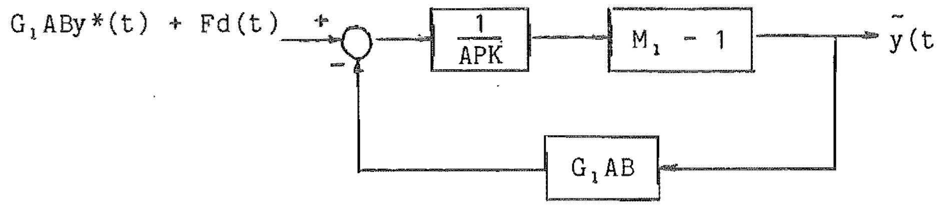



Fig. 5.4 Output error feedback system

Theorem 5.1

The feedback system of Fig. 5.4 is input-output stable if any of the following conditions is met:

(1) $M_1$ linear (as defined by (5.12))

$$\left| \frac{B_\Omega}{A_0} - \frac{B}{A} \right| < \left| \frac{P\,K}{G_1 A} \right|$$

for all $|d|=1$ (i.e., on the unit circle)

(2) $\frac{G_1 A}{P\,K}\left(\frac{B_\Omega}{A_0} - \frac{B}{A}\right)$ obeys the usual Nyquist stability criterion with respect to the point $-1+j0$.

(3) $\sup\limits_{|d|=1} \left| \frac{G_1 B}{P\,K} \right| < \frac{1}{\text{gain } (M_1 - 1)}$

Proof

(1), (3) follows the small-gain theorem and a well-known result for the gain of an $l_2$ mapping in the frequency domain.

(2) follows directly from the Nyquist criterion in the open-loop transfer function.

Theorem 5.1-(1), (3) are only sufficient, while Theorem 5.1-(2) is both sufficient and necessary.

Remarks

If $\sup\limits_{|d|=1} \left| \frac{G_1 B}{P\,K} \right| < \frac{1}{\delta}$ then the Nyquist plot of $\frac{G_1 B}{P\,K}$ must lie within a circle of radius $\frac{1}{\delta}$, centred at (0,0) in the complex plane.

It has been shown (Zames 1981) that in linear systems a control law may be decomposed into two stages, consisting of:

(i) filtering of plant uncertainty

(ii) design of a control law for the nominal plant

Fig. 5.4 can easily be decomposed as such, as shown in Fig. 5.5.

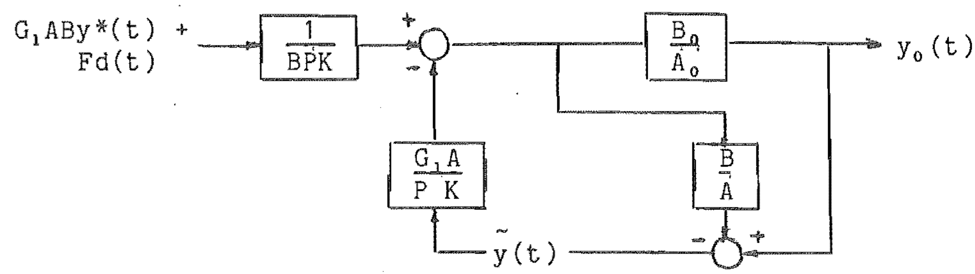



Fig. 5.5 Decomposition of feedback systems for output error case

Here, $\frac{G_1A}{P\ K}$ can be regarded as a filter on plant uncertainty, and $\frac{1}{BPK}$ a precompensator. The former is chosen for stability reasons in the presence of plant uncertainty, while the latter is designed to meet some other specifications.

When a feedback system consists of a linear subsystem and a non-linear subsystem, the circle theorem (e.g. Zames 1966) asserts that if the nonlinearity is inside a sector $(\alpha,\beta)$ and if the Nyquist locus of the linear part avoids a critical region, then the closed-loop is bounded. The following result gives a circle criterion type of interpretation of Theorem 5.1.

Theorem 5.2

If $(M_1 - 1)$ is a bounded memoryless operator in the sector $(\alpha,\beta)$, then for stability the following conditions must hold:

(1) If $0 < \alpha < \beta$, the Nyquist locus of $\frac{G_1B}{P\ K}$ does not enclose or intersect the disc $D(-\frac{1}{\alpha}, -\frac{1}{\beta})$.

(2) If $0 = \alpha < \beta$, the Nyquist locus of $\frac{G_1B}{P\ K}$ is such that

$$\mathrm{Re}\left(\frac{G_1B}{P\ K}\right) > -\frac{1}{\beta} \quad \text{for all } |d|=1$$

Proof

Note that $\frac{G_1B}{PK}$ has no poles outside the stable region, and the result can be established from p. 141, Desoer and Vidyasagar (1975).

5.5.2 Input error

Assume $\tilde{y}(t) = 0$ and $M_2^{-1}$ exists. Then the control law is given by

$$Fu_0(t) = G_1A(y^*(t) - y(t)) \tag{5.15}$$

(5.6), (5.10) and (5.15) imply

$$u(t) = (PK)^{-1}(G_1Ay^*(t) - F\tilde{u}(t) - G_1d(t)) \tag{5.16}$$

(5.10) and (5.16) yields the system of Fig. 5.6.

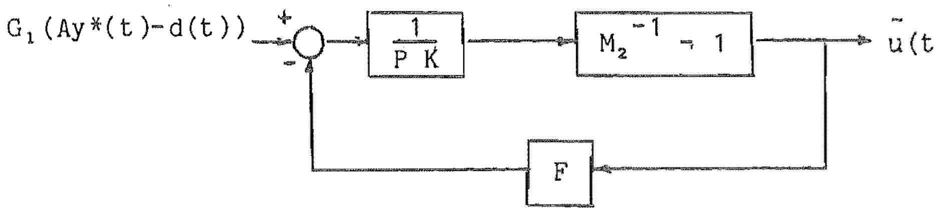



Fig. 5.6 Input error feedback system

Theorem 5.3

The feedback system of Fig. 5.6 is input-output stable if any of the following conditions holds:

(1) $M_2$ linear (as defined by (5.12))

$$\left| \frac{A_\Omega}{B_0} - \frac{A}{B} \right| < \left| \frac{A\,P\,K}{B\,F} \right| \qquad \text{for all } |d|=1$$

(2) $\frac{B\,F}{APK}\left(\frac{A_\Omega}{B_0} - \frac{A}{B}\right)$ obeys the usual Nyquist criterion with respect to the point $-1+j0$ .

(3) $\sup_{|d|=1} \left| \frac{F}{P\,K} \right| < \left[\text{gain}(M_2^{-1} - 1)\right]^{-1}$

Proof

The proof parallels that of Theorem 5.1.

In the case where $M_2$ is not necessarily invertible, an error feedback system can be obtained as shown in Fig. 5.7.

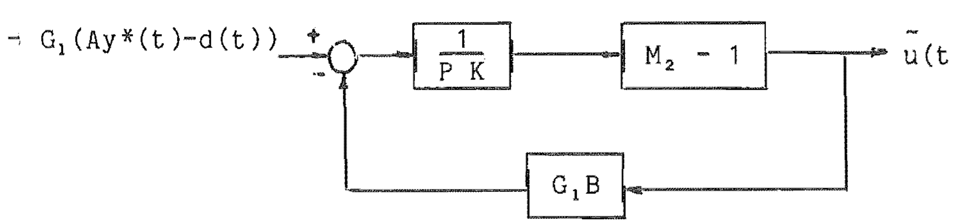



Fig. 5.7 Input error feedback system

Theorem 5.1 is equally applicable if $M_1$ is replaced by $M_2$.

## 5.6 Instability of the Fixed Strategy

In the previous section, the Nyquist criterion provides the necessary and sufficient conditions for stability of the error feedback systems. Thus instability can be shown as well. The stability results established via the small-gain theorem are only sufficient, and hence, conservative. In this section, the instability counterpart of Fig. 5.4

is given.

Theorem 5.4

If $(M_1 - 1)$ is a time-invariant, memoryless operator in the sector $(\alpha, \beta)$, $0 < \alpha < \beta$ and if the Nyquist locus of $\frac{G_1 B}{P\ K}$ does not intersect the disc

$$D\left(-\frac{1}{\alpha}, -\frac{1}{\beta}\right)$$

and encircles it at least once in the clockwise direction, then there exists some input in $l_2$ such that $\tilde{y}$, $y \notin l_2$.

Proof

Fig. 5.4 can be redrawn to consist of $(M_1 - 1)$ in the forward path and $\frac{G_1 B}{PK}$ in the feedback path. Since the latter subsystem has no unstable poles, the result can be established from p. 161, Desoer and Vidyasagar (1975).

## 5.7 Robust Disturbance Attenuation

In this section, the problem of maximizing disturbance attenuation is examined.

Consider the output error feedback system of Fig. 5.4 redrawn as shown in Fig. 5.8.

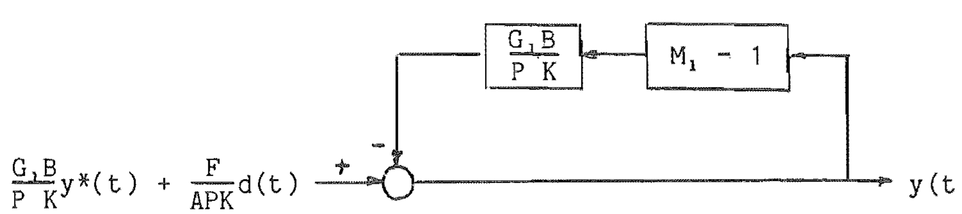



Fig. 5.8 Output error feedback system

Define the following:

$$w(t) = \frac{G_1 B}{P\ K} y^*(t) \tag{5.17}$$

$$v(t) = \frac{F}{APK}\ d(t) \tag{5.18}$$

Also define

$$H_0 = \frac{G_1 B}{P\ K}\ (M_1 - 1) \tag{5.19}$$

$$H_c = H_0 (1 + H_0)^{-1} \tag{5.20}$$

Thus from Fig. 5.8 and using (5.19), (5.20)

$$y = (1 + H_0)^{-1}(w+v)$$

$$= (1 - H_c)(w+v) \tag{5.21}$$

Rearranging,

$$y - (w+v) = - H_c(w+v) \tag{5.22}$$

Here, $(w+v)$ may be regarded as a disturbance.

The following definition is useful for the understanding of the rest of the section.

Definition: Let $x$ and $h$ be arbitrary elements of the normed space $X$ and let $F$ be an operator on $X$. If there exists an operator $A(x)$ linear in $h$ (which may depend on $x$) such that

$$F(x+h) - F(x) = A(x)h + A(x,h)$$

where

$$\lim_{||h||\to 0} \frac{||A(x,h)||}{||h||} = 0$$

then $F$ is said to be Frechet differentiable at $x$, and $A(x)h$ is called the Frechet differential of $F$ at $x$ for incremental $h$. The linear operator $A(x)$ is called the Frechet derivative of $F$ at $x$.

Theorem 5.5

If the plant is not necessarily linear and assume that it is Frechet differentiable, then

$$y - (w+v) = - H_c(w+v)$$

$$= - H_c(v) - \int_0^1 DH_c(v+\gamma w)w\,d\gamma$$

where $DH_c$ is the Frechet derivative of $H_c$.

Proof

Using the identity

$$f(a+h) - f(a) = \int_0^h f'(a+h-t)dt$$

where $f'$ is the derivative of the function with respect to its argument $(a+h-t)$ gives

$$f(a+h) - f(a) = \int_0^1 f'(a+\gamma h)h\,d\gamma$$

From Theorem 5.5,

$$||y - (w+v)|| \leq ||H_c(v)|| + ||\int_0^1 DH_c(v+\gamma w)\, d\gamma||$$

$$\leq ||H_c(v)|| + \sup_{0<\gamma<1} ||DH_c(v+\gamma w)||\,||w|| \qquad (5.23)$$

A reasonable strategy would be to minimize both terms on the right-hand side of (5.23).

From (5.20),

$$||H_c|| \leq ||H_0||\ ||(1 + H_0)^{-1}||$$

$$\leq ||H_0||\ (1 - ||H_0||)^{-1} \qquad (5.24)$$

since for robust stability $||H_0|| < 1$ (see Theorem 5.1). Also,

$$DH_c = D(H_0(1 + H_0)^{-1})$$

$$= DH_0(1 + DH_0)^{-1} \qquad (5.25)$$

where

$$DH_0 = \frac{G_1 B}{P\ K}\,(D(M_1) - 1) \quad \text{using (5.19)} \qquad (5.26)$$

Thus

$$\sup_{0<\gamma<1} ||DH_0|| = \sup_{|d|=1} \left|\frac{G_1 B}{P\ K}\right| \sup_{0<\gamma<1} ||D(M_1) - 1||$$

$$= \left|\left|\frac{G_1 B}{P\ K}\right|\right|\ ||M_1 - 1||_\Delta \qquad (5.27)$$

where $||.||_\Delta$ is the incremental norm.

By definition, $DH_0$ is a linear operator. Thus, the well-known M-contours (Ogata 1970) can be used to design for minimum $||DH_c||$ from the Nyquist plot of $DH_0$. It is now clear that minimizing the norm of $\frac{G_1 B}{P\ K}$ minimizes the two terms on the right-hand side of (5.23). However, $||y||$ is not the minimum value possible.

When the plant perturbations are linear, and since the Frechet derivative of a linear operator is the linear operator itself, the result below follows immediately from Theorem 5.5.

Theorem 5.6

If (w+v) is regarded as a disturbance, then minimizing the gain of the closed-loop operator $H_c$ at each frequency maximizes the attenuation of (w+v) at that frequency.

For linear plant perturbations,

$$H_c = \frac{G_1(AB_0 - BA_0)}{A_0PK + G_1(AB_0 - BA_0)} \tag{5.28}$$

The pole-zero polynomials P and $G_1$ should be selected such that $H_c$ is minimized over the frequency range of the disturbance. Also this can be done via the M-contours and the Nyquist locus of $H_0$ .

## 5.8 Examples

In this section, examples are given to illustrate the results obtained in the previous sections.

### 5.8.1 Linear systems

Consider a continuous-time multiplicatively perturbed plant given by

$$G_0(s) = G(s)(1 - \Delta G(s))$$

where

nominal plant: $$G(s) = \frac{1}{s + 1}$$

perturbation: $$\Delta G(s) = \frac{\mu s}{\mu s + 1}$$

Assuming an input zero-order hold with a sampling period of 0.5 sec the resulting discrete-time nominal plant is given by

$$G(d) = \frac{0.3935d}{1 - 0.6065d}$$

Figs. 5.9 - 5.11 show the Nyquist loci ($0 < \omega < \pi$) of

$$H_0 = \frac{AB_0 - BA_0}{A_0PK}$$

for various P(d)'s and $\mu$ 's, where $0 < \mu < 1$.

In Fig. 5.9 the closed-loop system is stable for $\mu \leq 0.3$ with P = 1.0. In Fig. 5.10 - 5.11 the robust stability margin has increased with a higher order P. Thus using a higher order P is good for improving robust stability. Of course, the order of P is normally less than or equal to 2.

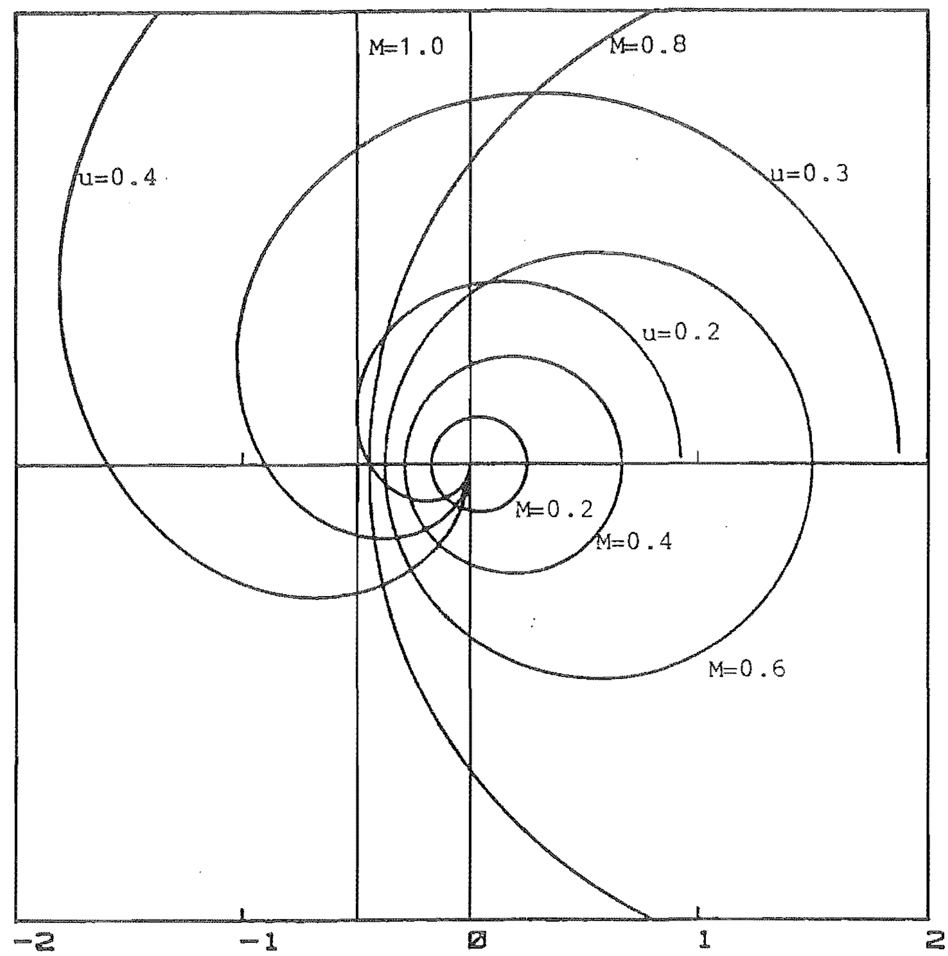



Fig. 5.9 Nyquist plot of Ho: P(d)=1; u=0.2, 0.3, 0.4

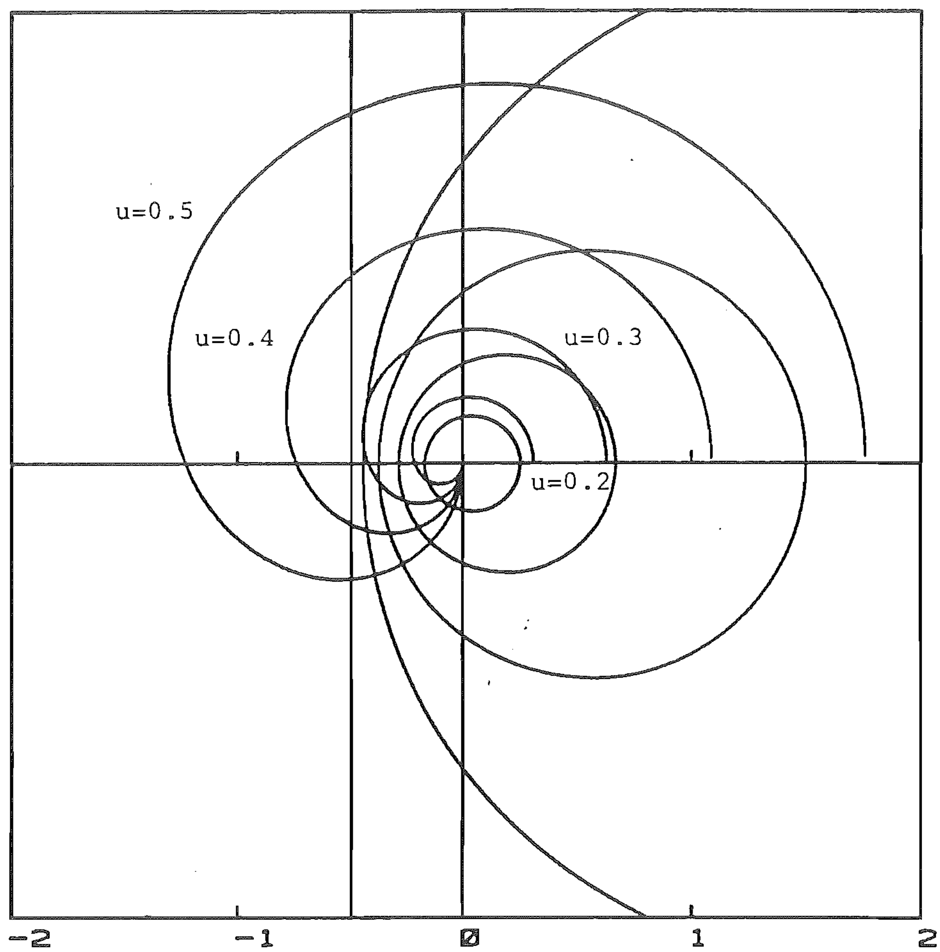



Fig. 5.10 Nyquist plot of Ho: P(d)=1-0.5d;
u=0.2, 0.3, 0.4, 0.5

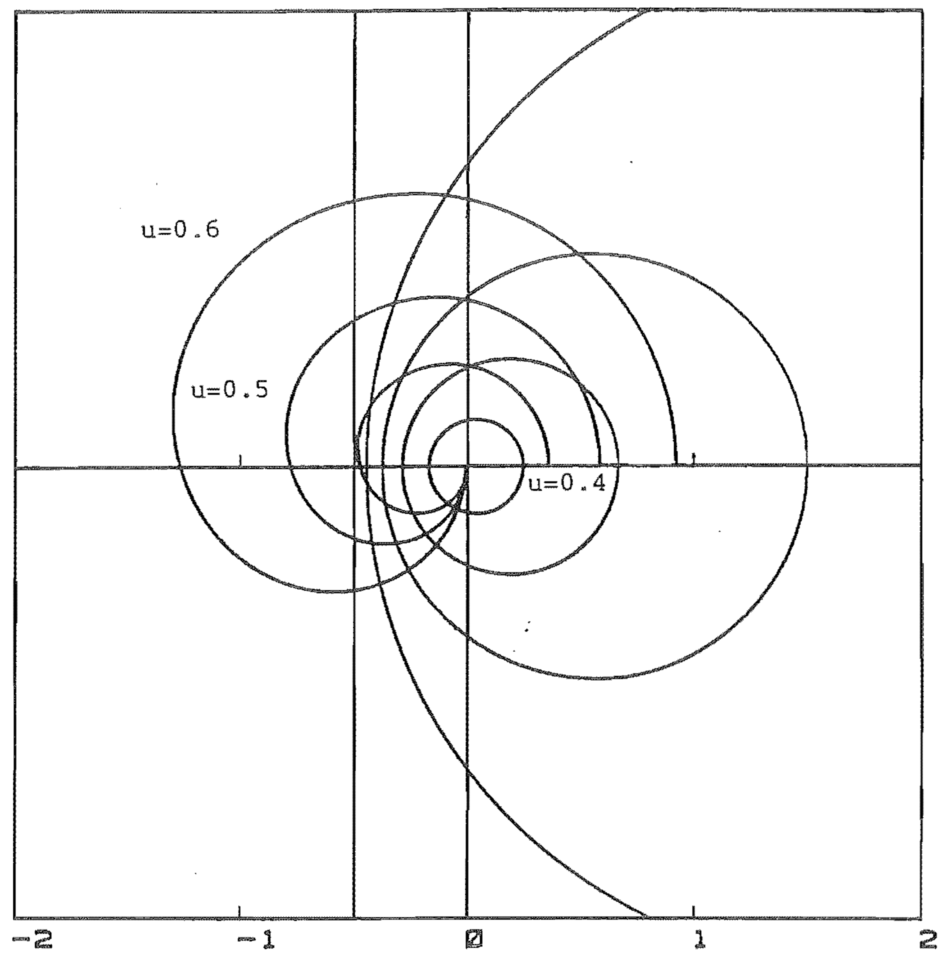



Fig. 5.11 Nyquist plot of Ho: $P(d)=1-d+0.25d^2$;
u=0.4, 0.5, 0.6

Fig. 5.9 – 5.11 also show the M-contours for M = 0.2, 0.4, 0.6, 0.8, and 1.0. Now, it is of interest to obtain the largest frequency ranges over which $|H_c|$ is less than some maximum value, say, 1. In Fig. 5.9, for the case $\mu = 0.2$ the frequency ranges over which $|H_c| > 1$ are

$$0.211\omega_s < \omega < 0.2278\omega_s$$

$$0.772\omega_s < \omega < 0.7917\omega_s$$

where $\omega_s$ is the sampling frequency. For $\mu = 0.3$, the frequency ranges now become

$$0.1167\omega_s < \omega < 0.3056\omega_s$$

$$0.6944\omega_s < \omega < 0.8833\omega_s$$

It is seen that more high frequency disturbances are amplified for larger plant perturbations. With the same magnitude of plant perturbation, the disturbance attenuation is better for higher order P. Thus an appropriate choice of P serves to improve robust stability and robust disturbance attenuation.

The role of $G_1$ is also of significance as it helps to improve robust stability, as seen from Fig. 5.12.

5.8.2 Nonlinear system

Suppose a nonlinearity defined by Fig. 5.13 exists at the output of a first-order nominal system given by

$$G(d) = \frac{d(1 + 2d)}{1 - 0.7d}$$

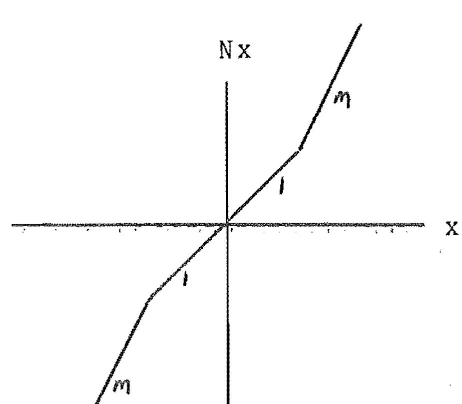



Fig. 5.13 Nonlinearity

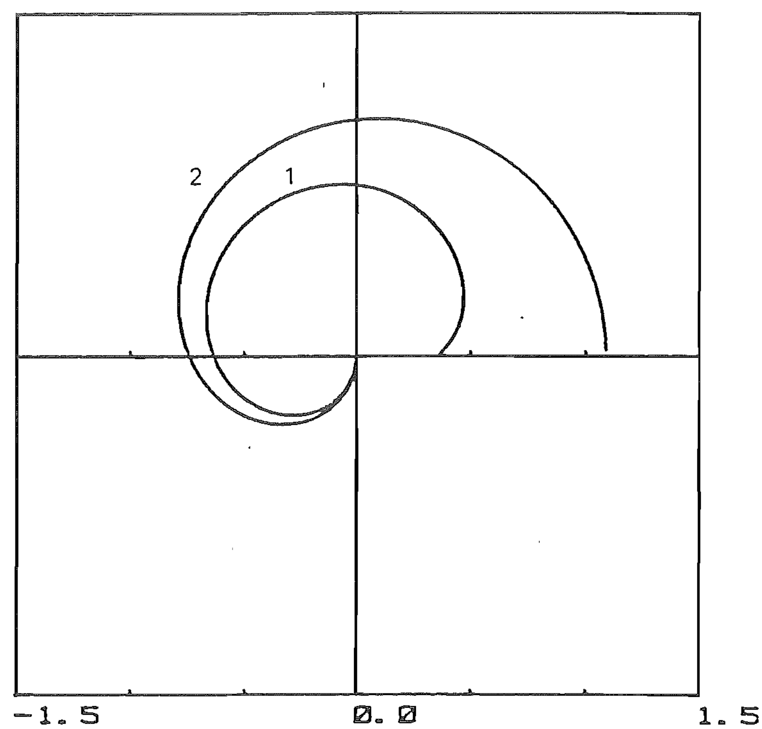



Fig. 5.12 Nyquist plot of Ho: u=0.4; P(d)=1-0.5d

1 -- $G_1(d)=(1+0.5d)/1.5$

2 -- $G_1(d)=1$

Fig. 5.14 shows the Nyquist plot of $\frac{B}{P\,K}$ for various P's. For stability, Theorem 5.1 requires that

$$\sup_{|d|=1} \left| \frac{B}{P\,K} \right| < \frac{1}{\text{gain }(N-1)} = \frac{1}{m-1}$$

This is equivalent to requiring the Nyquist locus of $\frac{B}{P\,K}$ to be within a circle of radius given by $1/(m-1)$. Thus the radius of the circle must be greater than 1 (see Fig. 5.14), which implies that $1 < m < 2$.

If the nonlinearity defined by Fig. 5.13 is at the input of the plant, an identical condition is obtained (see Fig. 5.7).

If $(M_1 - 1)$ is in the sector $(0,\beta)$, Theorem 5.2 states that the closed-loop system will be stable if

$$\text{Re}\left(\frac{B}{P\,K}\right) > -\frac{1}{\beta} \quad \text{for all } |d| = 1$$

Fig. 5.14 shows that, for a fixed $\beta$, the robust stability can be improved by the use of higher order P. The following table shows the variation of the sector bounds with P.

| P | $(M_1 - 1)$ | $\lVert\frac{B}{P\,K}\rVert$ |
|---|---|---|
| 1 | (0, 1.4546) | 1.0 |
| $1 - 0.5d$ | (0, 2.662) | 1.0 |
| $(1 - 0.5d)^2$ | (0, 3.333) | 1.0 |

It is seen that the robustness sectors for $(M_1 - 1)$ increases with a higher order P. Note that the norm of $\frac{B}{P\,K}$ does not change. This is because at zero frequency, the magnitude of the frequency response is greatest. But at other frequencies the magnitude does change, as clear -ly seen from Fig. 5.14.

Also, instability of the closed-loop system can be shown using Theorem 5.4 for $(M_1 - 1)$ in the complementary sectors. For example, for $P = 1$, if $(M_1 - 1)$ in the sector $(1.4546,\infty)$, then there exists an input in $l_2$ which will lead to instability.

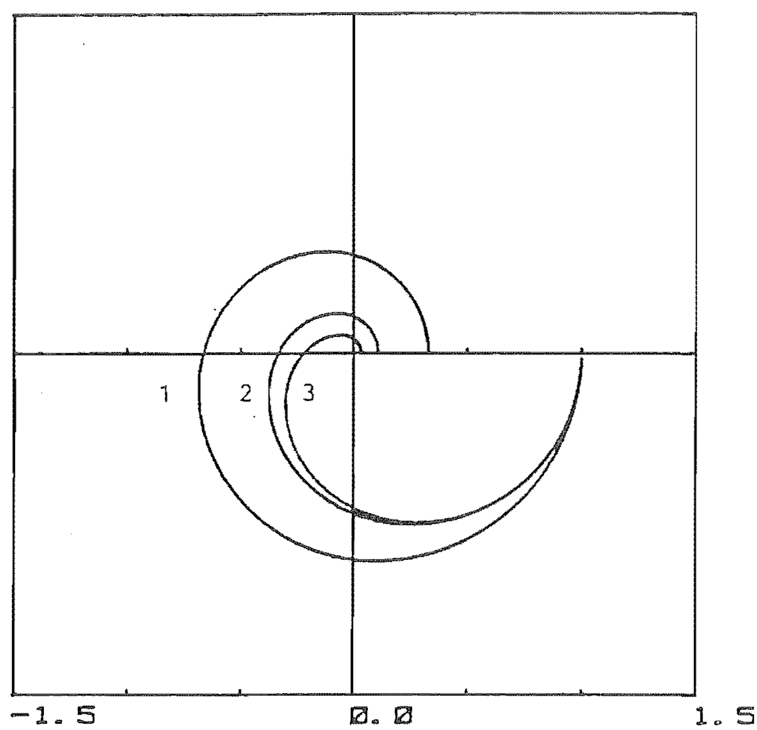



Fig. 5.14 Nyquist plot of B/(PK) for

1 -- $P(d)=1$
2 -- $P(d)=1-0.5d$
3 -- $P(d)=1-d+0.25d^2$

## 5.9 Concluding Remarks

In this chapter some frequency domain conditions for the robust-ness of the pole-zero placement controller with respect to some plant perturbations have been given. The results provide engineering guide-lines as to how the pole-zero polynomials may be selected to improve robustness. They also indicate the types of plant perurbations that are likely to cause difficulties.

The results were for the nonadaptive case, but intuitively must also be relevant for the adaptive case. The robustness of the adaptive control algorithm is addressed in the next chapter.

# CHAPTER 6

# ROBUSTNESS: ADAPTIVE CASE

## 6.1 Introduction

In the previous chapters, deterministic and stochastic versions of the explicit pole-zero placement algorithm have been developed and stability and convergence for a class of nonminimum phase systems have also been demonstrated. A key feature of the algorithm is that it avoids the solution of a Diophantine equation. Consequently, the algorithm is not only computationally efficient but also side steps the estimated stabilizability problem.

However, stability and convergence have been established under somewhat ideal conditions. Specifically, the unknown plant is linear, time-invariant and the upper bound of the system order is assumed to be known. Rohrs et al (1985) have, in fact, shown that a stable adaptive control algorithm is not necessarily stable in the presence of unmodelled dynamics. That is, it is not robustly stable.

In this chapter, the ideas of Praly (1983a, b) are extended to derive a robust version of the pole-zero placement algorithm described previously, and its robustness with respect to unmodelled dynamics and bounded external disturbances is studied.

This chapter is organized as follows. Section 6.2 introduces the modelling error associated with the plant and the control objective in the presence of modelling errors. Section 6.3 presents the adaptation algorithm and the basic properties of the parameter estimator. The properties of the adaptive controller are given in Section 6.4. In Section 6.5 conditions for stability are given.

## 6.2 The Plant and Control Objective

Let $u(t)$ and $y(t)$ be the plant input and output, respectively. To model the input and output behaviour of the plant, define an n-th order model:

$$A(d)y(t) = B(d)u(t) \tag{6.1}$$

where $A(d)$ and $B(d)$ are polynomials in the unit delay operator d defined as

$$A(d) = 1 + a_1 d + \ldots\ldots + a_n d^n$$

$$B(d) = b_1 d + \ldots\ldots + b_n d^n$$

(6.1) can be written as

$$y(t) = \phi(t-1)^T \theta_0 \tag{6.2}$$

where

$$\phi(t-1) = [\ y(t-1),\ldots\ldots,y(t-n),u(t-1),\ldots\ldots,u(t-n)\ ]^T$$

$$\theta_0 = [\ -a_1,\ldots\ldots,-a_n,b_1,\ldots\ldots,b_n\ ]^T$$

Associated with these chosen polynomials A(d) and B(d), the modelling error is defined as

$$w(t) = A(d)y(t) - B(d)u(t) \tag{6.3}$$

Define a sequence of nonnegative real numbers $\{s(t)\}$ by

$$s(t)^2 = \sigma^2 s(t-1)^2 + \max(s^2,\ ||\phi(t-1)||^2) \tag{6.4}$$

$$0 < \sigma < 1, \quad s > 0$$

The modelling error is said to be relatively bounded if there exists a finite $\eta$ such that

$$\frac{|w(t)|}{s(t)} \leqq \eta \tag{6.5}$$

Note that a relatively bounded modelling error is not guaranteed to be bounded unless the input and output signals are bounded.

Example

Suppose the true plant is represented by

$$y(t) = \psi(t-1)^T \Theta + \omega(t) \tag{6.6}$$

where

$$\psi(t-1) = [\ y(t-1),\ldots,y(t-N),u(t-1),\ldots,u(t-N)\ ]^T$$

$\Theta$ is a vector representing the parameters of the plant,

$\omega(t)$ is an external disturbance.

Comparing (6.1) with (6.6) and using (6.3) gives

$$w(t) = \psi(t-1)^T \Delta\Theta + \omega(t) \tag{6.7}$$

where $\Delta\Theta$ is the difference between $\Theta$ and $\theta$ (dimension augmented).

Now, suppose that for $t \geqq 0$,

$$||\Delta\Theta|| \leqq \eta_1 \tag{6.8}$$

$$|\omega(t)| \leqq m_1 \tag{6.9}$$

(6.8) implies that some fast poles or zeros of the parameters have been neglected in (6.1). Using (6.7) and (6.8) - (6.9) gives

$$\frac{|w(t)|}{||\psi(t-1)||} \leqq \eta_1 + \frac{m_1}{||\psi(t-1)||} \tag{6.10}$$

However, $||\psi(t-1)||$ is not available because the true plant is unknown. Now introduce the sequence in (6.4). Noting that

$$||\psi(t-1)||^2 \leqq \sum_{i=0}^{N} ||\phi(t-1-i)||^2 \tag{6.11}$$

and

$$\begin{aligned} s(t)^2 &\geqq \sum_{i=0}^{N} \sigma^{2i}\max(s^2, ||\phi(t-1-i)||^2) \\ &\geqq \sigma^{2N}\max((N+1)s^2, ||\psi(t-1)||^2) \\ &\geqq \sigma^{2N}||\psi(t-1)||^2 \end{aligned} \tag{6.12}$$

and since $s(t) \geqq s$ for an arbitrary large s, (6.10) becomes

$$\begin{aligned} \frac{|w(t)|}{s(t)} &\leqq \frac{\eta_1}{\sigma^N} + \frac{m_1}{s} \\ &\leqq \eta_w \end{aligned} \tag{6.13}$$

Hence, the modelling error is relatively bounded if (6.8) - (6.9) hold.

Motivated by the above example, it is assumed that the true plant can be represented by

$$A(d)y(t) = B(d)u(t) + w(t) \tag{6.14}$$

$$s(t)^2 = \sigma^2 s(t-1)^2 + \max(s^2, ||\phi(t-1)||^2) \tag{6.15}$$

$$0 < \sigma < 1, \quad s > 0$$

Thus (6.14) can be written as

$$y(t) = \phi(t-1)^T\theta_0 + w(t) \tag{6.16}$$

The following assumptions about the plant are made:

Assumption 6A

Given a vector $\theta_s \varepsilon R^{2n}$ such that $A_s(d)$ is stable and a scalar $\rho_0$, there exists an unknown vector $\theta_0 \varepsilon R^{2n}$ such that:

(1) $||\theta_0 - \theta_s|| \leqq \rho_0$

(2) $A(d)$ is stable

(3) $\frac{|w(t)|}{s(t)} \leq n_w$ for some $n_w$

After the above discussions, the control objective may be stated as follows. Given a model of the unknown but stable plant, the object is to find a pole-zero controller so that, despite the presence of the modelling error, acceptable behaviour of the input and output signals is retained.

## 6.3 The Adaptation Algorithm

The adaptation algorithm consists of two parts, one of which is a parameter estimation algorithm, the other a control law calculation.

### 6.3.1 The control calculation

In the following, a circumflex means estimated value.

The control signal is determined from the control law

$$\hat{F}(t,d)u(t) = \hat{G}(t,d)\left(y^*(t) - y(t)\right) \tag{6.17}$$

where $y^*(t)$ is the reference signal, assumed to be a series of steps, and

$$\hat{G}(t,d) = G_1(d)\hat{A}(t,d)$$

$$\hat{F}(t,d) = P(d)\hat{K}(t) - G_1(d)\hat{B}(t,d)$$

$$\hat{K}(t) = \begin{cases} \frac{\hat{B}(t,1)}{P(1)} & \text{if } |\hat{B}(t,1)| > \varepsilon \\ \\ \frac{\varepsilon}{P(1)} & \text{otherwise} \end{cases} \tag{6.18}$$

$\varepsilon$ is a positive scalar.

The following additional assumptions are made:

Assumption 6B

(1) $P(z^{-1})$ has all roots inside the unit circle

(2) $|y^*(t)| < M_1 < \infty$

6.3.2 Parameter estimation algorithm

The parameter estimation algorithm to be considered is a modified version of the algorithm used by Praly (1983). It is described as follows:

$$\hat{\theta}'(t) = \hat{\theta}(t-1) + \frac{P(t-2)\phi(t-1)}{\mu s(t)^2 + \phi(t-1)^T P(t-2)\phi(t-1)} e(t) \tag{6.19}$$

where

$$e(t) = y(t) - \phi(t-1)^T \hat{\theta}(t-1) \tag{6.20}$$

$\mu$ is a strictly positive scalar.

$$P(t-1) = \left(1 - \frac{\lambda_0}{\lambda_1}\right)\left[P(t-2) - \frac{P(t-2)\phi(t-1)\phi(t-1)^T P(t-2)}{\mu s(t)^2 + \phi(t-1)^T P(t-2)\phi(t-1)}\right] + \lambda_0 I \tag{6.21}$$

$$\hat{\theta}''(t) = \theta_s + \min\left(1, \frac{\rho_0\sqrt{\frac{\lambda_0}{\lambda_1}}}{||\hat{\theta}'(t) - \theta_s||}\right)(\hat{\theta}'(t) - \theta_s) \tag{6.22}$$

The symbols in (6.19) have the following meaning:

$$\hat{\theta}(t) = [\ -\hat{a}_1(t),\dots,-\hat{a}_n(t),\hat{b}_1(t),\dots,\hat{b}_n(t)\ ]^T$$

$$\hat{A}(t,d) = 1 + \hat{a}_1(t)d + \dots\dots + \hat{a}_n(t)d^n$$

$$\hat{B}(t,d) = \hat{b}_1(t)d + \dots\dots + \hat{b}_n(t)d^n$$

The parameter vector $\hat{\theta}''(t)$ is modified according to the following projection facility:

$$\hat{\theta}(t) = \begin{cases} \hat{\theta}''(t), & \text{if } \hat{\theta}''(t) \in C \\ \\ \hat{\theta}^*(t), & \text{if } \hat{\theta}''(t) \notin C \end{cases} \tag{6.23}$$

where C is a closed-convex set satisfying:

(1) $\theta_0,\ \theta_s \in C$

(2) $C \subset \{\ \theta(t)\colon\ \hat{\rho}_i(t) = 1 - \rho < 1,\ i=1,\dots,n$ (6.24)

$\hat{\rho}_i(t)$ are the roots of $\hat{A}(t,d)\ \}$

In (6.21), P(t) is a positive symmetric definite matrix whose eigenvalues satisfy:

$$0 < \lambda_0 \leq \lambda_{min}(P(t)) \leq \lambda_{max}(P(t)) \leq \lambda_1 \tag{6.25}$$

where $\lambda_{min}$( $\lambda_{max}$ ) is the minimum (maximum) eigenvalue of P(t).

To choose P(0) and $\mu$, note that in the stochastic estimation context, P(0) corresponds to the a priori parameter error covariance, and $\mu$ the covariance of $\frac{|w(t)|}{s(t)}$.

$\hat{\theta}^*(t)$ in (6.23) can be computed following the simple scheme in chapter 3.

The basic properties of the algorithm (6.19)-(6.23) are summarized in the following lemma.

Lemma 6.1

Subject to Assumption 6A and (6.14), the algorithm (6.19)-(6.23) ensures that:

(1) $||\hat{\theta}(t)|| \leq M_2$

(2) $\sum_{t=q+1}^{q+k} \frac{|e(t)|}{s(t)} \leq \sqrt{k}M_3 + kL_1n_W$ for all (q,k)

(3) $||\hat{\theta}(t) - \hat{\theta}(t-1)|| \leq L_2 \frac{|e(t)|}{s(t)}$

where $M_2$ and $M_3$ are positive constants independent of $n_W$, and

$$L_1^2 = 1 + \frac{\lambda_1}{\mu}$$

$$L_2 = 2 \frac{\lambda_1}{\mu + \lambda_0}$$

Proof

In the sequel, $(\bar{.})$ will be used to denote a normalized variable and is defined as:

$$(\bar{.}) = s(t)^{-1}(.)(t)$$

Define the following errors:

$$\tilde{\theta}(t) = \hat{\theta}(t) - \theta_0 \tag{6.26}$$

$$\tilde{\theta}'(t) = \hat{\theta}'(t) - \theta_0 \tag{6.27}$$

$$\tilde{\theta}''(t) = \hat{\theta}''(t) - \theta_0 \tag{6.28}$$

$$\tilde{\theta}'_S(t) = \hat{\theta}'(t) - \theta_S \tag{6.29}$$

$$\tilde{\theta}''_S(t) = \hat{\theta}''(t) - \theta_S \tag{6.30}$$

$$\tilde{\theta}_0 = \theta_0 - \theta_S \tag{6.31}$$

Also define

$$\Psi'(t) = \tilde{\theta}'(t)^T P'(t-1)^{-1}\, \tilde{\theta}'(t) \tag{6.32}$$

$$P'(t) = P(t-1) - \frac{P(t-1)\bar{\phi}(t)\bar{\phi}(t)^T P(t-1)}{\mu + \sigma(t)} \tag{6.33}$$

$$\sigma(t) = \bar{\phi}(t)^T P(t-1)\bar{\phi}(t) \tag{6.34}$$

$$V(t) = \tilde{\theta}(t)^T P(t-1)^{-1}\, \tilde{\theta}(t) \tag{6.35}$$

$$V'(t) = \tilde{\theta}'(t)^T P(t-1)^{-1}\, \tilde{\theta}'(t) \tag{6.36}$$

$$V''(t) = \tilde{\theta}''(t)^T P(t-1)^{-1}\, \tilde{\theta}''(t) \tag{6.37}$$

Some preliminary properties of the algorithm (6.19)-(6.23) are summarized in the following lemma.

Lemma 6.2

Subject to Assumption 6A and (6.14), the algorithm (6.19)-(6.23) has the following properties:

(1) $\Psi'(t) = V(t-1) + \frac{1}{\mu}\left(\bar{z}(t)^2 + 2\bar{z}(t)\bar{e}(t) + \Pi(t-1)\bar{e}(t)^2\right)$

where

$$\bar{z}(t) = \bar{w}(t) - \bar{e}(t)$$

$$\Pi(t) = \frac{\sigma(t)}{\mu + \sigma(t)}$$

(2) $V(t) \leq \Psi'(t)$

(3) $||\hat{\theta}(t) - \theta_0|| \leq \rho_0 \left(1 + \sqrt{\frac{\lambda_0}{\lambda_1}}\right)$

(4) $||\hat{\theta}(t) - \hat{\theta}(t-1)|| \leq 2\, \frac{||P(t-2)\bar{\phi}(t-1)||\,|\bar{e}(t)|}{\mu + \sigma(t-1)} \sqrt{\frac{\lambda_1}{\lambda_0}}$

Proof

(1) Subtracting $\theta_0$ from both sides of (6.19) gives

$$\tilde{\theta}'(t) = \tilde{\theta}(t-1) + \frac{P(t-2)\bar{\phi}(t-1)}{\mu + \sigma(t-1)} \bar{e}(t) \tag{6.38}$$

Note that from (6.33)

$$P'(t)^{-1} = P(t-1)^{-1} + \mu^{-1} \bar{\phi}(t)\bar{\phi}(t)^T \tag{6.39}$$

Use (6.16), (6.20), and (6.26) to obtain

$$\bar{\phi}(t-1)^T \tilde{\theta}(t-1) = \bar{\phi}(t-1)^T \hat{\theta}(t-1) - (\bar{y}(t) - \bar{w}(t))$$

$$= \bar{w}(t) - \bar{e}(t) \tag{6.40}$$

(6.38) - (6.40) imply

$$\mathcal{V}'(t) = V(t-1) + \mu^{-1} \left[\bar{z}(t)^2 + 2\bar{z}(t)\bar{e}(t) + \Pi(t-1)\bar{e}(t)^2\right] \tag{6.41}$$

(2) Clearly, from (6.21) and (6.33)

$$P(t) \geq P'(t) \tag{6.42}$$

Thus using the above inequality yields

$$V'(t) \leq \mathcal{V}'(t) \tag{6.43}$$

From (6.22), it is seen that

$$\min\left(1, \frac{\rho}{||\hat{\theta}'(t) - \theta_s||}\right) \leq 1 \tag{6.44}$$

Therefore, there are two cases to be considered.

Case (1)

Consider the case where

$$||\hat{\theta}'(t) - \theta_s|| \leq \rho \tag{6.45}$$

Hence,

$$V''(t) \leq \mathcal{V}'(t) \qquad \text{using (6.43)} \tag{6.46}$$

$$||\hat{\theta}''(t) - \hat{\theta}(t-1)|| = ||\hat{\theta}'(t) - \hat{\theta}(t-1)|| \tag{6.47}$$

Due to the projection facility (see Chapter 3)

$$V(t) \leq V''(t)$$

$$\leq \mathcal{V}'(t) \quad \text{using (6.46)} \tag{6.48}$$

$$||\hat{\theta}(t) - \theta_0|| \leq ||\hat{\theta}''(t) - \theta_0|| \sqrt{\frac{\lambda_1}{\lambda_0}} \tag{6.49}$$

and it follows that

$$||\hat{\theta}(t) - \hat{\theta}(t-1)|| \leq ||\hat{\theta}'(t) - \hat{\theta}(t-1)|| \sqrt{\frac{\lambda_1}{\lambda_0}} \tag{6.50}$$

Using (6.19) leads to

$$||\hat{\theta}(t) - \hat{\theta}(t-1)|| \leq \frac{||P(t-2)\bar{\phi}(t-1)|| \, |\bar{e}(t)|}{\mu + \sigma(t-1)} \sqrt{\frac{\lambda_1}{\lambda_0}} \tag{6.51}$$

Case (2)

$$\text{Consider} \quad r(t) = \frac{\rho}{||\hat{\theta}'(t) - \theta_s||} < 1 \tag{6.52}$$

where

$$\rho = \rho_0 \sqrt{\frac{\lambda_0}{\lambda_1}}$$

From (6.22)

$$\tilde{\theta}''_s(t) = r(t)\tilde{\theta}'_s(t) \tag{6.53}$$

Now, using (6.29) and (6.31) in (6.36) gives

$$\begin{aligned} V'(t) &= \tilde{\theta}'(t)^T P(t-1)^{-1} \tilde{\theta}'(t) \\ &= (\tilde{\theta}'_s(t) - \tilde{\theta}_0)^T P(t-1)^{-1} (\tilde{\theta}'_s(t) - \tilde{\theta}_0) \end{aligned} \tag{6.54}$$

Subtracting $V''(t)$ from both sides of (6.54) and using (6.30) lead to

$$\begin{aligned} V'(t) - V''(t) &= (\tilde{\theta}'_s(t) - \tilde{\theta}_0)^T P(t-1)^{-1} (\tilde{\theta}'_s(t) - \tilde{\theta}_0) \\ &\quad - (\tilde{\theta}''_s(t) - \tilde{\theta}_0)^T P(t-1)^{-1} (\tilde{\theta}''_s(t) - \tilde{\theta}_0) \end{aligned} \tag{6.55}$$

Using (6.53) gives

$$\begin{aligned} V'(t) - V''(t) &= (1 - r(t))[(1 + r(t))\tilde{\theta}'_s(t)^T P(t-1)^{-1} \tilde{\theta}'_s(t) \\ &\quad - 2\, \tilde{\theta}'_s(t)^T P(t-1)^{-1} \tilde{\theta}_0 \,] \end{aligned} \tag{6.56}$$

The aim here is to show that $V'(t) - V''(t) \geq 0$. Clearly, if the term 2 (.) is less than or equal to zero, then the problem is solved. The problem now reduces to showing that $V'(t) - V''(t) \geq 0$ for a positive term 2 (.). Now use the following inequality

$$x^T Py \leq (x^T Px)^{1/2} (y^T Py)^{1/2}$$

to obtain

$$V'(t) - V''(t) \geq (1 - r(t))(\tilde{\theta}'_s(t)^T P(t-1)^{-1} \tilde{\theta}'_s(t))^{1/2} \kappa(t) \tag{6.57}$$

where

$$\kappa(t) = (1 + r(t))(\tilde{\theta}'_s(t)^T P(t-1)^{-1} \tilde{\theta}'_s(t))^{1/2} - 2(\tilde{\theta}_0^T P(t-1)^{-1} \tilde{\theta}_0)^{1/2}$$

Now, Assumption 6A-(1) leads to

$$\begin{aligned} 2\sqrt{\frac{\lambda_1}{\lambda_0}}\,||\tilde{\theta}_0|| &\leqq 2\rho_0\sqrt{\frac{\lambda_1}{\lambda_0}} \\ &\leqq 2\rho \;(= \rho + \rho) \\ &\leqq (1 + r(t))||\tilde{\theta}'_s(t)|| \end{aligned} \tag{6.58}$$

after making use of (6.52) and the following inequality

$$r(t)||\tilde{\theta}'_s(t)|| = \rho < ||\tilde{\theta}'_s(t)||$$

With the help of the following relations

$$\tilde{\theta}_0^T P(t-1)^{-1} \tilde{\theta}_0 \leqq \lambda_{max}(P(t-1)^{-1})||\tilde{\theta}_0||^2$$

$$\lambda_{min}(P(t-1)^{-1})||\tilde{\theta}'_s(t)||^2 \leqq \tilde{\theta}'_s(t)^T P(t-1)^{-1} \tilde{\theta}'_s(t)$$

$$\lambda_{max}(P(t-1)) = (\lambda_{min}(P(t-1)^{-1}))^{-1}$$

$$\lambda_{min}(P(t-1)) = (\lambda_{max}(P(t-1)^{-1}))^{-1}$$

and (6.25), it follows that

$$2(\tilde{\theta}_0^T P(t-1)^{-1} \tilde{\theta}_0)^{1/2} \leqq (1 + r(t))(\tilde{\theta}'_s(t)^T P(t-1)^{-1} \tilde{\theta}'_s(t))^{1/2} \tag{6.59}$$

Thus with (6.59), (6.57) leads to

$$V''(t) \leqq V'(t) \tag{6.60}$$

(6.43), (6.60) and the projection facility lead to

$$V(t) \leqq V'(t) \tag{6.61}$$

Lemma 6.2-(2) follows immediately from (6.48) and (6.61).

(3) From (6.22)

$$\begin{aligned} ||\hat{\theta}(t) - \theta_0|| &= ||(\theta_s - \theta_0) + \frac{\rho}{||\hat{\theta}'(t) - \theta_s||}(\hat{\theta}'(t) - \theta_s)|| \\ &\leqq ||\theta_s - \theta_0|| + \rho \\ &\leqq \rho_0 + \rho \end{aligned} \tag{6.62}$$

where Assumption 6A-(1) has been used.

(4) From (6.22)

$$\hat{\theta}''(t) - \hat{\theta}'(t) = -\left(1 - \frac{\rho}{||\hat{\theta}'(t) - \theta_s||}\right)(\hat{\theta}'(t) - \theta_s) \tag{6.63}$$

Taking norms leads to

$$||\hat{\theta}''(t) - \hat{\theta}'(t)|| = -\rho + ||\hat{\theta}'(t) - \theta_s||$$

Using (6.19) gives

$$||\hat{\theta}''(t) - \hat{\theta}(t-1)|| \leqq -\rho + ||\hat{\theta}(t-1) - \theta_s||$$

$$+ 2\frac{||P(t-2)\bar{\phi}(t-1)||}{(\mu + \sigma(t-1))}|\bar{e}(t)| \tag{6.64}$$

Due to the projection facility,

$$||\hat{\theta}(t) - \hat{\theta}(t-1)|| \leqq ||\hat{\theta}''(t) - \hat{\theta}(t-1)||\sqrt{\frac{\lambda_1}{\lambda_0}} \tag{6.65}$$

$$||\hat{\theta}(t) - \theta_s|| \leqq ||\hat{\theta}''(t) - \theta_s||\sqrt{\frac{\lambda_1}{\lambda_0}} \tag{6.66}$$

From (6.22)

$$||\hat{\theta}''(t) - \theta_s|| = \rho \tag{6.67}$$

Thus, (6.64) - (6.67) and $\frac{\lambda_1}{\lambda_0} \geqq 1$ lead to the announced property.

With the above properties in hand, Lemma 6.1 can now be established.

(1) Lemma 6.2-(3) leads directly to Lemma 6.1-(1). Then with (6.25), boundedness of V(t) follows.

(2) From Lemma 6.2-(1) and (2)

$$\bar{e}(t)^2 \leqq (\mu + \sigma(t-1))(V(t-1) - V(t)) + \frac{\mu + \sigma(t-1)}{\mu}\bar{w}(t)^2 \tag{6.68}$$

Noting that

$$\sigma(t-1) \leqq \lambda_1||\bar{\phi}(t-1)||^2$$

$$||\bar{\phi}(t)|| \leqq 1$$

(6.68) may be replaced by

$$\bar{e}(t)^2 \leqq (\mu + \lambda_1)(V(t-1) - V(t)) + \left(1 + \frac{\lambda_1}{\mu}\right)\bar{w}(t)^2 \tag{6.69}$$

Hence for any k, any q

$$\sum_{t=q+1}^{q+k} \bar{e}(t)^2 \leqq (\mu + \lambda_1)(V(q) - V(q+k)) + \left(1 + \frac{\lambda_1}{\mu}\right) kn_w^2 \tag{6.70}$$

With the Schwarz inequality, Lemma 6.1-(2) follows immediately.

(3) From Lemma 6.2-(4)

$$||\hat{\theta}(t) - \hat{\theta}(t-1)||^2 \leqq 4 \frac{||P(t-2)\bar{\phi}(t-1)||^2}{(\mu + \sigma(t-1))^2} \bar{e}(t)^2 \frac{\lambda_1}{\lambda_0}$$

$$\leqq 4 \frac{\lambda_1{}^2||\bar{\phi}(t-1)||^2}{(\mu + \sigma(t-1))^2} \bar{e}(t)^2 \frac{\lambda_1}{\lambda_0} \qquad (6.71)$$

Noting that

$$\lambda_0 \leqq \sigma(t-1)$$

(6.71) and $\frac{\lambda_1}{\lambda_0} \geqq 1$ lead to the announced property.

## 6.4 Properties of the Adaptive Controller

In this section, the properties of the adaptive controller are studied. The analysis technique of Praly and Trulsson (1986) will be used.

Define the a posteriori error as:

$$n(t) = y(t) - \phi(t-1)^T \hat{\theta}(t) \qquad (6.72)$$

(6.72) can be written as

$$n(t) = y(t) - \phi(t-1)^T \hat{\theta}(t-1) - \phi(t-1)^T(\hat{\theta}(t) - \hat{\theta}(t-1))$$

$$= e(t) - \phi(t-1)^T(\hat{\theta}(t) - \hat{\theta}(t-1)) \quad \text{using (6.20)} \qquad (6.73)$$

Thus

$$|n(t)| \leqq |e(t)| + ||\phi(t-1)||\,||\hat{\theta}(t) - \hat{\theta}(t-1)||$$

$$\leqq (1 + L_2)|e(t)| \quad \text{using Lemma 6.1-(3)} \qquad (6.74)$$

Let

$$P(d) = 1 + p_1 d + \ldots\ldots + p_{np} d^{np}; \quad np \leqq n \qquad (6.75)$$

The control law (6.17) can be written as (for simplicity, $G_1(d) = 1$)

$$u(t) = \Psi^T\phi(t-1) + \hat{f}(t)[\hat{B}(t,d)u(t) - \hat{A}(t,d)y(t)] + \hat{f}(t)\hat{A}(t,d)y^*(t)$$

$$= \Psi^T\phi(t-1) + \hat{f}(t)[\phi(t-1)^T \hat{\theta}(t) - y(t)] + \hat{f}(t)v^*(t) \qquad (6.76)$$

where

$$\hat{f}(t) = \hat{K}(t)^{-1} \tag{6.77}$$

$$\Psi = [\ 0 \ \ldots \ 0 \mid -p_1 \ldots\ldots -p_{np}\ ]^T$$

$$v^*(t) = \hat{A}(t,d)y^*(t) \tag{6.78}$$

Notice that $\{v^*(t)\}$ is uniformly bounded due to Lemma 6.1-(1) and Assumption 6B-(2).

Combining (6.72) and (6.76) yields the following closed-loop state-space representation:

$$X(t+1) = F(t)X(t) + B(t)n(t) + D(t)v^*(t) \tag{6.79}$$

wHere

$$X(t+1) = \phi(t)$$

$$F(t) = \begin{bmatrix} -\hat{a}_1(t) & -\hat{a}_2(t) & \ldots & -\hat{a}_n(t) & \hat{b}_1(t) & \hat{b}_2(t) & \ldots & \hat{b}_n(t) \\ 1 & 0 & . & 0 & 0 & 0 & . & 0 \\ 0 & 1 & . & 0 & 0 & . & . & 0 \\ \vdots & . & 1 & 0 & . & . & & \vdots \\ 0 & 0 & \ldots & 0 & -p_1 & -p_2 & \ldots & -p_{np} \\ . & & & . & 1 & 0 & . & 0 \\ . & & & . & 0 & 1 & . & 0 \\ & & & & & & . & \\ 0 & . & . & & 0 & . & 1 & 0 \end{bmatrix}$$

$$B(t) = [\ 1 \ 0 \ \ldots \ 0 \ -\hat{f}(t) \ 0 \ \ldots \ 0\ ]^T$$

$$D(t) = [\ 0 \ \ldots \ 0 \ \hat{f}(t) \ 0 \ \ldots \ 0\ ]^T$$

Some facts about $F(t)$ are given as follows.

(1) $||F(t)|| < M_4$ for all $t$ (6.80)

(2) $||F(t)^i|| \leq c\nu^i$ $(i \geq 0)$; $0 \leq \nu < 1$ for all $t$ (6.81)

(3) $||F(t) - F(t-1)|| \leq L_2 \dfrac{|e(t)|}{s(t)}$ (6.82)

Property (1) follows from Lemma 6.1-(1). Notice that $F(t)$ is an exponentially stable matrix due to (6.24) and Assumption 6B-(1). Hence, Property (2) follows from the results of Desoer (1970) or Fuchs (1982). Property (3) follows from Lemma 6.1-(3). It is assumed that by choice of the polynomial $P(d)$ and the estimated model stability region that $\nu$ can be chosen such that

$$0 \leq \nu < \sigma \tag{6.83}$$

Now, (6.79) can be rewritten as

$$X(t+1) = F(t)^{t+1}X(0) - F(t)^{t+1}X(0) + F(t)^{t}X(1) - F(t)^{t}X(1) + \dots +$$
$$F(t)^{t-k+1}X(k) - F(t)^{t-k+1}X(k) + \dots + F(t)X(t) + B(t)n(t) +$$
$$D(t)v^*(t)$$
$$= F(t)^{t+1}X(0) - F(t)^{t}[(F(t) - F(0))X(0) - B(0)n(0) -$$
$$D(0)v^*(0)] + \dots + F(t)^{t-k}[(F(t) - F(k))X(k) - B(k)n(k)$$
$$- D(k)v^*(k)] + \dots + B(t)n(t) + D(t)v^*(t)$$

using (6.79)

$$= F(t)^{t+1}X(0) - \sum_{i=0}^{t} F(t)^{t-i}[(F(t) - F(i))X(i) - B(i)n(i)$$
$$- D(i)v^*(i)] \tag{6.84}$$

Taking norms and assuming that $||X(0)||$ is zero gives

$$||X(t+1)|| \leq ||\sum_{i=0}^{t} F(t)^{t-i}[(F(t) - F(i))X(i) - B(i)n(i) -$$
$$- D(i)v^*(i)]||$$
$$\leq c \sum_{i=0}^{t} \nu^{t-i}[||F(t) - F(i)||\,||X(i)|| + ||B(i)||\,|n(i)|$$
$$+ ||D(i)||\,|v^*(i)|] \quad \text{using (6.81)}$$
$$\leq c \sum_{i=0}^{t} \nu^{t-i}[||F(t) - F(i)||\,||X(i)|| + M_5|n(i)| + M_6] \tag{6.85}$$

after using the boundedness of $||B(t)||$, $||D(t)||$ and $|v^*(t)|$.

Next, it will be shown that, using (6.82), (6.83), (6.85), and the definition of $s(t)$, the following property of the adaptive controller can be derived.

Lemma 6.3

There exists constants $C_1$, $C_2$ and $C_3$ such that for all $t_0$ and t

$$\sigma^{-2t}s(t) \leq C(t_0)\sigma^{-2t_0}s(t_0)^2 + C_2\,\sigma^{-2t} + C_3 \sum_{i=t_0+1}^{t-1} \sigma^{-2i}|e(i)|^2$$

where

$$C(t_0) = 1 + C_1 \max_{0 \leq t \leq t_0} \left[\frac{e(t)}{s(t)}\right]^2$$

Proof

From (6.82)

$$||F(t) - F(i)|| \leq L_2 \sum_{j=i+1}^{t} \frac{|e(j)|}{s(j)} \tag{6.86}$$

Hence

$$\sum_{i=0}^{t} \nu^{t-i}||F(t) - F(i)||\,||X(i)|| \leq L_2 \sum_{i=0}^{t} \nu^{t-i} \sum_{j=i+1}^{t} \frac{|e(j)|}{s(j)}||X(i)||$$

$$\leq L_2 \sum_{j=1}^{t} \nu^{t-j+1} \frac{|e(j)|}{s(j)} \sum_{i=0}^{j-1} \nu^{j-1-i}||X(i)|| \tag{6.87}$$

Consider the second summation term on the right-hand side of (6.87). It can be written as

$$\sum_{i=0}^{j-1} \nu^{j-1-i}||X(i)|| = \sum_{i=0}^{j-1}\left[\frac{\nu}{\sigma}\right]^{j-1-i} \sigma^{j-1-i}||X(i)|| \tag{6.88}$$

Using the Schwarz inequality gives

$$\sum_{i=0}^{j-1} \nu^{j-1-i}||X(i)|| \leq \left[\sum_{i=0}^{j-1}\left[\frac{\nu}{\sigma}\right]^{2(j-1-i)}\right]^{1/2}\left[\sum_{i=0}^{j-1}\sigma^{2(j-1-i)}||X(i)||^2\right]^{1/2}$$

$$\leq \frac{\left[1 - \left[\frac{\nu}{\sigma}\right]^{2(j-1)}\right]^{1/2}}{\left[1 - \left[\frac{\nu}{\sigma}\right]^2\right]^{1/2}}\left[\sum_{i=0}^{j-1}\sigma^{2(j-1-i)}||\bar{X}(i)||^2\right]^{1/2} \tag{6.89}$$

where

$$||\bar{X}(i)||^2 = \max(s^2, ||X(i)||^2)$$

From (6.15)

$$s(j)^2 = \sigma^{2j}s(j_0)^2 + \sum_{i=1}^{j} \sigma^{2(j-i)}\max(s^2, ||X(i)||^2)$$

and since $\nu < \sigma$, (6.89) becomes

$$\sum_{i=0}^{j-1} \nu^{j-1-i}||X(i)|| \leq \frac{\sigma}{(\sigma^2 - \nu^2)^{1/2}} s(j-1)$$

$$\leq \frac{s(j)}{(\sigma^2 - \nu^2)^{1/2}} \tag{6.90}$$

after using the inequality (see (6.15))

$$s(j)^2 \geq \sigma^2 s(j-1)^2$$

Thus (6.87) becomes

$$\sum_{i=0}^{t} \nu^{t-i}||F(t) - F(i)||\,||X(i)|| \leq \frac{L_2}{(\sigma^2 - \nu^2)^{1/2}} \sum_{j=1}^{t} \nu^{t-j+1}|e(j)| \tag{6.91}$$

Substituting (6.91) into (6.85) and using (6.74) gives

$$||X(t+1)|| \leq c \sum_{j=1}^{t} \nu^{t-j+1}\frac{L_2}{(\sigma^2 - \nu^2)^{1/2}} |e(j)| +$$

$$c \sum_{i=0}^{t} \nu^{t-i}[M_5(1 + L_2)|e(i)| + M_6]$$

$$\leq c \sum_{i=0}^{t} \nu^{t-i}\left[\left(\frac{\nu L_2}{(\sigma^2 - \nu^2)^{1/2}} + M_5(1 + L_2)\right)|e(i)| + M_6\right] \tag{6.92}$$

Notice that the operator: $e(t) \longrightarrow X(t)$ is $(1,\nu)$ exponentially stable.

Now, (6.92) can be rearranged to yield

$$||X(t+1)|| \leq M_7 \sum_{i=0}^{t} \nu^{t-i}|e(i)| + M_8 \sum_{i=0}^{t} \nu^{t-i} \tag{6.93}$$

where

$$M_7 = c\left[\frac{\nu L_2}{(\sigma^2 - \nu^2)^{1/2}} + M_5(1 + L_2)\right]$$

$$M_8 = cM_6$$

Using the Schwarz inequality gives

$$||X(t+1)||^2 \leq 2M_7{}^2\left[\sum_{i=0}^{t} \nu^{t-i}|e(i)|\right]^2 + 2M_8{}^2\left[\sum_{i=0}^{t} \nu^{t-i}\right]^2 \tag{6.94}$$

Consider the first summation term of (6.94). It can be written as

$$\sum_{i=0}^{t} \nu^{t-i}|e(i)| = \sum_{i=0}^{t} \left[\frac{\nu}{\sigma}\right]^{(t-i)/2} (\nu\sigma)^{(t-i)/2}|e(i)| \tag{6.95}$$

Applying the Schwarz inequality gives

$$\left[\sum_{i=0}^{t} \nu^{t-i}|e(i)|\right]^2 \leq \left[\sum_{i=0}^{t} \left[\frac{\nu}{\sigma}\right]^{t-i}\right]\left[\sum_{i=0}^{t} (\nu\sigma)^{t-i}|e(i)|^2\right]$$

$$\leq \frac{\left(1 - \left|\frac{\nu}{\sigma}\right|^{t+1}\right)}{1 - \frac{\nu}{\sigma}} \sum_{i=0}^{t} (\nu\sigma)^{t-i}|e(i)|^2$$

$$\leq \frac{\sigma}{\sigma - \nu} \sum_{i=0}^{t} (\nu\sigma)^{t-i}|e(i)|^2 \tag{6.96}$$

The second summation term of (6.94) is given by

$$\left[\sum_{i=0}^{t} \nu^{t-i}\right]^2 = \left[\frac{1 - \nu^{t+1}}{1 - \nu}\right]^2$$

$$\leq \left[\frac{1}{1 - \nu}\right]^2 \tag{6.97}$$

Thus (6.94) becomes

$$||X(t+1)||^2 \leq 2M_7{}^2\frac{\sigma}{\sigma - \nu} \sum_{i=0}^{t} (\sigma\nu)^{t-i}|e(i)|^2 + 2M_8{}^2\left[\frac{1}{1 - \nu}\right]^2 \tag{6.98}$$

The following relation follows from (6.98)

$$\max(s^2, ||X(t)||^2) \leq M_9 \sum_{i=0}^{t-1} (\sigma\nu)^{t-i-1}|e(i)|^2 + M_{10} \tag{6.99}$$

where

$$M_9 = \frac{2c^2\sigma}{\sigma - \nu}\left[\frac{\nu L_2}{(\sigma^2 - \nu^2)^{1/2}} + M_5(1 + L_2)\right]$$

$$M_{10} = 2\left[\frac{cM_6}{1 - \nu}\right]^2 + s^2$$

Now, the definition of $s(t)^2$ gives

$$\sigma^{-2T}s(T)^2 - \sigma^{-2t_0}s(t_0)^2 = \sum_{t=t_0+1}^{T} \sigma^{-2t}\max(s^2, ||X(t)||^2) \qquad (6.100)$$

Hence, using (6.99) yields

$$\sigma^{-2T}s(T)^2 - \sigma^{-2t_0}s(t_0)^2 \leq \sum_{t=t_0+1}^{T} \sigma^{-2t} M_9 \sum_{i=0}^{t-1} (\sigma\nu)^{t-i-1}|e(i)|^2$$

$$+ \sum_{t=t_0+1}^{T} M_{10}\sigma^{-2t} \qquad (6.101)$$

Consider the first right-hand side term of (6.101). It can be rearranged to give

$$M_9 \sum_{t=t_0+1}^{T} \sigma^{-2t} \sum_{i=0}^{t-1} (\sigma\nu)^{t-i-1}|e(i)|^2 = M_9[\sum_{i=t_0+1}^{T-1} (\sigma\nu)^{-1-i}|e(i)|^2 \sum_{t=t_0+1}^{T} \left[\frac{\nu}{\sigma}\right]^t$$

$$+ \sum_{i=0}^{t_0} (\sigma\nu)^{-1-i}|e(i)|^2 \sum_{t=t_0+1}^{T} \left[\frac{\nu}{\sigma}\right]^t]$$

$$\leq M_9[\sum_{i=t_0+1}^{T-1} |e(i)|^2\sigma^{-2i} \left[\frac{1}{\sigma^2 - \sigma\nu}\right]$$

$$+ \left[\frac{\nu}{\sigma}\right]^{t_0} \sum_{i=0}^{t_0} |e(i)|^2(\sigma\nu)^{-i}\left[\frac{1}{\sigma^2 - \sigma\nu}\right]]$$

$$(6.102)$$

Similarly, the second right-hand side term of (6.101) is given by

$$M_{10} \sum_{t=t_0+1}^{T} \sigma^{-2t} = M_{10}\, \sigma^{-2T} \frac{(1 - \sigma^{2(T-t_0+2)})}{1 - \sigma^2}$$

$$\leq M_{10} \frac{\sigma^{-2T}}{1 - \sigma^2} \qquad (6.103)$$

Thus with (6.102) and (6.103), (6.101) becomes

$$\sigma^{-2T}s(T)^2 - \sigma^{-2t_0}s(t_0)^2 \leq \frac{M_9}{\sigma^2 - \sigma\nu} \left[\sum_{i=t_0+1}^{T-1} \sigma^{-2i}|e(i)|^2 + \right.$$

$$\left.\left[\frac{\nu}{\sigma}\right]^{t_0} \sum_{i=0}^{t_0} (\sigma\nu)^{-i}|e(i)|^2\right] + \frac{M_{10}\,\sigma^{-2T}}{1 - \sigma^2}$$

$$(6.104)$$

The second summation term in (6.104) can be rearranged to give

$$\left[\frac{\nu}{\sigma}\right]^{t_0} \sum_{i=0}^{t_0} (\sigma\nu)^{-i} |e(i)|^2 = \left[\frac{\nu}{\sigma}\right]^{t_0} \sum_{i=0}^{t_0} (\sigma\nu)^{-i} \left[\frac{e(i)}{s(i)}\right]^2 s(i)^2$$

$$\leq \left[\frac{\nu}{\sigma}\right]^{t_0} e(t_0) \sum_{i=0}^{t_0} (\sigma\nu)^{-i} s(i)^2 \qquad (6.105)$$

where

$$e(t_0) = \max_{0 \leq t \leq t_0} \left[\frac{e(t)}{s(t)}\right]^2 \qquad (6.106)$$

Now, since $\sigma^{-2i} s(i)^2$ is increasing

$$\left[\frac{\nu}{\sigma}\right]^{t_0} \sum_{i=0}^{t_0} (\sigma\nu)^{-i} s(i)^2 = \sum_{i=0}^{t_0} \sigma^{-i-t_0} \nu^{t_0-i} s(i)^2$$

$$= \sum_{i=0}^{t_0} \left[\frac{\nu}{\sigma}\right]^{t_0-i} \sigma^{-2i} s(i)^2$$

$$\leq \sigma^{-2t_0} s(t_0)^2 \frac{1}{1 - \frac{\nu}{\sigma}}$$

$$\leq \frac{\sigma}{\sigma - \nu} \sigma^{-2t_0} s(t_0)^2 \qquad (6.107)$$

Thus (6.104) can be rewritten as

$$\sigma^{-2T} s(T) \leq \left[1 + \frac{M_9}{(\sigma - \nu)^2} e(t_0)\right] \sigma^{-2t_0} s(t_0)^2 + \frac{M_{10}}{1 - \sigma^2} \sigma^{-2T}$$

$$+ \frac{M_9}{\sigma^2 - \sigma\nu} \sum_{i=t_0+1}^{T-1} \sigma^{-2i} |e(i)|^2 \qquad (6.108)$$

or

$$\sigma^{-2t} s(t) \leq \left[1 + \frac{M_9}{(\sigma - \nu)^2} e(t_0)\right] \sigma^{-2t_0} s(t_0)^2 + \frac{M_{10}}{1 - \sigma^2} \sigma^{-2t}$$

$$+ \frac{M_9}{\sigma^2 - \sigma\nu} \sum_{i=t_0+1}^{t-1} \sigma^{-2i} |e(i)|^2 \quad \text{for } t \text{ in } [t_0, T] \qquad (6.109)$$

Note that taking $t_0 = 0$, the operator: $e(t) \longrightarrow s(t)$ is $(2,\sigma)$ exponentially stable.

## 6.5 Conditions for Stability

In the following, it will be shown that boundedness of the input-output signals follows from (6.109) and Lemma 6.1-(2).

From Lemma 6.1-(2)

$$\left[\frac{e(t)}{s(t)}\right]^2 \leq M_3 + L_1 n_w^2 \tag{6.110}$$

Now using (6.109)

$$\sigma^{-2t} s(t) \leq \left[1 + \frac{M_9}{(\sigma - \nu)^2} e(t_0)\right] \sigma^{-2t_0} s(t_0)^2 + \frac{M_{10}}{1 - \sigma^2} \sigma^{-2t}$$

$$+ \frac{M_9}{\sigma^2 - \sigma\nu} \sum_{i=t_0+1}^{t-1} \sigma^{-2i} \left[\frac{e(i)}{s(i)}\right]^2 s(i)^2$$

$$\leq K_1(n_w)\sigma^{-2t_0} s(t_0)^2 + K_2 \sigma^{-2t} + K_3(n_w) \sum_{i=t_0+1}^{t-1} \sigma^{-2i} s(i)^2$$

$$\leq K_2 \sigma^{-2t} + K_4(n_w) \sum_{i=t_0}^{t-1} \sigma^{-2i} s(i)^2 \tag{6.111}$$

where

$$K_1(n_w) = \left[1 + \frac{M_9}{(\sigma - \nu)^2} (M_3 + L_1 n_w^2)\right]$$

$$K_2 = \frac{M_{10}}{1 - \sigma^2}$$

$$K_3(n_w) = \frac{M_9}{\sigma^2 - \sigma\nu} (M_3 + L_1 n_w^2)$$

$$K_4(n_w) = \max\left(K_1(n_w), K_3(n_w)\right)$$

Thus applying the Bellman-Gronwall Lemma (p. 254, Desoer and Vidyasagar 1975) gives

$$\sigma^{-2t} s(t)^2 \leq K_2 \sigma^{-2t} + \sum_{i=t_0}^{t-1} \prod_{j=i+1}^{t-1} (1 + K_{4j}(n_w))\, K_{4i}(n_w) K_2 \sigma^{-2i}$$

$$\leq K_2 \sigma^{-2t} + K_4(n_w) K_2 \left[\sum_{i=t_0}^{t-1} \prod_{j=i+1}^{t-1} (1 + K_{4j}(n_w)) \sigma^{-2i}\right] \tag{6.112}$$

Now the product term yields

$$\prod_{j=i+1}^{t-1} (1 + K_{4j}(n_W)) \le \prod_{j=i+1}^{t-1} e^{K_{4j}(n_W)}, \qquad (\text{since } e^x \ge 1 + x)$$

$$\le e^{\left[\sum_{j=i+1}^{t-1} K_{4j}(n_W)\right]}$$

$$\le e^{(t-i-1)K_4(n_W)} \tag{6.113}$$

Hence, (6.112) becomes

$$s(t)^2 \le K_2 + K_4(n_W)K_2e^{-K_4(n_W)} \sum_{i=t_0}^{t-1} \sigma^{2(t-i)} e^{(t-i)K_4(n_W)}$$

$$\le K_2 + K_4(n_W)K_2e^{-K_4(n_W)} \sum_{i=t_0}^{t-1} \left[\sigma e^{K_4(n_W)/2}\right]^{2(t-i)} \tag{6.114}$$

Now, if there exists a $K_4(n_W)$ such that

$$x = \sigma e^{K_4(n_W)/2} < 1 \tag{6.115}$$

then

$$s(t)^2 \le K_2 + K_4(n_W)K_2e^{-K_4(n_W)} \sum_{i=t_0}^{t-1} x^{2(t-i)}$$

$$\le K_2 + K_4(n_W)K_2e^{-K_4(n_W)} x^2\left[\frac{1 - x^{2(t-t_0)}}{1 - x^2}\right]$$

$$\le S \quad \text{for all } t \tag{6.116}$$

That is, the input-output signals are bounded.

The main result is summarized in the following theorem.

Theorem 6.1

Subject to Assumptions 6A and 6B, and if the following condition

$$\sigma e^{K_4(n_W)/2} < 1$$

is met, the algorithm (6.19) - (6.23) will be robustly stable.

Remark

The proof of the boundedeness of the regressor vector relies on the existence of a $K_4(n_W)$ such that (6.115) is satisfied. This implies that the unmodelled dynamics considered must be sufficiently small relative to the normalizing signals.

## 6.6 Conclusion

In this chapter, a 'robustified' pole placement adaptive controller has been derived and its stability in the presence of unmodelled dynamics and bounded disturbances has been studied. The controller is designed for the reduced-order plant, which is assumed stable but not necessarily minimum phase. The overall plant is assumed stable but may be nonminimum phase. Using a normalization signal and projection in the adaptive law, robust stability of the adaptive controller has been demonstrated. The additional a priori information required is that certain bounds on the unknown plant parameters are known. The unmodelled dynamics considered are required to be sufficiently small relative to the normalizing signals.

# CHAPTER 7

# CONCLUSIONS

In this thesis, a pole-zero placement algorithm for explicit adaptive control of single-variable and multivariable discrete-time plants has been presented. The stability of the resulting adaptively controlled systems subject to external bounded disturbances and unmodel-led dynamics has also been investigated.

## 7.1 Adaptive Controller Based on Pole-zero Placement

The main conclusions are given in the following.

The adaptive control algorithm has several attractive features:

(1) the pole polynomial P(d) determines the transient response while the zero polynomial $G_1(d)$ specifies additional closed-loop zeros which may modify the control action;

(2) it possesses inherent integral action which ensures zero steady state error under constant inputs and biases;

(3) although explicit, the controller design step is trivial and the estimated model stabilizability problem does not exist;

(4) it is applicable to nonminimum phase plants, which in the discrete-time context, are common;

(5) it can handle unknown and varying but bounded time delays.

However, the algorithm is applicable only to open-loop stable plants, although this is not a serious limitation because most practical processes are inherently stable.

Convergence properties have been established for the resulting adaptive control algorithms applied to linear time-invariant systems having purely deterministic or stochastic disturbances.

In the deterministic single-variable case, it has been shown that the adaptive control algorithm ensures that the disturbances are removed from the system output and that asymptotic perfect tracking is achieved. Examples illustrating the performance and robustness of the algorithm have also been given.

In the deterministic multivariable case, it has been shown that the adaptive control algorithm ensures that asymptotic perfect tracking is achieved. Examples illustrating the performance of the algorithm have been given. It has been demonstrated that by choosing appropriate pole-zero polynomials, it may be possible to reduce interaction between channels, although a systematic procedure for doing this has yet to be developed.

In the stochastic case, it has been shown that the adaptive control algorithm leads to the required stability properties of the closed-loop system. A noteworthy feature is that no asymptotic behaviour of the parameter estimates have been imposed. However, the results obtained preclude purely deterministic disturbances.

The robustness of the nonadaptive control algorithm to nonideal conditions has also been studied. It has been shown that given a nominal plant and some plant uncertainty quantified in terms of gain measure or frequency response, the design polynomials may be appropriately chosen to ensure stability. Also, robust disturbance attenuation may be enhanced by using appropriate design polynomials.

In the adaptive case, in order to achieve robustness of the adaptively controlled systems to modelling errors, a modified adaptive law has been used. Specifically, a normalizing signal to bound the modelling error and a projection to keep the parameter estimates bounded have been introduced in the adaptive law. The modelling error is assumed sufficiently small relative to the normalizing signals. Bounds on the unknown plant parameters are required to design the adaptive law.

## 7.2 Further Areas of Study

Most of the results presented in this thesis have been established for single-variable cases. The extension to the multivariable cases can be made proceeding along the lines of the work for the single-variable cases. Particular attention should be paid to the decoupling problem.

The extension of the results obtained should be made to include both stochastic and purely deterministic disturbances. Perhaps the case which involves constant bias should be studied since the controller has an inherent integrator.

The adaptive controller has been designed to be stable under the assumption of unknown plant parameters which are time-invariant. The stability of the adaptive controller should be studied in the case where the unknown plant parameters are time-varying, possibly in an unknown manner.

Recently, there has been some uses of expert system techniques in control systems and it appears that an area needing further research would be the incorporation of expert system methodolgy in the algorithm to improve its versatility and reliability for practical applications.

# REFERENCES


Anderson, B.D.O. and Johnstone, R.M. (1985)
Global adaptive pole-positioning,
IEEE Trans. AC-30, No. 1, 11-22.

Anderson, B.D.O. and Moore, B.J. (1979)
Optimal filtering,
Prentice-Hall, Englewoods Cliffs, NJ.

Astrom, K.J. (1980)
Robustness of a design method based on assignment of poles and zeros,
IEEE Trans. AC-25, No. 3, 588-591.

Astrom, K.J. (1983)
Theory and applications of adaptive control- A survey,
Automatica, Vol. 19, No.5, 471-486.

Astrom, K.J., Hagander, P. and Sternby, J. (1984)
Zeros of sampled systems,
Automatica, Vol. 20, 31-39.

Astrom, K.J. and Wittenmark, B. (1973)
On self-tuning regulators,
Automatica, Vol. 9, No. 2, 185-199.

Astrom, K.J. and Wittenmark, B. (1980)
Self-tuning controllers based on pole-zero placement,
Proc. IEE, Pt. D, Vol. 127, No. 3, 120-130.

Clark, D.W. (1984)
Self-tuning control of nonminimum-phase systems,
Automatica, Vol. 20, No. 5, 501-517.

Clark, D. and Gawthrop, P.J. (1975)
Self-tuning controller,
Proc. IEE, Pt. D, Vol. 122, No.9, 929-934.

Cluett, W.R., Shah, S.L. and Fisher, D.G. (1987)
Stable robust adaptive controller,
Int. J. Control, Vol. 45, No. 4, 1265-1273.

De Larminat, Ph. (1984)
On the stabilizability condition in indirect adaptive control,
Automatica, Vol. 20, No. 6, 793-795.

Desoer, C.A. (1970)
Slowly varying discrete system $x(i+1) = A(i)x(i)$,
Electron. Lett., Vol. 6, 339-340.

Desoer, C.A. and Vidyasagar, M. (1975)
Feedback systems: Input and output properties,
Academic Press.

Egardt, B. (1979)
Stability of adaptive controllers,
Lecture notes in Control and Information Sciences, Vol. 20,
Springer Verlag, New York.

Elliot, G. (1982)
Direct adaptive pole-placement with application to nonminimum phase systems,
IEEE Trans. AC-27, No. 3, 720-722.

Elliot, G., Cristi, R. and Das, M. (1985)
Global stability of of adaptive pole-placement algorithms,
IEEE Trans. AC-30, No. 4, 348-356.

Fuchs, J.J. (1980)
Explicit self-tuning methods,
Proc. IEE, Pt. D, Vol. 127, No. 6, 259-264.

Fuchs, J.J. (1982)
On the good use of the spectral radius of a matrix,
IEEE Trans. AC-27, No. 5, 1134-1135.

Gawthrop, P.J. and Lim, K.W. (1982)
Robustness of self-tuning controllers,
Proc. IEE, Vol. 129, Pt. D, No. 1, 21-29.

Goodwin, G.C. and Chan, S.W. (1983)
Model reference adaptive control of systems having purely deterministic disturbances,
IEEE Trans. AC-28, No. 8, 855-858.

Goodwin, G.C., Hill, D.J. and Palaniswami, M. (1984)
A perspective on convergence of adaptive control algorithms,
Automatica, Vol. 20, No. 5, 519-531.

Goodwin, G.C., Leal, R.L., Maynes, D.Q. and Middleton, R.H. (1986)
Rapproachment between continuous and discrete model reference adaptive control,
Automatica, Vol. 22, 199-207.

Goodwin, C.G., Ramadge, P.J. and Caines, P.E. (1980)
Discrete-time multivariable adaptive control,
IEEE Trans. AC-25, No. 3, 449-456.

Goodwin, G.C. and Sin, K.S. (1984)
Adaptive filtering, prediction and control,
Prentice-Hall, New Jersey.

Goodwin, G.C. and Teoh, E.K. (1985)
Persistency of excitation in the presence of possibly unbounded signals,
IEEE Trans. AC-30, No. 6, 595-597.

Hersh, M.A. and Zarrop, M.B. (1986)
Stochastic adaptive control of nonminimum-phase systems,
Opt. Control Appl & Meth., Vol. 7, No. 2, 153-161.

Hsu, L. and Costa, R.R. (1987)
Adaptive control with discontinuous $\sigma$-factor and saturation for improved robustness,
Int. J. Control, Vol. 45, No. 3, 843-859.

Ioannou, P.A. (1986)
Robust Adaptive controller with zero residual tracking errors,
IEEE Trans. AC-31, No. 8, 773-776.

Ioannou, P.A. and Kokotovic, P. (1984)
Instability analysis and improvement of robustness of adaptive control,
Automatica, Vol. 20, No. 5, 583-594.

Ioannou, P.A. and Tsakalis, K.S. (1986)
A robust direct adaptive controller,
IEEE Trans. AC-31, No. 11, 1033-1043.

Isermann, R. (1981)
Digital control systems,
Springer-Verlag, New York.

Janecki, D. (1987)
Direct adaptive pole placement for plants having purely deterministic disturbances,
IEEE Trans. AC-32, No. 2, 187-189.

Kalman, R.E. (1958)
Design of a self-optimizing control system,
Trans. ASME, Vol. 80, 468-478.

Karam, K.Z., Warwick, K. and Farsi, M. (1986)
An implicit pole-assignment self-tuning regulator,
Opt. Control Appl & Meth., Vol. 7, No. 2, 127-134.

Kosut, R.L. and Friedlander, B. (1985)
Robust adaptive control: Conditions for global stability,
IEEE Trans. AC-30, No. 7, 610-624.

Kreisselmeier, G. (1985)
An approach to stable indirect adaptive control,
Automatica, Vol. 21, No. 4, 425-431.

Kreisselmeier, G. (1986)
A robust indirect adaptive control approach,
Int. J. Control, Vol. 43, No. 1, 161-175.

Kreisselmeier, G. and Anderson, B.D.O. (1986)
Robust model reference adaptive control,
IEEE Trans. AC-31, No. 2, 127-133.

Kreisselmeier, G. and Narendra, K.S. (1982)
Stable model reference adaptive control in the presence of bounded disturbances,
IEEE Trans. AC-27, No. 6, 1169-1176.

Kumar, P.R. (1985)
Survey of results in stochastic adaptive control,
SIAM J. Control & Opt., Vol. 23, No. 3, 329-380.

Landau, I.D. (1979)
Adaptive control. The model reference approach,
Marcel Dekker, New York.

Lim, K.W. (1982)
Robustness of self-tuning controllers,
Ph.D. Thesis, University of Oxford.

Lin Chih-Min and Chen Bor-Sen (1986)
Adaptive controller with desired pole/zero assignment,
Proc. IEE, Vol. 133, Pt. D, No. 6, 301-306.

Lozano, R. and Goodwin, G.C. (1985)
A globally convergent adaptive pole-placement algorithm without a persistency of excitation requirement,
IEEE Trans. AC-30, No. 8, 795-798.

Matko, D. and Schumann, R. (1984)
Self-tuning deadbeat controllers,
Int. J. Control, Vol. 40, No. 2, 393-402.

McDermott, P.E. and Mellichamp, D.A. (1986)
A decoupling pole placement self-tuning controller for a class of multivariable processes,
Optimal Control Appl & Meth., Vol. 7, 55-79.

M'saad, M., Ortega, R. and Landau, I.D. (1985)
Adaptive controllers for discrete-time systems with arbitrary zeros:
An overview,
Automatica, Vol. 21, No. 4, 413-423.

Narendra, K.S. and Annaswamy, A.M. (1986)
Robust adaptive control in the presence of bounded disturbances,
IEEE Trans. AC-31, No. 4, 306-315.

Narendra, K.S. and Annaswamy, A.M. (1987)
A new adaptive law for robust adaptation without persistent excitation,
IEEE Trans. AC-32, No. 2, 134-145.

Ogata, K. (1970)
Modern control engineering,
Prentice-Hall, Englewoods Cliffs, NJ.

Ortega, R., Praly, L. and Landau, I.D. (1985)
Robustness of discrete-time direct adaptive controllers,
IEEE Trans. AC-30, No. 12, 1179-1187.

Peterson, B.B. and Narendra, K.S. (1982)
Bounded error adaptive control,
IEEE Trans. AC-27, No. 6, 1161-1168.

Prager, D.L. and Wellstead, P.E. (1981)
Multivariable pole-assignment self-tuning regulator,
Proc. IEE, Vol. 128, Pt. D, No. 1, 9-18.

Praly, L. (1983)
Robustness of indirect adaptive control based on pole placement design,
IFAC Workshop on adaptive control, San Francisco, CA, June,
CAI Rep Ecole des Mines A134, October.

Praly, L. (1984)
Towards a globally stable direct adaptive control scheme for not necessarily minimum-phase systems,
IEEE Trans. AC-29, 946-949.

Praly. L. and Trulsson, E. (1986)
Decentralized indirect adaptive control,
APII, Vol. 20, 295-315.

Rohrs, C.E., Valavani, L., Athans, M. and Stein, G. (1985)
Robustness of continuous-time adaptive control algorithms in the presence of unmodelled dynamics,
IEEE Trans. AC-30, No. 9, 881-889.

Samson, C. (1983)
Stability analysis of adaptively controlled systems subject to bounded disturbances,
Automatica, Vol. 19, No. 1, 81-86.

Schumann, R. (1979)
Proc. IFAC Symposium on computer aided design of control systems,
Pergamon Press, Oxford.

Sirisena, H.R. and Teng, F.C. (1986)
Multivariable pole-zero placement self-tuning controller,
Int. J. Systems Sci., Vol. 17, No. 2, 345-352.

Wellstead, P.E., Prager, D. and Zanker, P. (1979)
Pole-assignment self-tuning regulator,
Proc. IEE, Pt. D, Vol. 126, No.8, 781-787.

Wolovich, W.A. and Falb, P.L. (1976)
Invariants and canonical forms under dynamic compensation,
SIAM J. Control & Opt., Vol. 14, 996-1008.

Zames, G. (1963)
Functional analysis applied to nonlinear feedback systems,
IEEE Trans. CT-10, No. 3, 392-404.

Zames, G. (1966)
The input-output stability of time-varying nonlinear feedback systems ( Part 1 and 2 ),
IEEE Trans. AC-11, No. 2, 228-238, 465-476.

Zames, G. (1981)
Feedback and optimal sensitivity: Model reference transformations, multiplicative seminorms and approximate inverses,
IEEE Trans. AC-26, No. 2, 301-320.

# APPENDIX

## THE DECOUPLING PROBLEM

In this section, the dynamic decoupling problem is briefly addressed. No solution is offered, however.

Definition: A linear multivariable system is said to be dynamically decoupled if and only if its transfer matrix is diagonal and non-singular.

The transfer function matrix corresponding to (2.111) is given by

$$T_c(z) = A(z^{-1})^{-1}B(z^{-1})(P(z^{-1})K)^{-1}G_1(z^{-1})A(z^{-1}) \tag{A.1}$$

and let the transfer function matrix corresponding to (2.80) be

$$T(z) = A(z^{-1})^{-1}B(z^{-1}) \tag{A.2}$$

The following result is required ( Wolovich and Falb 1976).

Lemma A.1

If T(z) is full rank and strictly proper, then there exists an interactor matrix $\xi(z)$ which satisfies

$$\lim_{z \to \infty} \xi(z)T(z) = K_o, \qquad \det K_o \neq 0$$

Now, assuming that $T_c(z)$ is of full rank and strictly proper, then there exists an interactor matrix $\xi c(z)$ which satisfies

$$\lim_{z \to \infty} \xi_c(z)T_c(z) = K_c\ , \quad \det K_c \neq 0 \tag{A.3}$$

Using (A.1) gives

$$\lim_{z \to \infty} \left(\xi_c(z)A(z^{-1})^{-1}B(z^{-1})\right)\left(P(z^{-1})K\right)^{-1}G_1(z^{-1})A(z^{-1})$$

$$= K_c \tag{A.4}$$

or

$$\left[\lim_{z \to \infty} \xi_c(z)T(z)\right] K^{-1} = K_c \tag{A.5}$$

Since the delay structure of T(z) is specified in terms of $\xi(z)$, it can be concluded that

$$\xi_c(z) = \xi(z) \tag{A.6}$$

$$K_c = K_o\, K^{-1} \tag{A.7}$$

Thus, it can be seen the interactor matrix polynomial plays a role in the decoupling problem. Now it appears that the problem cannot be totally solved because $K_c$ is, in general, nondiagonal and unknown. The solution to this decoupling problem can be the subject of further research.