The Third Conference on
Control Engineering 1986
Sydney 12-14 May

Marrying Frequency Domain Intuition and Adaptive Control

B.D.O. ANDERSON
Head, Department of Systems Engineering, Australian National University
I.M.Y. MAREELS
Post-graduate Student, Department of Systems Engineering, Australian National University; Research Student, National Fund for Scientific Research, Belgium

SUMMARY Most adaptive control algorithms make little or no connection with the frequency domain concepts so crucial to an understanding of classical control. In this paper, we outline how frequency domain concepts can influence the design of adaptive algorithms. These influences arise in two areas: determining the rate of convergence of an adaptive algorithm, and determining the robustness of an algorithm, that is, its ability to function satisfactorily in the presence of departures from the idea, such as undermodelling or failure to model high frequency plant dynamics. The end conclusions from these considerations involve frequency control or external driving signals, use of regression vector filtering, and consideration of frequency bands in which the plant may be difficult to model.

1. INTRODUCTION

A classical approach to adaptive control is to identify on line a parametrised model for the plant and to use then the identified parameters to update the control parameters. This certainty equivalence implementation of adaptive control yields generally a highly nonlinear closed loop. It is therefore difficult to predict the behaviour of the adaptive system, or even to have a conceptual framework for thinking about its operation using concepts familiar from classical control.

In this paper we present some intuitively appealing results, almost design guidelines, which are interpretable in frequency domain terms, offering a clear insight into the adaptive system's response.

Our results are valid for stable closed loop systems with "slow" adaptation, meaning that the control parameters are updated slowly compared to the controller dynamics. For most industrial applications, this is not a real restriction, as the plant variations an adaptive controller is supposed to track usually are significantly slower than the process dynamics. When the adaptation is slow, we can separate the adaptive system state into two distinct categories: the slow state, comprising the parameters to be updated, and the fast state describing the plant controller loop. This separation and the possibility of treating the slow and fast states almost independently greatly simplifies the analysis. The underlying mathematical theory is known as "averaging"[1].

The paper is organised as follows. In section 2 we develop the main ideas. The discussion in this section is only concerned with a single parameter feedforward control scheme - known as the MIT-rule for adaptive control [2]. Section 3 generalizes the results, clearly indicating that we indeed did capture all the relevant effects. Section 4 concludes the paper.

2. THE MIT RULE FOR ADAPTIVE CONTROL

2.1. Problem Formulation

We refer to Figure 1. Using a feedforward adjustment k_c for the plant $k_pZ_p(s)$ the control objective is that the plant output $y_p(t)$ would track the output $y_m(t)$ of the model $k_mZ_m(s)$, prescribed by the reference intput $r(t)$. Obviously the transfer function $Z_m(s)$ must be a good approximation for the transfer function $Z_p(s)$ over the relevant frequency range (dictated by the reference input $r(t)$) in order for this control strategy to be meaningful. (Both $Z_m(s)$ and $Z_p(s)$ are assumed to be exponentially stable.) Assuming that this is the case, minimising the integral squared error

$$I = \lim_{T\to\infty} \frac{1}{T}\int_0^T e^2(t,k_c)dt \qquad (2.1)$$

where

$$e(t,k_c) = y_p(t)-y_m(t)$$

becomes a natural criterion with which to select k_c. The optimal controller setting will depend, in particular, on the reference input $r(t)$ and the plant's gain k_p.

Suppose now that the magnitude of the gain k_p is not known and that k_p may be slowly time varying. Without loss of generality assume that $Z_p(0)=Z_m(0)=1$, k_p and k_m positive; otherwise introduce sign changes including one in k_c. The MIT rule is an adaptive law to update the gain k_c on line as to minimize the integral squared error. It is proposed that

$$\dot{k}_c \propto \partial e^2(t,k_c)/\partial k_c \qquad (2.2)$$

which is equivalent to

$$\dot{k}_c \propto - e(t,k_c)(Z_p(s)(k_pr))(t) \qquad (2.3)$$

This update mechanism is however not implementable as the signal $(Z_p(s)(k_pr))(t)$ is not available to us. Under the above assumptions it is reasonable to approximate (2.3) by

$$\dot{k}_c \propto - e(t,k_c)(Z_m(s)k_mr)(t) \qquad (2.4)$$

which can be implemented as

$$\dot{k}_c = -g(y_p(t)-y_m(t))y_m(t) \qquad (2.5)$$

Here g is a positive constant, scaling the adaptation speed. This is the MIT rule for adaptive control (Figure 2).

With the algorithm being derived via heuristic arguments, the pertinent questions are: "Does it work?", "Does it come anywhere near minimising I?"

For purposes of analysis we rewrite (2.5) as to make the dependence of the right hand side on k_c explicit:

$$\dot{k}_c = -g(Z_m(s)k_mr)(t)(Z_p(s)k_pk_cr)(t) + g(Z_m(s)k_mr)^2(t) \qquad (2.6)$$

This is a linear, time-varying system with a forcing input.

2.2. Response of the Adaptive System with Constant Reference Input

When the reference input $r(t)$ is a constant $r(t)=R$, (2.5) reduces to:

$$\dot{k}_c = -gR^2k_mk_p(Z_p(s)k_c)(t) + gR^2k_m^2 \qquad (2.8)$$

To obtain (2.8) we disregard the transient response of the model. This does not affect the stability properties. Also, we treated k_p as a constant; the non-stationarity of k_p will be dealt with later.

The stability of (2.8) can be verified using the Nyquist criterion for: $1+\gamma\, G(s)$ with $G(s) = Z_p(s)/s$ and $\gamma = gR^2k_pk_m$. The following lemma is immediate:

Lemma 1: The MIT rule for input $r(t)\equiv R$ has an infinite gain margin iff:

$$-\pi/2 < \arg Z_p(j\omega) < 3\pi/2 \qquad \forall\, \omega\, \varepsilon \qquad (2.9)\ \square$$

Infinite gain margin means that for any positive design parameter g, R^2, k_m, the algorithm is stable. In particular, all strictly positive real Z_p satisfy the condition of the Lemma. On the other hand if for example Z_p has non-minimum phase zeros, or has at least two poles more than zeros, the gain margin is finite. This means that there exists a positive constant γ^*, depending on $Z_p(s)$, such that the algorithm is strictly stable for all $gR^2k_m \varepsilon (0,\gamma^*/k_p)$. Hence, in order to implement a stable algorithm, more a priori information is required. Even where $Z_m(s)\equiv Z_p(s)$, hence γ^* is known, it is necessary to have an upper bound for k_p to be able to guarantee the stability of an implementation.

Assuming that the algorithm is stable, k_c converges exponentially fast to the steady state value

$$k_c \rightarrow k_m/k_p \quad \text{as } t\rightarrow\infty \qquad (2.10)$$

This means that the adaptive algorithm converges to the value of k which minimizes the integral squared error criterion. For a constant input I becomes

$$I = (k_pk_cR-Rk_m)^2$$

which is obviously minimised by (2.10). This is not surprising, as at d.c. $Z_m(s)$ and $Z_p(s)$ match exactly. We summarise:

Theorem 1: Assuming that the transfer functions $Z_m(s)$ and $Z_p(s)$ are stable, and match at d.c., i.e. $Z_m(0) = Z_p(0) = 1$ the MIT rule converges to the global minimum of the integral squared error criterion, when the reference input $r(t)\equiv R$ and the zeros of $1+gR^2k_pk_mZ_p(s)/s$ have negative real parts. $\square$

We note two important prerequisites necessary for the MIT rule to work:

(a) at d.c. the sign of the model and the plant should match;

(b) in general, the algorithm's design parameter gR^2k_m is limited by a constant, which is inversely proportional to the plant's gain.

2.3. The Algorithm's Response for Sinusoidal Inputs

The effect of the input on the algorithm is clearly nonlinear in nature (cfr R^2-dependence of the gain in the constant input case). Furthermore, it has been observed that an algorithm designed to work for constant input can become unstable for higher frequency input signals. This is clearly illustrated in Figure 3, which displays the stability domain of the algorithm in the frequency (ω) / gain (g) parameter plane for the algorithm driven by $r(t)=\cos(\omega t)$ and applied to a first order plant and model which match exactly :

$$k_pZ_p(s) = k_mZ_m(s) = 1/(s+1). \qquad (2.11)$$

(The figure is obtained using Floquet Theory.) For similar results see [3].

The stability/instability boundary is very complex. This is due to the strong interaction between the input's excitation and the system's dynamics. Notice also that the gain margin is drastically reduced, especially around the cut off frequency of the plant. This motivates us to consider more specifically the algorithm's response for slow adaptation - the more so, as these simulation reslts are valid for the ideal plant-model matching case only. Any modelling errors are likely to reduce the gain margin even further. Figure 4 illustrates this point. Figure 4 is obtained in the same way as Figure 3, but for the algorithm applied to the following plant and model :

$$k_pZ_p(s)=1/(s+1) \quad k_mZ_m(s)=e^{-st}/(s+1) \qquad (2.12)$$

with input $r(t)=\cos(\omega t)$. The extra phase errors introduced by the mismodelling of the delay of the plant diminishes the algorithm's performance considerably, limiting both the range of allowable gains and the range of allowable frequencies.

2.4 Slow Adaptation

Having partially motivated the need for slow adaptation, we consider it now in greater detail. The algorithm is

$$\dot{k}_c = -g(Z_m(s)k_mr)(t)(Z_p(s)k_pk_cr)(t) + g(Z_m(s)k_mr)^2(t) \qquad (2.6)$$

Assuming that g is small, i.e. k_c is slowly time-varying, it seems reasonable to approximate (2.6) by:

$$\dot{k}_c^* = -g(Z_m(s)k_mr)(t)(Z_p(s)k_pr)(t)k_c^* + g(Z_m(s)k_mr)^2(t) \qquad (2.13)$$

which is obtained from (2.6) by formally treating k_c as a constant in the right hand side of the equation (2.6). This step reduces the (stability) study from a possibly infinite dimensional, linear, time varying system to a first order, linear, time varying system! Under very general conditions it is possible to demonstrate rigorously that (2.13) is a good approximation for (2.6) and that it retains its stability properties for sufficiently small g.

The stability of (2.13) is directly accessible. We have:

Lemma 2: The homogeneous part of (2.13) is exponentially stable iff:

$$\lim_{t\to\infty}\inf \frac{1}{t}\int_0^t (Z_m(s)k_mr)(\tau)(Z_p(s)k_pr)(\tau)d\tau \geq \alpha > 0 \qquad (2.14)$$

for some positive constant α. Under this condition k_c^* converges exponentially fast to

$$k_c^* \rightarrow g\int_0^t \phi(t,\tau)\ (Z_m(s)k_mr)^2(\tau)d\tau \qquad (2.15)$$

where

$$\phi(t,\tau) = \exp -g\int_\tau^t (Z_m(s)k_mr)(\alpha)(Z_p(s)k_pr)(\alpha)d\alpha$$

$\square$

The Lemma follows by direct integration of (2.13). Inputs satisfying the condition (2.14) are referred to as "persistently exciting" inputs. Moreover, averaging theory states that the previous result carries over to the original algorithm (2.6) if g is sufficiently small.

Theorem 2: Under the conditions that $Z_p(s)$ and $Z_m(s)$ are stable and that the reference input $r(t)$ is bounded, there exists a positive constant g^*, such that for all $g\varepsilon(0, g^*)$ the gain k_c updated by the MIT rule (2.6) converges exponentially fast to

$$k_c^\infty(t) = g\int_0^t \phi(t,\tau)(Z_m(s)k_mr)^2(\tau)d\tau + O(g) \qquad (2.17)$$

if condition (2.14) is satisfied. For (2.17), $\phi(t,\tau)$ is defined by (2.16), and $O(g)$ stands for a function of t bounded above by C.g for some constant C>0. $\square$

How can one satisfy the stability condition (2.14)? Consider a reference input with a limited number of spectral lines:

$$r(t) = \sum_{n=-N}^{+N} \alpha_n e^{j\omega_n} \qquad (2.18a)$$

$$\alpha_n = \alpha_{-n}^*, \quad \alpha_0\ \varepsilon \qquad (2.18b)$$

$$\omega_n = -\omega_{-n}, \quad \omega_0 = 0 \qquad (2.18c)$$

(* stand for complex conjugate) (2.14) becomes, after some simple algebra:

$$|\alpha_0|^2+2\sum_{n=1}^{N} \mathrm{Re}(Z_m(j\omega_n)Z_p(-j\omega_n))|\alpha_n|^2 \geq \bar{\alpha} \qquad (2.19)$$

for some positive constant $\bar{\alpha}$; or equivalently

$$|\alpha_0|^2 + 2 \sum_{n=1}^{N} \{|Z_m(j\omega_n)||Z_p(j\omega_n)\}|\alpha_n|^2 \le \bar{\alpha}$$

$$- \arg Z_p(j\omega)\}.|\alpha_n|^2 \ge \bar{\alpha} \qquad (2.20)$$

Hence, condition (2.20), or (2.14), will be satisfied for every input $r(t)$ of the form (2.18), if the phase errors are less than $\pi/2$ in magnitude.

<u>Corollary 1</u>: The MIT rule for almost periodic reference input $r(t)$, (2.18), and strictly stable model and plant is stable (for sufficiently small g) if:

$$\sup_{\omega\in\mathbb{R}} |\arg Z_m(j\omega) - \arg Z_p(j\omega)| < \frac{\pi}{2} \qquad (2.21)$$

□

However, condition (2.14), or (2.20), is much less restrictive. It merely requires the reference input to have dominant energy in the frequency band where the phase errors are less than $\pi/2$ – the *energy* at a certain frequency being weighted according to the amplitude response of Z_m and Z_p.

$$|\alpha_0|^2 + \sum_{n\in\Omega_1} |Z_m(j\omega_n)||Z_p(j\omega_n)|\cos(\phi_{m,n}-\phi_{p,n})|\alpha_n|^2$$

$$\ge \sum_{n\in\Omega_2} |Z_m(j\omega_n)||Z_p(j\omega_n)||\cos(\phi_{m,n}-\phi_{p,n})|$$

$$|\alpha_n|^2 + \bar{\alpha} \qquad (2.22)$$

where

$$\Omega_1 = \{n : |\phi_{m,n} - \phi_{p,n}| < \frac{\pi}{2}\} \qquad (2.23a)$$

$$\Omega_2 = \{n : |\phi_{m,n} - \phi_{p,n}| > \frac{\pi}{2}\} \qquad (2.23b)$$

$$\phi_{m,n} = \arg Z_m(j\omega_n) \quad \phi_{p,n} = \arg Z_p(j\omega_n) \qquad (2.23c)$$

Before proceeding to interpret the asymptotic expression (2.17) in terms of the integral squared error criterion, we given an instability result. *Notice that Lemma 2 gives necessary and sufficient* conditions for exponential stability, but that Theorem 2 only gives a sufficient condition for stability. The instability result which follows from Lemma 2 via averaging is the following.

<u>Theorem</u> 3: Under the conditions that $Z_p(s)$ and $Z_j(s)$ are stable, that the reference input $r(t)$ is bounded and that there exists a positive constant β such that

$$\limsup_{t\to\infty} \frac{1}{t}\int_0^t (Z_m(s)k_m r)(\tau)(Z_p(s)k_p r)d\tau) \le -\beta \qquad (2.24)$$

there exists a positive g^*, such that for all $g\varepsilon(0,g^*)$, the MIT rule is exponentially unstable. □

Interpreting (2.23) in terms of a reference input $r(t)$ of the form (2.18) we obtain that the MIT rule is exponentially unstable if:

$$|\alpha_0|^2 + \sum_{n\in\Omega_1} |Z_m(j\omega_n)||Z_p(j\omega_n)|\cos\phi_{m,n}-\phi_{p,n}|\alpha_n|^2 + \beta$$

$$\sum_{n\in\Omega_2} |Z_n(j\omega_n)||Z_p(j\omega_n)| \cos(\phi_{m,n}-\phi_{p,n})|\alpha_n|^2 \qquad (2.25)$$

for some positive $\bar{\beta}$, Ω_1 and Ω_2 are as before (2.23). Hence, (2.24), together with (2.12a) delineate a sharp instability/stability boundary. Also (2.22) and (2.25) confirm theoretically the observations from the previous section. Moreover, they explain in precise terms both the stability and instability mechanisms for a broad class of reference inputs under the restriction of small gain.

Notice in particular that Theorem 2 and Theorem 3 explain, for small gain, the simulation results displayed in Section 2.3 (Figures 3 and 4). For exact matching, a single sinusoid always satisfies the condition (2.14) (see 2.20; Figure 3), however if phase errors larger than $\pi/2$ in magnitude are possible (cfr. example with the delay, Figure 4), then there is only a limited range of frequencies in which a single sinusoidal input guarantees stability – outside this region instability occurs (Figure 4). Delay modelling errors always cause this problem.

There remains the question of how (2.17) compares with the minimization of the integral squared error. Again we limit ourselves to inputs of the form (2.18), and we make the extra assumption that:

$$|\omega_n - \omega_m| = O(1) \quad \forall n\neq m, m, n \in \{-N,\ldots,0,1,\ldots,N\}$$

After some simple, but messy, algebra, we obtain:

$$k_c^{\infty}(t) = \frac{k_m}{k_p}\frac{\alpha_{m,m}}{\alpha_{m,p}} + O(g) \qquad (2.26)$$

where

$$\alpha_{m,m} = \lim_{t\to\infty} \frac{1}{t}\int_0^t (Z_m(s)r)^2(\tau)d\tau \qquad (2.27a)$$

$$\alpha_{m,p} = \lim_{t\to\infty} \frac{1}{t}\int_0^t (Z_m(s)r)(\tau)(Z_p(s)r)(\tau)d\tau \qquad (2.27b)$$

Notice that (2.14) and (2.20) imply that $\alpha_{m,p}>0$. This expression has to be compared with

$$k_{opt} = \frac{k_m}{k_p}\frac{\alpha_{m,p}}{\alpha_{p,p}} \qquad (2.28)$$

where

$$\alpha_{p,p} = \lim_{t\to\infty} \frac{1}{t}\int_0^t (Z_p(s)r)^2(\tau)d\tau \qquad (2.29)$$

and k_{opt} minimises the integral squared error I (see (2.1). In the situation that $Z_m(s)$ matches $Z_p(s)$ exactly on all the spectral lines ω_n of the reference input r, we obviously have that:

$$k_c^{\infty}(t) = k_{opt} + O(g) \qquad (2.30)$$

In general however k_c^{∞} will be biased positively, as the Cauchy-Shwartz inequality says that

$$\alpha_{m,m}/\alpha_{m,p} \quad \alpha_{m,p}/\alpha_{p,p} \geqslant (\alpha_{m,p}>0) \qquad (2.31)$$

Actually, we can prove the slightly more general result.

<u>Theorem</u> 4: Under the conditions that the limits (2.27a), 2.27b) exist and that $\alpha_{m,p}>0$ k_c converges exponentially fast to (2.26) for sufficiently small g. $(g\varepsilon(0,g^*)$, $g^*>0$ □

[1] (Notice that the existence of the limits is a very mild assumption, which we actually implicitly made when writing down the integral squared error. Indeed, without these assumptions, I in (2.1) makes little sense.)

We summarise the results of this section, especially paying attention to the time scale separation introduced in the derivations. The gain of the plant k_p is treated as "a constant" throughout. More precisely, k_p is effectively a constant over a time interval of $O(1/g)$ which makes it slow compared to both the adaptation dynamics and the plant and model dynamics. Likewise, the adapted gain k_c is treated as a constant compared to the plant and model dynamics. This means that $Tg \ll 1$, where T is the dominant time constant for the plant. If this separation is satisfied, the algorithm performs well when there is sufficient input excitation in that region of the frequency spectrum where Z_m is a good approximation for Z_p.

2.5 Effects of Exponential Forgetting and Error Filtering

Exponential forgetting (defined below) and error filtering are two popular algorithm modifications [4]. Their respective effects can be readily analysed using the foregoing results. Both modifications can be interpreted in terms of a changed integral squared error criterion.

Exponential forgetting corresponds to the following criterion:

$$I_e = \lim_{T\to\infty} \frac{1}{T} \int_0^T \{e^2(t,k_c) + \sigma(k_c-\bar{k}_c)^2\}dt \qquad (2.32)$$

The second term is introduced to avoid large gains, hoping to stabilise the modified MIT rule even when the input is not persistently exciting. The algorithm is

$$\dot{k}_c = -g\{\sigma(k_c-\bar{k}_c) + y_m(t)(y_p(t)-y_m(t))\} \quad (2.33).$$

Obviously, in the absence of any excitation ($r(t)\equiv0, y_m(t)\equiv0$) this modification forces k_c to converge to $\bar{k}_c$ which may or may not be a very good thing to have. How this exponential forgetting effects the stablity properties in general is readily seen from:

$$\dot{k}_c = -g\{\sigma k_c + (Z_m(s)k_m r)(t)(Z_p(s)k_c k_p r)(t)\} + g\{\sigma \bar{k}_c + (Z_m(s)k_m r)^2(t)\} \quad (2.34)$$

which is the analogue of (2.6) with exponential forgetting. Averaging reduces the stability question to the study of

$$\dot{k}_c^* = -g\{\sigma + \_(Z_m(s)k_m r)(t)(Z_p(s)k_p)r(t)\}k_c^* + g\sigma\bar{k}_c + g(Z_m(s)k_m r)^2(t)$$

and

$$\dot{k}_c(t) = k_c^*(t) + O(g)$$

if we assume that $\sigma=O(1)$. Obviously, the stability condition (2.14) now modifies to

$$\liminf_{t\to\infty} \frac{1}{t}\int_0^t [Z_m(s)k_m r](\tau)[Z_p(s)k_p r](\tau)d\tau > -\sigma$$

which indeed allows for insufficient excitation (see 2.14). However, relaxing the excitation requirement comes with a price: more steady state error. Indeed, the asymptotic expression (2.27) becomes:

$$\dot{k}_c^\infty(t) = \frac{\sigma\,\bar{k}_c + k_m^2\,\alpha_{m,m}}{\sigma + k_m k_p \alpha_{m,p}} + O(g) \quad (2.36)$$

where $\alpha_{m,m}$ and $\alpha_{m,p}$ are defined in (2.26). This last expression together with (2.35) indicates the trade-off between guaranteed stability and steady state errors.

Error filtering is beneficial to reduce unwanted spectral components in the input excitation, and to reduce the influence of additive noise at the output side of the plant. Referring to Figure 5, the algorithm (2.6) is modified to:

$$\begin{aligned}\dot{k}_c = g(G(s)Z_m(s)k_m r)(t)(G(s)Z_p(s)k_p k_c r)(t) \\ + g(G(s)Z_m(s)k_m r)^2(t) \\ - g(G(s)Z_m(s)k_m r)(t)(G(s)\varepsilon)(t) \quad (2.37)\end{aligned}$$

from which the role of G is quite obvious. This algorithm corresponds to the modified integral squared error criterion :

$$I_G = \lim_{t\to\infty} \frac{1}{t}\int_0^t (G(s)e(.,k_c))^2(\tau)d\tau \quad (2.38)$$

Using the same heuristic argument as in section (2.1), the algorithm (2.37) can be derived from (2.38). All the consequent remarks [sections (2.2) - (2.4)] follow mutatis mutandis. It simply boils down to replacing Z_m and Z_p respectively by Z_mG and Z_pG. The extra-noise input $\varepsilon(t)$ causes no analysis difficulties.

3. GENERALIZATIONS

In the previous section we were concerned with a scalar parameter feedforward control problem. We now generalize the obtained results. First we retain the linearity in the state, but increase the dimension of the parameter vector. This is achieved in an equation error identification context. Next we indicate how to generalize to the adaptive feedback control problem, which is both nonlinear in the state and the input.

3.1 Equation Error Identification

We refer to Figure 6. The plant can be represented as:

$$P(s)y(t) = Q(s)u(t) + P(s)d(t) \quad (3.1)$$

y is the output, u the input and d is an output disturbance. P(s) is a monic Hurwitz polynomial. P,Q are given by

$$P(s) = s^n + p_1 s^{n-1} + \ldots + p_n \quad (3.2a)$$

$$Q(s) = q_1 s^{n-1} + q_2 s^{n-2} + \ldots + q_n \quad (3.2b)$$

An equivalent representation, more useful for identification purposes is:

$$y(t) = \frac{F(s)-P(s)}{F(s)}y(t) + \frac{Q(s)}{F(s)}u(t) + \frac{P(s)}{F(s)}d(t) \quad (3.3)$$

with F(s) a monic Hurwitz polynomial of degree n:

$$F(s) = s^n + f_1 s^{n-1} + \ldots + f_n \quad (3.4)$$

Defining

$$\theta* \triangleq (f_1-p_1 \ldots f_n-p_n \quad q_1 \ldots q_n)^T \quad (3.5a)$$

$$\phi(t) \triangleq \left(\left(\frac{s^{n-1}}{F(s)}\cdots\frac{1}{F(s)}\right)y(t)\ \left(\frac{s^{n-1}}{F(s)}\cdots\frac{1}{F(s)}\right)u(t)\right)^T \quad (3.5b)$$

we have that

$$y(t) = \theta_*^T\phi(t) + \frac{P(s)}{F(s)}d(t) \quad (3.6)$$

(ϕ is called the regressor, or information vector.)

Suppose now that the plant parameters are not available to us. A reasonable criterion to identify the parameters with, on the basis of input output measurements is a weighted integral squared error.

$$I(\theta) = \lim_{T\to\infty} \frac{1}{T}\int_0^T [H(s)(y-\phi^T\theta)]^2(t)dt \quad (3.7)$$

(In writing down this expression, we implicitly assume that the input and the disturbance are such that the limit indeed exists.) H(s) is a frequency weighting on the instantaneous error $y(t)-\phi^T(t)\theta$, introduced to minimise the effect of the disturbance d on the parameter estimate. Using this criterion, assuming that the regressor contains enough information i.e.

$$\Psi \triangleq \lim_{T\to\infty} \frac{1}{T}\int_0^T [H(s)\phi](t)[H(s)\phi]^T(t)dt > 0 \quad (3.8)$$

we obtain the unique optimal estimate

$$\theta_{opt} = \theta* + \Psi^{-1}\lim_{T\to\infty}\frac{1}{T}\int_0^T [H(s)\phi](t)[H(s)\frac{P(s)}{F(s)}d](t)dt \quad (3.9)$$

$$\triangleq \theta* + \Psi^{-1}\xi \quad (3.10)$$

Expressions (3.8) and (3.9) clearly indicate the role of H(s): suppressing d as much as possible whilst retaining the positive definiteness of Ψ.

Our on-line identification procedure can be derived from (3.7) along the same lines we have obtained the MIT rule from (2.1). We find:

$$\dot{\theta}(t) = g[H(s)\phi](t)\{[H(s)y](t)-[H(s)\phi^T\theta](t)\} \quad (3.11)$$

Defining $\tilde{\theta}(t) = \theta(t)-\theta*$, this can be rewritten as

$$\begin{aligned}\dot{\tilde{\theta}}(t) = -g[H(s)\phi](t)[H(s)\phi^T\tilde{\theta}](t) \\ +g[H(s)\phi](t)[H(s)\frac{P(s)}{F(s)}d](t) \quad (3.12)\end{aligned}$$

This is the obvious generalization of (2.6) for non scalar parameters.

We investigate the properties of (3.11) under the assumption of slow adaptation. As before, θ being slow, averaging theory allows us to analyse the system

$$\begin{aligned}\dot{\bar{\theta}}(t) = -g[H(s)\phi](t)[H(s)\phi]^T(t)\bar{\theta}(t) \\ +g[H(s)\phi](t)[H(s)\frac{P(s)}{F(s)}d](t) \quad (3.13)\end{aligned}$$

instead of (3.12). (3.13) is obtained form (3.12) by formally treating $\theta(t)$ as a constant in the right hand side of (3.12). Moreover, if the homonogeneous part of equation (3.13) is exponentially stable, so is the homogeneous part of equation (3.12), and we have that

$$\tilde{\theta}(t) \to \Psi^{-1}\lim_{T\to\infty}\frac{1}{T}\int_0^T [H(s)\phi](t)[H(s)\frac{P(s)}{F(s)}d]dt + O(g) \quad (3.15)$$

[cfr. Ref[1] pp 48-] Here Ψ is as defined in (3.8). The positive definiteness of Ψ is a necessary and sufficient condition for the exponential stability of (3.12) and (3.13) for sufficiently small ε.

Expression (3.15) is valid only if all limits indeed exist. We summarise in the form of the following theorem.

Theorem 5: Under the conditions that:

A1: u(t) is bounded; d(t) is bounded

A2: H(s), P(s), F(s) are strictly stable

A3: The limits Ψ and ξ in (3.8) and (3.10) exist.

A4: Ψ is a positive definite matrix.

the equation error identification scheme is stable for all $g\varepsilon(0,g^*)$, for some small positive g^*; and the parameter estimate converges exponentially fast to:

$$\theta^{\infty}(t) = \theta^* + \Psi^{-1}\xi + O(g) \qquad (3.17)$$

□

Condition (3.8) or assumption A3 is a persistency of excitation condition, which can be easily interpreted in terms of the input spectrum. The condition number of the matrix Ψ determines the rate of exponential convergence (which is proportional to ε for ε small). Taking u(t) of the form (2.18) one can compute Ψ and show that it is positive definite iff P and Q are coprime, and the input u has at least n distinct frequences ω_m for which $H(j\omega_m)\neq 0$.

Theorem 5 extends the results of the previous section to the vector case. The interpretation is really identical. The only phenomenon this scheme does not display is the "undermodelling" effect, because H(s) is at our disposal.

3.2 Adaptive Feedback Control

We merely indicate how the previous results can be extended to the adaptive feedback control problem. To fix the ideas, assume we are adaptively controlling a linear time invariant plant, using a gradient algorithm - corresponding to an error criterion - for updating the control parameter(s). Assume that the adaptation is slow. Then with one additional provisor, namely that for all values of the controller parameter encountered during adaptation the associated closed loop system is exponentially stable, the earlier conclusions on the convergence of the adaptive algorithm continue to hold. For details, we refer to [5,6,7,8].

4. CONCLUSIONS

We presented an analysis of adaptive algorithms which update the parameters slowly, compared to the dynamics of the closed loop plant. In this framework, we demonstrated that good performance, i.e. nearly optimal with respect to some error measure, requires persistently exciting inputs. The frequency content of the input - perhaps as modified by regression vector filtering - bears on the performance in three different ways:

(a) when there is no noise, and modelling is exact, the rate of convergence of the parameter is effected by the condition number of the matrix defined by a time-average of the product of the regression vector with itself. Persistently exciting inputs are necessary to secure a positive definite matrix, which is a necessary requirement for convergence

(b) When it is not possible for the actual plant to be exactly modelled, it is desirable that the excitations be confined to those frequencies where close modelling can be achieved.

(c) it is desirable to try to filter out the effects of noise, so far as this is possible.

Finally, notice that the implementation of a well performing adaptive algorithm requires a certain amount of a priori knowledge about the plant to be controlled. In particular one needs to have an idea of the relevant time constants of the plant and its d.c. gain in order to quantify the gain of the adaptive algorithm and to specify the required excitation of the input.

REFERENCES

[1] Volosov, 'Averaging in ODE', Russian Mathematical Surveys, Vol.XVII, 1982.

[2] Osburn P.V., Whitaker M.P., Keser A., 'New developments on the design of model reference adaptive control systems', Paper 61-39, Institute of Aeronautical Sciences, 1961.

[3] D.J.G. James, 'Stability of a model reference control system', AIAA Journal, Vol.9, No.5, pp950-952.

[4] P.A. Ioannou and P.V. Kokotovic, Adaptive Systems With Reduced Models, Springer Verlag, 1983.

[5] B.D.O. Anderson, 'Exponential stability of linear equations arising in adaptive identification', IEEE Trans.Auto.Control, Vol.AC-22, Feb. 1976, pp83-88.

[6] R. Kosut, B.D.O. Anderson, I.M.Y. Mareels, 'Stability of theory for adaptive systems: methods of averaging and persistency of excitation', IEEE Trans. Aut. Control, to appear 1986

[7] B. Riedle and P. Kokotovic, 'A stability/instability boundary for distrubance - free slow adaptation and unmodelled dynamics', Proc. 22nd IEEE CDC, Las Vegas, December 1984, pp998-1001.

[8] R.R. Bitmead, B.D.O. Anderson, R. Kosut, C.R. Johnson, 'Stability theorem for the realisation of the strictly positive real condition', Proceedings 23rd CDC, Las Vegas, December 1984, pp1286-1291.

Fig 1 : The MIT rule's problem.

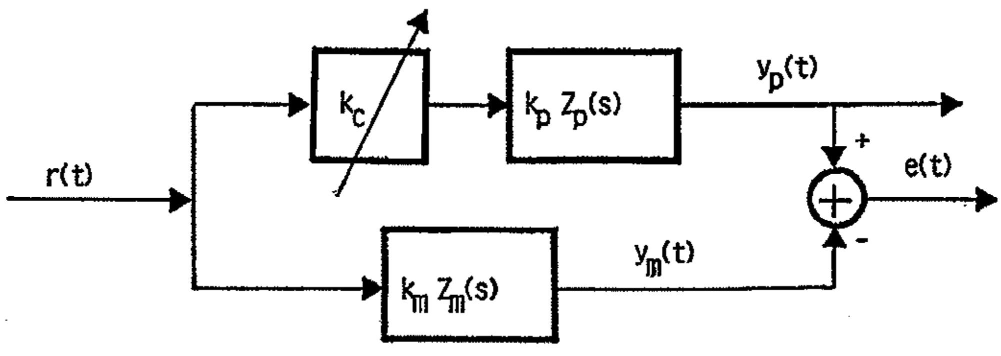
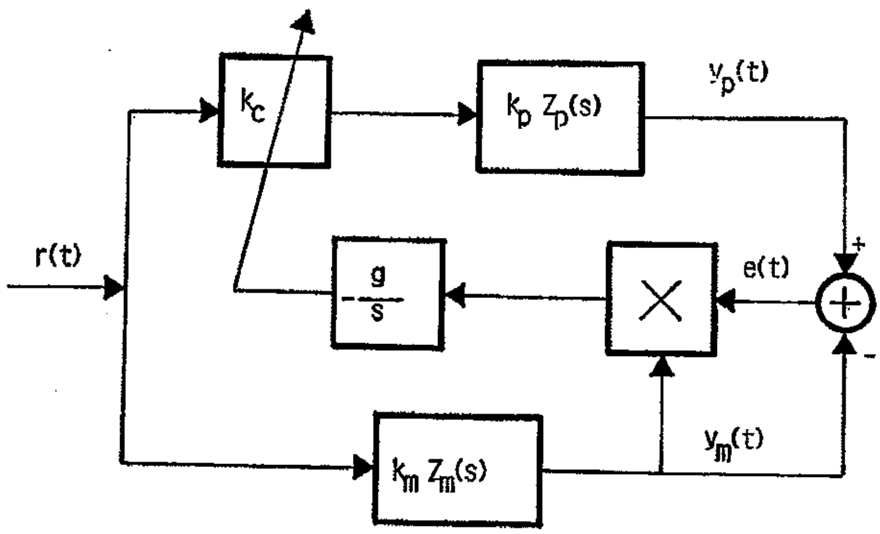


Fig 2. The MIT rule.



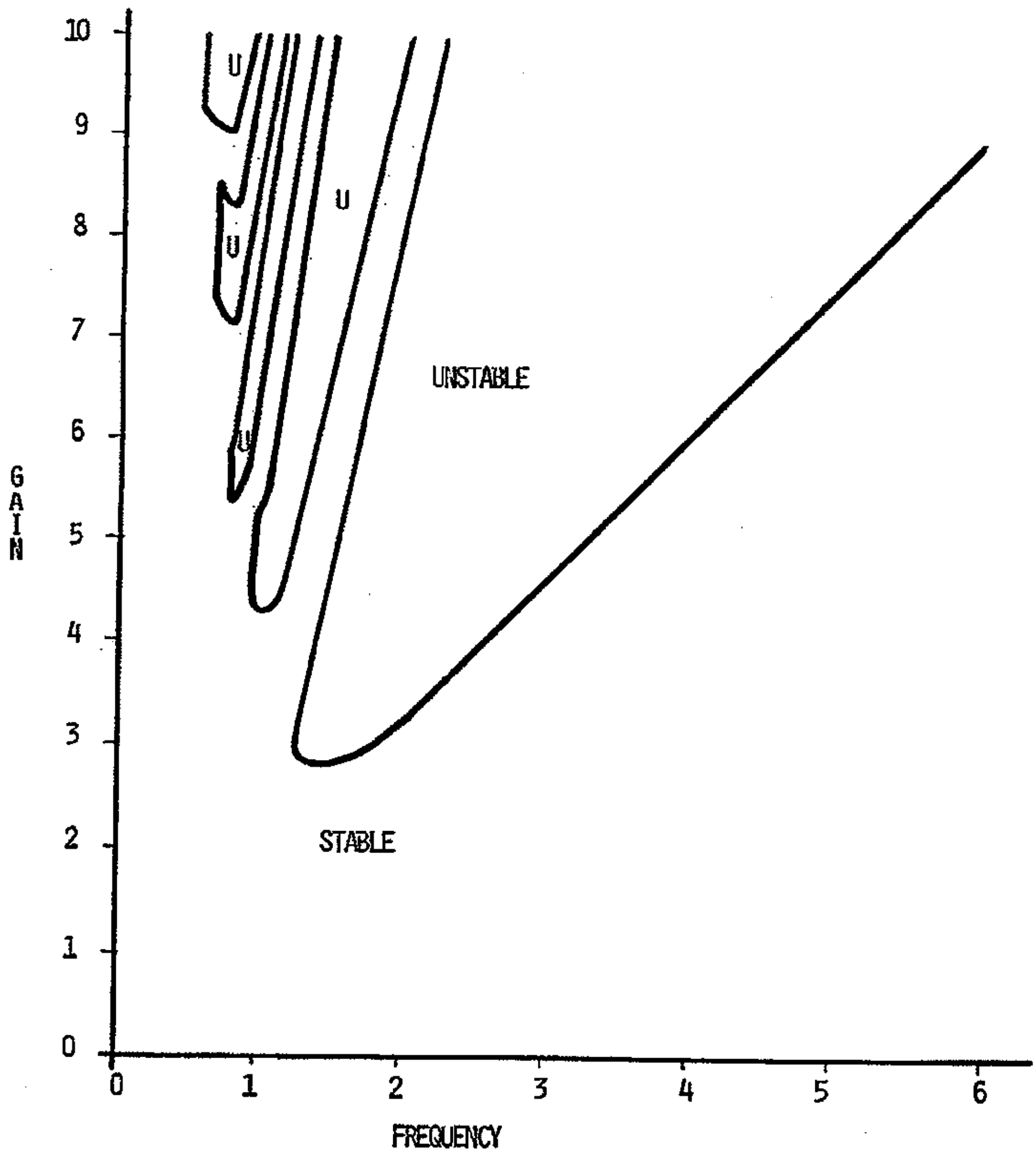


Fig 3. Stability Boundary in Parameter Plane

Fig. 4 Stability boundary in parameter plane

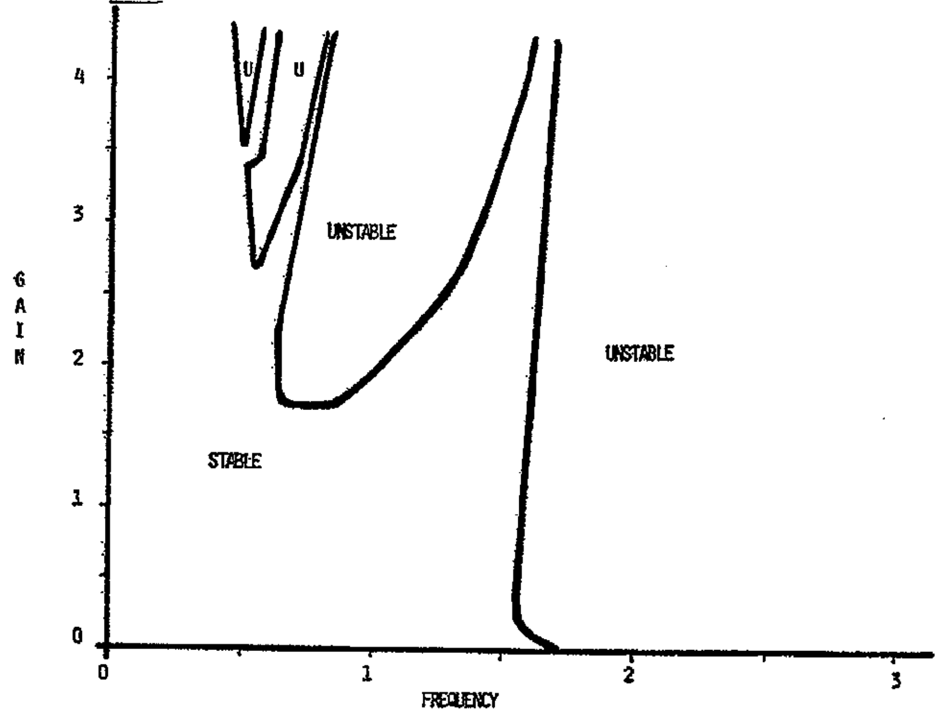


Fig 5. Error Filtering

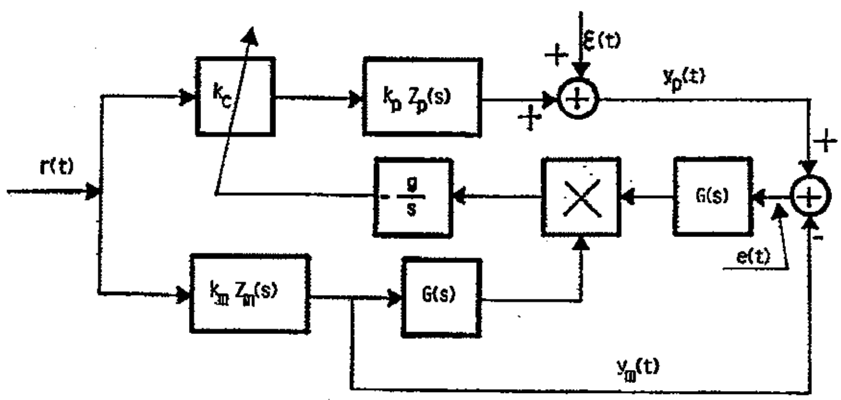


Fig 6. Equation Error Model