INSTRUMENTATION AND AUTOMATION IN THE PAPER, RUBBER, PLASTICS AND POLYMERISATION INDUSTRIES

Edited by

A. R. VAN CAUWENBERGHE

IFAC

INSTRUMENTATION AND AUTOMATION IN THE PAPER, RUBBER, PLASTICS AND POLYMERISATION INDUSTRIES

Proceedings of the 5th International IFAC/IMEKO Conference
Antwerp, Belgium, 11-13 October 1983

Edited by

A. R. VAN CAUWENBERGHE

Published for the

INTERNATIONAL FEDERATION OF AUTOMATIC CONTROL

by

PERGAMON PRESS

OXFORD · NEW YORK · TORONTO · SYDNEY · PARIS · FRANKFURT

U.K.	Pergamon Press Ltd., Headington Hill Hall, Oxford OX3 0BW, England
U.S.A.	Pergamon Press Inc., Maxwell House, Fairview Park, Elmsford, New York 10523, U.S.A.
CANADA	Pergamon Press Canada Ltd., Suite 104, 150 Consumers Road, Willowdale, Ontario M2J 1P9, Canada
AUSTRALIA	Pergamon Press (Aust.) Pty. Ltd., P.O. Box 544, Potts Point, N.S.W. 2011, Australia
FRANCE	Pergamon Press SARL, 24 rue des Ecoles, 75240 Paris, Cedex 05, France
FEDERAL REPUBLIC OF GERMANY	Pergamon Press GmbH, Hammerweg 6, D-6242 Kronberg-Taunus, Federal Republic of Germany

First edition 1984

Library of Congress Cataloging in Publication Data

International IFAC/IMEKO Conference on Instrumentation & Automation in the Paper, Rubber, Plastics & Polymerization Industries (5th : 1983 : Antwerp, Belgium) Instrumentation & Automation in the paper, rubber, plastics & polymerization industries.
(IFAC proceedings series)
Sponsored by the International Federation of Automatic Control, Technical Committee on Applications; co-sponsor, International Measurement Confederation. Includes Index.
1. Paper mills — Automation — Congresses. 2. Rubber industry and trade — Automation — Congresses. 3. Pulp mills — Automation — Congresses. 4. Chemical plants — Automation — Congresses. I. Cauwenberghe, A. van. II. International Federation of Automatic Control. III. International Federation of Automatic Control. Technical Committee on Applications. IV. International Measurement Confederation. V. Title. VI. Title: Instrumentation and automation in the paper, rubber, plastics and polymerization industries. VII. Series.
TS1117.6.153 1983 660.2'81 84–9212

British Library Cataloguing in Publication Data

International IFAC/IMEKO Conference on Instrumentation & Automation in the Paper, Rubber, Plastics & Polymerization Industries (5th : 1983 : Antwerp)
Instrumentation & automation in the paper, rubber, plastics & polymerization industries — (IFAC proceedings series)
1. Process control 2. Engineering instruments
I. Title II. Cauwenberghe, A. van
III. International Federation of Automatic Control IV. Series
670.42'7 TS156.8
ISBN 0-08-031112-1

These proceedings were reproduced by means of the photo-offset process using the manuscripts supplied by the authors of the different papers. The manuscripts have been typed using different typewriters and typefaces. The lay-out, figures and tables of some papers did not agree completely with the standard requirements; consequently the reproduction does not display complete uniformity. To ensure rapid publication this discrepancy could not be changed; nor could the English be checked completely. Therefore, the readers are asked to excuse any deficiencies of this publication which may be due to the above mentioned reasons.

The Editors

Printed in Great Britain by A. Wheaton & Co. Ltd., Exeter

FIFTH INTERNATIONAL IFAC/IMEKO CONFERENCE ON INSTRUMENTATION AND AUTOMATION IN THE PAPER, RUBBER, PLASTICS AND POLYMERIZATION INDUSTRIES

Sponsored by
International Federation of Automatic Control
Technical Committee on Applications (IFAC)

Co-sponsored by
International Measurement Confederation (IMEKO)

Organized by
Belgian Institute of Automatic Control (BIRA)

International Programme Committee
A. R. van Cauwenberghe, Belgium (Chairman)
C. Bekaert, Belgium (Secretary)
R. A. Baxley, Jr., USA
G. De Blende, Belgium
M. De Coster, Belgium
F. D'Hulster, Belgium
G. Dumont, Canada
S. G. Edlund, Sweden
C. Foulard, France
J. Hill, Sweden
K. Hoogendoorn, The Netherlands
F. Huyten, The Netherlands
M. Jedlicka, Czechoslovakia
R. E. Johnston Australia
E. Jutila, Finland
A. Kaya, USA
T. Kemeny, Hungary
W. Koch, Federal Republic of Germany
M. MacLaurin, United Kingdom
P. Raichle Federal Republic of Germany
H. Ruch, Federal Republic of Germany
A. C. Saxena, Canada
K. H. Schmitt, Federal Republic of Germany
J. Seynaeve, Belgium
T. Takamatsu, Japan
P. Uronen, Finland
J. Virkkunen, Finland
D. Waye, Canada
J. Willems, The Netherlands

National Organizing Committee
J. Baert, Belgium (Chairman)
R. Viroux, Belgium (Secretary)
R. van Keilegom, Belgium
L. Prenen, Belgium

CONTENTS

* Survey Paper

RUBBER

PAPER WEB

INJECTION MOULDING

DIAGNOSTICS AND MACHINE HELP MONITORS

BATCH BLENDING SYSTEMS

INTERNAL MODEL CONTROL — THEORY AND APPLICATIONS

M. Morari

California Institute of Technology, Pasadena, California, USA

Abstract. Process control is characterized by severe modelling problems. Therefore, robustness of the control system, that is, stability and acceptable performance in the event of plant parameter changes and sensor and/or actuator failure is of paramount importance. During the past few years a multivariable control system design method, Internal Model Control (IMC), has been developed which addresses specifically these issues. The IMC scheme is transparent, can be easily adjusted on-line and is therefore readily accepted by the operating personnel.

The basic theoretical principles behind IMC are described and parallels are drawn to other design schemes (Model Algorithmic Control, Dynamic Matrix Control, Linear Quadratic Optimal Control, Smith Predictor, etc.). Extensions of IMC to nonlinear systems are indicated. Applications of IMC both in simulations and on pilot plants are discussed.

Keywords. Process control, robustness, time lag systems, sampled data systems, nonlinear control systems.

INTRODUCTION

Many of the modern controller design techniques have not found their way into the process industries despite proven success in aerospace applications. The reasons are the different underlying systems on one hand and the different performance requirements on the other. Space structures are notorious for their large number of modes most of which are only slightly damped. On the contrary, most systems found in the process industries are sluggish, overdamped and their dynamic characteristics can generally be approximated well by a first or second order lag combined with a dead time. A second important distinguishing feature is that many chemical processes are strongly nonlinear and can only be poorly modeled, while the mathematical descriptions of sattelite motions are usually quite accurate. The requirements on the closed loop transient response of a chemical processing system are generally quite lose; the steady state performance is of major importance (no offset). On the other hand, in space applications the problem is most often of the servotype and there is no steady state to worry about. Another important requirement is that for controllers to be accepted by the process industries they have to be easily adjustable on-line by operating personnel without university level training. A final important issue is that the operating region of a chemical plant is usually highly constrained and a controller must be able to take the constraints into account explicitly in order to keep the plant safely within the prescribed bounds.

The dissatisfaction with the ability of the available control system design methods to deal effectively with these issues and the increased power of readily available computer hardware have led a number of research groups in industry and academia to search for new alternatives. The prominent ones have become known as Model Algorithmic Control (Richalet, 1978), Dynamic Matrix Control (Cutler, Ramaker, 1980), Inferential Control (Brosilow, 1979) and Internal Model Control (Garcia, Morari, 1982). Though this was clearly not recognized by most of the developers the principle features which give these methods their power, are identical and will be elucidated next. The key issue is the capability of the new techniques to combine the advantages of open-loop (feed forward) and feedback control and to eliminate their disadvantages.

The advantage of the open loop scheme (Fig. 1A) is that the stability question is trivial (the system is stable when both the controller and the system are stable) and that the controller is easy to design ($g_c=g^{-1}$). The disadvantages are the sensitivity of the performance to modelling errors and the inability to cope with unmeasured disturbances.

With the feedback arrangement (Fig. 1B) the situation is reversed. Modelling errors and unmeasured disturbances can be dealt with effectively but the tuning is complicated by the closed loop stability problem.

We can now augment the open-loop and closed-loop systems as indicated in Fig. 1 C & D without effecting their performance: In Fig. 1C, $\tilde{d}=0$, and therefore the system is still open-loop, in Fig. 1D the two blocks $\tilde{g}$ ($\tilde{g}$ indicates a model of the plant g) cancel each other. Comparing Fig. 1C and D and using the appropriate definitions we arrive at the general structure in Fig. 1E which has all the advantages of both the open-loop and the closed-loop structures: When the model of the plant is perfect ($\tilde{g}=g$) and there are no disturbances (d=0), feedback is not needed and structure E behaves identically as structure A. When there are modelling errors and/or disturbances feedback is needed and structure E behaves identically as structure B. Because the plant model $\tilde{g}$ appears explicitely in E, this structure is referred to as the Internal Model Control (IMC) structure. As a first approximation we can say that the controller g_c in E can be designed with the simplicity of an open-loop controller but that the structure E has all the nice performance characteristics of a feedback system. Obviously, the situation is not quite as straightforward, but this argument should provide sufficient motivation to explore the theoretical properties of the IMC structure in more depth. Striving for clarity rather than generality in this expository paper we will start with a discussion of continuous single-input-single-output (SISO) systems. After that the results will be extended to multi-input-multi-output (MIMO) and sampled data systems. The paper will conclude with a number of comparative simulation and experimental studies.

SISO SYSTEMS

From the block diagram for the IMC structure (Fig. 1E) follow the relationships

$$u = \frac{g_c}{1+g_c(g-\tilde{g})}(y_s-d) \tag{1}$$

$$y = \frac{gg_c}{1+g_c(g-\tilde{g})}(y_s-d)+d \tag{2}$$

The advantages of the IMC structure discussed qualitatively in the introduction can be stated more precisely in the form of three properties which can be proved easily from (1) and (2).

Property P1 (Dual Stability): Assume the model is perfect ($g=\tilde{g}$). Then the closed loop system in Fig. 1E is stable if the controller g_c and the plant g are stable.

Property P2 (Perfect Control): Assume that the controller is equal to the model inverse ($g_c=\tilde{g}^{-1}$) and that the closed loop system in Fig. 1E is stable. Then $y(t) = y_s(t)$ for all $t > 0$ and all disturbances d(t).

Property P3 (Zero Offset): Assume that the steady state gain of the contoller is equal to the inverse of the model gain ($g_c(0)=\tilde{g}(0)^{-1}$) and that the closed loop system in Fig. 1E is stable. Then for an asymptotically constant setpoint ($\lim_{t\to\infty} y_s(t)=\bar{y}_s$) and asymptotically constant disturbances there will be no offset ($\lim_{t\to\infty} y(t)=\bar{y}_s$).

P1 simply expresses the fact that unless there are modelling errors and as long as the open loop system is stable, the stability issue is trivial. P2 reasserts that the ideal open-loop controller leads to perfect closed-loop performance when the IMC structure is employed. P3 states that integral-type control action can be easily achieved without the need of introducing additional tuning parameters.

Superficially these properties seem too good to be true. However, it should be emphasized that structures B and E are equivalent (Fig. 1) as is apparent from the following transformation equations

$$g_c = \frac{c}{1+cg} \tag{3}$$

$$c = \frac{g_c}{1-\tilde{g}g_c} \tag{4}$$

and therefore the properties can be easily explained. Whatever is possible with structure B is possible with structure E and vice versa. We know intuitively that P2 requires an infinite controller gain and this is confirmed by substituting $g_c = \tilde{g}^{-1}$ in (4). By setting $g_c(0) = \tilde{g}(0)^{-1}$ as postulated for P3 we find $c(0) = \infty$ which implies integral control action as expected.

The advantage of the IMC structure is twofold. In simplified terms we can say the larger the "gain" the better the performance. In the conventional structure B the objective is to make the "gain" as large as possible without causing instability. Simultaneously attention is to be paid to other criteria like robustness to modelling errors and to constraints like input saturation. According to P2, with IMC we can start with a stable closed loop system with perfect control. Thus the first design problem is eliminated altogether and full attention can be devoted to the additional criteria and constraints. The second advantage of IMC is that the design philosophy lends itself much better to be extended to multivariable and nonlinear systems as we will see later.

There are several reasons why the "perfect controller" implied by P2 cannot be realized in practice.

1) Right half plane (RHP) zeros: If the model has a RHP zero, the controller $g_c = \tilde{g}^{-1}$ has a RHP pole and if $\tilde{g} = g$ the closed-loop system will be unstable according to P1.

2) Time delay: If the model/plant contains a time delay, the controller $g_c = \tilde{g}^{-1}$ is predictive and cannot be realized by a physical system.

3) Constraints on the manipulated variables: If the model is strictly proper then the perfect controller $g_c = \tilde{g}^{-1}$ is improper which implies $\lim_{\omega\to\infty} |g_c| = \infty$. Thus infinitesimally small high frequency disturbances would give rise to infinitely large excursions of the manipulated variables which is physically impossible.

4) Modelling error: If $\tilde{g} \neq g$, P1 does not hold and the closed loop system will generally be unstable for the controller $g_c = \tilde{g}^{-1}$.

In order to deal with these four issues the ideal of perfect control has to be abandoned. This is done in two steps.

1) The model $\tilde{g}$ is factored

$$\tilde{g} = \tilde{g}_+ \cdot \tilde{g}_- \tag{5}$$

such that $\tilde{g}_-^{-1}$ is stable and causal and $\tilde{g}_+(0) = 1$.

2) The controller is

$$g_c = \tilde{g}_-^{-1} f, \tag{6}$$

where f is an adjustable low-pass filter which guarantees that g_c is proper and the closed loop system is robust. By definition of the factorization (5) g_c is realizable.

The design of the IMC controller involves as a first step the factorization of $\tilde{g}$ in some suitable manner and subsequently the selection and tuning of the filter f. In the design procedure we have developed, a perfect model is assumed for the first step and the factorization is performed to optimize some performance measure. In the second step the filter is selected to make the closed loop system robust against modelling errors.

Factorization of g

In the absence of modelling error ($\tilde{g}=g$) and with $g_c = g_-^{-1}$ the control error e is

$$-e = y - y_s = (g_+-1)(y_s-d) \tag{7}$$

g_+ can be selected to minimize some function of the error for a specific input (y_s-d), for example

$$\int_0^\infty e^2 dt = \frac{1}{\pi}\int_0^\infty (g_+-1)^2(y_s-d)^2 d\omega \tag{8}$$

Frank (1974) proposes a general procedure for this optimization which is valid for arbitrary inputs. In summary, the integral square error optimal factorization for step inputs is as follows.

Theorem 1 (Frank, 1974): Let

$$g = \frac{(s-z_1)(s-z_2)\cdots(s-z_m)}{(s-p_1)(s-p_2)\cdots(s-p_n)} \tag{9}$$

where $z_1, \ldots z_i > 0$ and $z_{i+1}, \ldots z_m < 0$.

Then the optimal g_+ minimizing (8) is

$$g_+ = \frac{(-s+z_1)\cdots(-s+z_i)}{(s+z_1)\cdots(s+z_i)} \tag{10}$$

The optimal ISE is $\sum_{j=1}^{i} \frac{2}{z_j}$.

Theorem 3 (Frank, 1974): Let

$$g = \frac{n(s)}{d(s)} e^{-\theta s} \tag{11}$$

where n(s) and d(s) are polynomials. Then the optimal g_+ minimizing (8) is

$$g_+ = e^{-\theta s} \tag{12}$$

Filter Design

The filter is required to make the controlled g_c proper and thus realizable and to make the closed loop system robust to modelling errors. For realizability the order difference between numerator and denominator polynomial of the filter should be at least n-m (c.f. (9)). In the absence of modelling errors the closed loop response is

$$y = g_+ f (y_s-d) + d \tag{13}$$

and the filter could be selected to minimize, for example

$$\int_0^\infty e^2 dt = \frac{1}{\pi}\int_0^\infty (g_+f-1)^2(y_s-d)^2 d\omega \tag{14}$$

Frank (1974) provides a table of optimal filters with one adjustable parameter which has a direct effect on the speed but not the shape of the closed loop response. The main objective of the filter however is to guarantee reasonably good and at least stable behavior in the presence of plant/model mismatch.

The model uncertainty is commonly assumed to be of the multiplicative type

$$g(s) = \tilde{g}(s)\,(1+\ell(s)) \tag{15}$$

where $\ell(s)$ is constrained by a real non-negative function

$$|\ell(i\omega)| \leq \bar{\ell}(\omega) \tag{16}$$

or

$$\frac{|g(i\omega) - \tilde{g}(i\omega)|}{|\tilde{g}(i\omega)|} \leq \bar{\ell}(\omega) \tag{17}$$

This implies that the Nyquist plot of the plant $g(s)$ can lie within a band around the Nyquist plot of the model $\tilde{g}(s)$. This band is described by a set of circles centered at $\tilde{g}(i\omega)$ with radius $|\tilde{g}(i\omega)|\ \bar{\ell}(\omega)$ (Fig. 2). At high frequencies the model is essentially always of lower order than the plant. Therefore

$$\lim_{\omega\to\infty} \bar{\ell}(\omega) = 1 \tag{18}$$

For the controller (6) and the uncertainty (16,17) the closed loop expression (2) becomes

$$y = \frac{\tilde{g}_+ f\,(1+\ell)}{1+\tilde{g}_+ f\ell}\,(y_s - d) + d \tag{19}$$

Theorem 3: The closed loop system is stable for all uncertainties $\ell(s)$ satisfying (16,17) if and only if $|f(i\omega)| \leq \frac{1}{\bar{\ell}(\omega)}$

Proof: Follows the same outline as in Garcia and Morari (1984). Thm. 3 provides a design rule for the filter given a specific uncertainty range for the plant. It guarantees stability but for adequate performance it is desirable to limit the maximum peak of the closed loop transfer function

$$\frac{|g_+ f(1+\ell)|}{|1 + g_+ f\ell|} \leq 1 + \alpha, \ \forall\omega \tag{20}$$

where $\alpha \approx 0.4$ would be a reasonable requirement for process control applications. Except for very simple analytic expressions for ℓ, the search for the filter f to satisfy (20) has to be performed numerically.

According to P3 the requirement for integral control is that

$$g_c(0) = \tilde{g}_-^{-1}(0) f(0) = \tilde{g}^{-1}(0) \tag{21}$$

and because of $\tilde{g}_+(0) = 1$, $f(0) = 1$. It then follows from Thm. 3 that integral control is impossible if the steady state gain error can exceed 100% ($\bar{\ell}(0)>1$). This is expressed more precisely in Thm. 4.

Theorem 4 (Morari, 1983a): There exists a stabilizing filter with $f(0) = 1$ for the closed loop system in Fig. 1E with the controller (6) if and only if $g(0)\tilde{g}(0)^{-1} > 0$ or with other words if the steady state gain of the model and the system have the same sign.

IMC and the Smith Predictor

A look at IMC for systems with a time delay (Fig. 3A) shows the complete equivalence with the Smith Predictor (Fig. 3B) where

$$c = \frac{1}{\tilde{g}_-}\ \frac{f}{1+f} \tag{22}$$

This leads to a number of important conclusions:

- IMC includes time delay compensators in a natural manner.
- The factorization (5) implied by the Smith Predictor is only optimal in the sense of ISE for step inputs.
- Robustness studies via the IMC structure have led to very simple filter design rules for the case that the modelling error is only in the time delay (Brosilow, 1979; Clinch, 1982). Let the possible time delay error be $\pm\varepsilon$ and let a first order filter $1/(\tau s+1)$ be sufficient to make the controller g_c realizable. If τ is selected equal to (1.4 times) ε, then the maximum closed loop amplitude ratio peak will not exceed 2 (1.4). With these filter settings it can be shown that the combination of a PI controller with a Smith Predictor always outperforms a PID controller alone. Therefore all the reports about the impracticality and poor performance of Smith Predictors because of their sensitivity are myths generated by incorrect tuning procedures.

Summary

The IMC design procedure consists of two steps.

Step 1: Factor the model transfer function into an invertible part $\tilde{g}_-$ and a noninvertible part $\tilde{g}_+$. If the factorization is performed according to Thm. 1 & 2 and if f is a low pass filter chosen to make g_c proper the controller $g_c = \tilde{g}_-^{-1} f$ minimizes the ISE for step changes in the inputs.

Step 2: In the presence of model uncertainty the filter time contants have to be increased to satisfy the condition of Thm. 3.

The simplicity of the design procedure should be apparent. In the absence of modelling errors there is only <u>one</u> adjustable parameter, the speed of response determined by the filter, which can be selected by the designer at will. If modelling errors are present the problem of robustness has to be addressed by adjusting the filter.

The tuning is so transparent because the designer selects the closed loop transfer function $g_+ f$ directly (see (13)). A "fast" filter pushes the system hard and increases the possibility of an instability if the model is inaccurate. If not much is demanded from the system (conservative filter) it will be stable even when the modelling errors are severe. The same could be accomplished with the classic feedback structure (Fig. 1B) but IMC uses the inverse of $\tilde{g}_-$ explicitely instead of approximating it indirectly by selecting a high controller gain. Also in the classic structure a series of parameters in c would have to be adjusted simultaneously to have the effect of the single IMC filter parameter.

Several questions of theoretical and practical interest are currently the focus of our research efforts:

1) The uncertainty description (17) destroys phase information and can therefore lead to very conservative control systems.

2) No simple, practically effective filter design methods to satisfy criteria like (20) have been proposed yet.

3) In deciding on the optimal factorization (5) attention should be paid to modelling errors and not only to the ISE. This can have a profound effect on the performance (Brosilow, 1983) but only demostrative case studies and no fundamental analyses are available to date.

MIMO SYSTEMS

The basic structure, properties, relationships and design philosophy carry over to the multivariable case and will not be elaborated on in detail. Transfer functions are replaced by transfer matrices which will be denoted by capital letters. Again all systems will be assumed to be strictly open-loop stable and to have the same number of inputs and outputs. Then for the IMC structure multivariable equivalent of Fig. 1E we find

$$y = G(I+G_c(G-\tilde{G}))^{-1}G_c(y_s-d) + d \qquad (23)$$

$$u = (I+G_c(G-\tilde{G}))^{-1}G_c(y_s-d) \qquad (24)$$

Properties P1-P3 carry over simply by substituting matrices for scalars. The factorization of G into an invertible and a noninvertible part and the design of the robustness filter need special attention. It should be pointed out at this point that zeros of transfer matrices can be defined in a number of ways (MacFarlane, 1974) but that in general they bear no connection to the zeros of the individual transfer matrix elements. For stable systems the RHP zeros can be determined from the determinant of the transfer matrix. The factorization of time delays is complicated by the fact that in general the time delays in the different matrix elements are different.

<u>Factorization of G</u>

In principle G_+ could be determined again by minimizing a scalar function of the error

$$-e = y - y_s = (G_+-I)(y_s-d) \qquad (25)$$

e.g. the ISE. Though such a procedure has been developed (Frank, 1974) it is not recommended in practice because it is extremely cumbersome and also requires a relative weighting of the different output errors which is usually quite arbitrary. The following simple results have emerged from the investigations by Holt and Morari (1983, 1984). Note that without modelling errors, and for

$$G_c = G_-^{-1} F \qquad (26)$$

reduces to

$$y = G_+ F (y_s-d) + d \qquad (27)$$

Thus, the type of factorization determines directly the closed loop response.

<u>Theorem 5</u> (Holt and Morari, 1984): Let the MIMO system G(s) have RHP zeros at $s=z_1, z_2, \ldots z_i$. Then, in general, the "bad" effect of the RHP zeros can be localized to any particular output,

$$G_+ = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & 0 & \\ & & 1 & & & \\ x \cdots x & & \frac{(-s+z_1)\cdots(-s+z_i)}{(s+z_1)\cdots(s+z_i)} & & x \cdots x \\ & & & 1 & & \\ & 0 & & & \ddots & \\ & & & & & 1 \end{bmatrix} \qquad (28)$$

where all the off-diagonal elements are zero except in the row which contains the RHP zeros.

For example, consider the system

$$G(s) = \frac{1}{s+1}\begin{bmatrix} 1 & 1 \\ 1+2s & 2 \end{bmatrix}$$

which has a zero at s = 1/2. Three possible factorizations are shown below together with the ISE resulting from a unit step change in both set points

$$G_+^1(s) = \begin{bmatrix} \frac{-2s+1}{2s+1} & 0 \\ 0 & \frac{-2s+1}{2s+1} \end{bmatrix}; \quad G_+^2 = \begin{bmatrix} 1 & 0 \\ \frac{8s}{2s+1} & \frac{-2s+1}{2s+1} \end{bmatrix};$$

ISE = 8 ISE = 4

$$G_+^3 = \begin{bmatrix} \frac{-2s+1}{2s+1} & \frac{2s}{2s+1} \\ 0 & 1 \end{bmatrix}$$

ISE = 1

The optimal G_+ can be found using the mentioned matrix factorization procedure (Frank, 1974):

$$G_+^4(s) = \frac{1}{5(1+2s)} \begin{bmatrix} 5-6s & 8s \\ 8s & 5+6s \end{bmatrix}; \quad ISE = \frac{4}{5}$$

For a different set of inputs or a different weighting of the outputs the ISE-optimal factor $G_+(s)$ would be different. Thus striving for ISE optimality does not appear a very practical proposition. Factorizations of the type $G_+{}^1$, $G_+{}^2$ & $G_+{}^3$ are much easier to obtain and allow the designer to clearly indicate his preference. If a decoupled response is sought $G_+{}^1(s)$ is the answer. If output 1 is more important $G_+{}^2$ should be selected, if output 2 is critical $G_+{}^3$ is the best candidate.

Similarly, in the case of time delays a trade-off between the speed of the closed loop response and decoupling is possible. For example three possible factorizations for

$$G = \begin{bmatrix} 0 & e^{-2s} \\ -e^{-2s} & 1 \end{bmatrix}$$

are

$$G_+^1 = \begin{bmatrix} e^{-4s} & 0 \\ 0 & e^{-2s} \end{bmatrix}; \quad G_+^2 = \begin{bmatrix} e^{-2s} & 0 \\ (1-e^{-2s}) & e^{-2s} \end{bmatrix}$$

$$G_+^3 = \begin{bmatrix} e^{-4s} & e^{-2s}(1-e^{-2s}) \\ 0 & 1 \end{bmatrix}$$

G indicates that output 1 can react only after two time invervals, output 2 can react immediately. These figures are a lower bound on the response time but they are not an indication of the actual settling time. If both outputs are equally important and decoupling is chosen, $G_+{}^1$ provides an upper bound on the settling time. This is verified by $G_+{}^2$, where preference is given to the first output which settles in minimum time (cf. G), at the cost of decoupling and a maximum settling time for the second output (cf. $G_+{}^1$). Analogously, in $G_+{}^3$ preference is given to the second output. Holt & Morari (1983) have shown that a diagonal G_+ which renders G_-^{-1} causal is "optimal" if and only if the rows and columns of G can be rearranged such that the smallest time delay of each row is on the diagonal. For example, the Wood & Berry (1973) distillation column has the transfer matrix

$$G(s) = \begin{bmatrix} \frac{12.8\ e^{-s}}{16.7s+1} & \frac{-18.9\ e^{-3s}}{21s+1} \\ \frac{6.6\ e^{-7s}}{10.9s+1} & \frac{-19.4\ e^{-3s}}{14.4s+1} \end{bmatrix} \qquad (29)$$

Here the smallest time delays are on the diagonal and therefore $G_+ = \text{diag}(e^{-s}, e^{-3s})$ is "optimal". The lower and upper bounds on the settling time coincide.

Contrary to the results obtained for systems involving RHP zeros, the effects of time delays are structured, i.e. they are generally associated with a particular output and cannot be shifted around.

Filter Design

Again the function of the filter is twofold: It serves to make the controller G_c (26) proper and thus relizable and to provide robustness against modelling errors for the closed loop system. The realizability issue can be resolved trivially simply by providing enough poles in the filter. The main objective of the filter, namely to guarantee reasonably good, but at least stable behavior is the presence of plant-model mismatch is more difficult to accomplish.

As shown in Fig. 4 the multivariable multiplicative uncertainties can act either on the imputs (L_I) or the outputs (L_O)

$$G(s) = \tilde{G}(s)(I+L_I(s)) \qquad (30A)$$

$$G(s) = (I+L_O(s))\tilde{G}(s) \qquad (30B)$$

$$\|\tilde{G}^{-1}(G-\tilde{G})\| < \ell_I(\omega) \qquad (31A)$$

$$\|(G-\tilde{G})\tilde{G}^{-1}\| < \ell_O(\omega) \qquad (31B)$$

where ℓ_I, ℓ_O are scalar functions defined on the positive reals. These functions do not allow to distinguish between uncertainty localized in one element and uncertainty "spread" over all elements. This might or might not be disadvantageous depending on how much uncertainty information is available. Also let us define

$$\bar{\ell}(\omega) = \text{Max}(\ell_I(\omega), \ell_O(\omega)) \qquad (32)$$

What norms whould be used will depend on the application. Here we will use the spectral norm

$$\|G\| = \max_i \lambda_i^{\frac{1}{2}} (G^*G) \qquad (33)$$

which is compatible with the Euclidean vector norm. We will employ the following notation for the singular values,

$$\sigma_M(G) = \lambda_{max}^{\frac{1}{2}}(G^*G)$$

$$\sigma_m(G) = \lambda_{min}^{\frac{1}{2}}(G^*G)$$

It can be shown that

$$\sigma_m(G)\ \|u\| \leqslant \|Gu\| \leqslant \sigma_M(G)\ \|u\| \qquad (34)$$

Thus the maximum singluar value is a natural definition of gain for multivariable systems. Substituting the uncertainty description (30-32) and the controller (26) into (23) the following guidelines emerge for the design of the multivatiable filter.

Theorem 6 (Grossmann & Morari, 1983): The closed loop system is stable for all uncertainties (30A) or (30B) satisfying (31A) or (31B) and (32) if

$$\|F\| \leqslant \frac{1}{\gamma(\tilde{G})\ell(\omega)} \qquad (35)$$

where

$$\gamma(\tilde{G}) = \|\tilde{G}^{-1}\|\,\|\tilde{G}\| = \frac{\sigma_M(\tilde{G})}{\sigma_m(\tilde{G})} \qquad (36)$$

is the condition number of $\tilde{G}$. γ is a measure of singularity.

Comparing Thm. 6 with Thm. 3 we note that the bound on the filter gain is not only inversely proportional to the uncertainty (ℓ) but also to the condition number. Because $\gamma \geqslant 1$ this implies that ill-conditioned systems can amplify modelling errors in a manner unknown in SISO systems. This is probably the main source of design difficulties in MIMO systems.

Thm. 6 guarantees stable but not necessarily good performance. In our experience (35) is generally too conservative because the uncertainty description (30), (31) is inherently conservative. It is likely that some recent results by Doyle (1983) lead to a more practical result.

According to P3 the requirement for MIMO integral control is that

$$G_c(0) = \tilde{G}_-^{-1}(0)F(0) = \tilde{G}^{-1}(0) \qquad (37)$$

and if we define $\tilde{G}_+(0) = I$, $F(0) = I$. Thus when $\gamma(\tilde{G}(0))\ell(0) > 1$ the existence of a filter with unity steady state gain ($F(0) = I$) is not guaranteed any more. Two tighter theorems provide more information.

Theorem 7 (Morari, 1983a): There exists no stabilizing filter with $F(0) = I$ for the MIMO closed loop system in Fig. 1E with the controller (26) if $\det(G(0)\tilde{G}(0)^{-1}) \leqslant 0$.

Theorem 8 (Morari, 1983a): There exists a stabilizing filter with $F(0) = I$ for the MIMO closed loop system in Fig. 1E with the controller (26) if all the eigenvalues of the matrix product $G(0)\tilde{G}(0)^{-1}$ are in the RHP.

Comparing Thm's. 4 and 7 we note that in MIMO system the eigenvalues of the steady state gain matrix play a similar role as the gain of SISO system. In SISO system the sign of the gain is usually known from physical arguments and thus the condition postulated in Thm. 4 can be satisfied easily. In ill-conditioned MIMO systems the accuracy of $\tilde{G}$ required by Thm. 8 can often be excessive. Also, Thm. 8 is only sufficient while Thm. 4 was also necessary. When the number of eigenvalues of $G(0)\tilde{G}(0)^{-1}$ in the LHP is odd there clearly exists no stabilizing filter (Thm. 7), when it is even there could exist one.

MIMO IMC and Multivariable Time Delay Compensation

After many attempts in the literature (Alevisakis & Seborg, 1973; Ogunnaike & Ray, 1979) of varying degree of success and restrictiveness the IMC structure points out a new way of multivariable time delay compensation. For the case of equal time delays in all transfer matrix elements the IMC proposed compensator structure (Fig. 3B) reduces to that of Alevisakis & Seborg (1973) or Ogunnaike & Ray (1979). Otherwise it generally does not remove all time delays from the transfer matrix but does so selectively (see the definition of $\tilde{G}_-$). This always leads to significant performance improvements as will be demonstrated in the example section. More details are available from Holt and coworkers (1984). To date the robustness results (Thm. 6) have not been translated into simple filter design rules to guarantee satisfactory performance in the presence of time delay modelling errors.

Summary

The IMC design procedure for MIMO systems consists of two steps:

Step 1: The model transfer matrix has to be factored into an invertible part $\tilde{G}_-$ and a noninvertible part $\tilde{G}_+$. In this factorization the designer has some flexibility to localize the detrimental effect of the nonminimum phase elements on one or the other output and to choose a decoupled response or to allow full or partial interactions.

Step 2: In the presence of model uncertainty a low pass filter has to be introduced, for example of the form

$$F = \mathrm{diag}(1/(\tau_i s+1)^{k_i}) \qquad (38)$$

By choosing the filter time constants and the filter order sufficiently large, Thm. 6 guarantees that the system can be stabilized without sacrificing integral control action as long as the condition of Thm. 8 is satisfied.

The tuning procedure is inherently simple and transparent. If the filter (38) is used then for each output there is a single tuning parameter τ_i which affects directly the speed of response of the particular output. If a fast response is demanded, a good model is required.

The technique bears some resemblance to the method of "decoupling" prominent in process control applications. The analysis here has shed light onto the old question when complete decoupling might be detrimental to performance. This is the case when time delays or RHP zeros are present in the transfer matrix. Furthermore some insight has been gained into the question of robustness.

In terms of open research questions the same types of problems as listed for SISO systems await solution. Some help should be available from the works of Zames (1981) and Doyle (1982).

DISCRETE TIME SYSTEMS

Most modern control systems are microprocessor or minicomputer based and unless the sampling rate is very fast a discrete time domain analysis and synthesis is more appropriate. All the results derived for SISO and MIMO continuous systems in the preceding sections can be easily rederived for discrete time systems. In most cases equivalent properties and theorems are found -- now formulated in terms of z-transforms instead of Laplace Transforms. However, it turns out that Thm. 8 can be considerably strengthened for sampled data systems.

Theorem 9 (Garcia & Morari, 1984): Assume that the robustness filter F(z) is diagonal and of the exponential type

$$F(\gamma) = \mathrm{diag}\,\frac{1-\alpha_i}{1-\alpha_i z^{-1}} \quad 0 \leqslant \alpha_i < 1 \qquad (39)$$

and that $G_c = \tilde{G}_-^{-1} F$. There exists an α^* $(0 \leqslant \alpha^* < 1)$ such that the system is closed loop stable for all α_i in the open interval $\alpha^* \leqslant \alpha_i < 1$ if and only if G and $\tilde{G}$ satisfy

$$\mathrm{Re}\{\lambda_j(G(1)\tilde{G}(1)^{-1})\} > 0 \quad \forall_j \qquad (40)$$

where $\lambda_j(A)$ denotes the jth eigenvalue of A.

Because of the discrete nature a first order filter is sufficient for stability as long as (40) is satisfied. Depending on the type of uncertainty a higher order filter can be required for continuous systems. Also, the condition (40) is necessary and sufficient for the existence of a range of α_i's $(\alpha^* \leqslant \alpha_i < 1)$ for which the system is closed loop stable. Note, however, that some specific set of α_i's (instead of the open interval extending to 1) might exist which stabilizes the closed loop system even when (40) is not satisfied. In this case the system will become unstable when the α_i's are increased. This conditional stability makes the on-line tuning much more difficult and is highly undesirable.

The role of the filter for the robustness of SISO control systems is illustrated by Reid and coworkers (1979) in a specific case study, but the general theoretical explanation offered by Thm. 9 is not provided.

The IMC structure starts to display its full power when instead of simply translating from continuous to discrete time, specific use is made of the discrete formulation in the computation of the control law. The controller (26) is composed of the inverse of the invertible part of the model and the filter. This combination can be interpreted as an approximate inverse of the model constructed to be stable and to avoid excessive actions of the manipulated variable or at least to be realizable. We can find approximate inverses in an alternate manner which offers increased flexibility.

The process model can be employed to predict the outputs resulting from a series of inputs. Or alternatively, desired outputs can be prescribed and the inputs could be calculated such that the predicted outputs follow the presctibed outputs in some "optimal" manner. If one requires the predicted values to agree with the prescribed ones exactly the system inputs resulting from the solution of this matching problem will be the same as would be obtained by an inversion of the process model. If one requires the predicted values only to be close to the desired ones in the least square sense, for example, the solution of the optimization problem will provide an approximate inverse of the process model. The characteristics of the approximate inverse can then be affected by the choice of weighting matrices in the least squares objective function. This method of computing the control law is referred to in the literature as "model-predictive control law formulation".

We can pose the following problem to be solved at time $\bar{k}$ subject to the model equations relating u and y