15th Triennial World Congress, Barcelona, Spain

A HISTORICAL VIEW OF MULTIVARIABLE FREQUENCY DOMAIN CONTROL

David H. Owens

Department of Automatic Control and Systems Engineering, University of Sheffield, Mappin Street, Sheffield S1 3JD

Abstract: The late sixties and early to mid seventies saw and explosion of interest and innovation in the area of frequency domain methods in the analysis and design of multivariable control systems. The emergence of multivariable frequency domain control design techniques in the 1960's and 1970's was based on traditional views of the role of the designer in the design process. Classical influences on the emerging Nyquist and root-locus theory were therefore natural and, it can be argued, very successful. This paper presents a personal review of the ideas, concepts and techniques from the perspective of a researcher active in the area at the time. The paper will aim to provide a personal insight into the issues and provide an opportunity for young researchers to review the wide range of contributions made before the current use of H-infinity (and related optimisation) methods become so prevalent.

Keywords: Feedback control, Frequency response methods, Multivariable control systems, Nyquist Stability Theory, Computer-aided system design, Root locus

1. INTRODUCTION

Multivariable control systems design is now a relatively mature discipline with many of the issues raised by the generalisation of classical frequency domain methods solved in a form that has found useful application and acceptance in the field. This is not to say that the area has been fully examined and satisfactory conclusions reached in all areas, but it does say that the progress has been substantial leaving issues to be resolved that are of some complexity and in urgent need of attention. The predominance of and increasing focus on H-infinity methods and concepts in current thinking has been the product of many years research and thought driven by need and the increasing power of computer systems and the increased sophistication and potential of software. Further progress can be expected but it may rely on new paradigms, techniques and tools of analysis.

At this time, it is essential that innovation is sought but, I argue, it is also essential to build on the past in an effective and constructive way. This past includes H-infinity approaches but it also includes methods that have been recorded, used and now omitted from the day-to-day vocabulary of the field. This can, and should, be regarded as a natural evolution of the essentially human process of research in the area but it should also be regarded with concern if the richness of the conceptual base is lost to new and young researchers in the field. This paper aims, in part, to address this issue by providing a personal view of the developments in the late sixties and early to mid seventies and the explosion of interest and innovation in the area of frequency domain methods in the analysis and design of multivariable control systems. The emergence of multivariable frequency domain control design techniques in the 1960's and 1970's assumed the availability of computing resources but was conceptually based on traditional views of the role of the designer in the design process. Classical influences on the emerging Nyquist and root-locus theory were therefore natural and, it can be argued, very successful. This reliance on the human input was both a strength and a weakness but in whatever emerges in the future, the development of the human input, either through education and training or through the encouragement of the tenuous and vital ingredient of intuition, will be essential.

As an employee (and external London) PhD student working for the United Kingdom Atomic Energy Authority in the period 1969-1973, the author was fortunate in hitting the research community at a time when multivariable frequency domain techniques were rapidly developing. As a trained physicist, he naturally sympathised with the model-based and physically relevant approach. This influenced his choice of research programme and it also influences the contents of this paper.

The paper and presentation are unashamedly based on the now classic edited text of reprints by MacFarlane (1979) and the 1978 text by the author. All readers are referred to these texts for a view of the period from the late seventies.

The presentation is guided by the historical sequence but this is modified to provide a coherent summary of the concepts and their interrelationships.

2. MULTIVARIABLE FEEDBACK CONTROL SYSTEMS

The times were dominated by linear theory and linear controllers where the l-input m-output system has a state space model (A,B,C,D) with $m \times l$ transfer function matrix (TFM)

$$G(s) = C(sI - A)^{-1}B + D \qquad (1)$$

The control system is assumed to be described by its own $l \times m$ TFM $K(s)$ in a familiar unity negative feedback system, although some generalisations are possible (see references). The simplified problem is to design $K(s)$ using acceptable components to ensure, amongst other things, acceptable stability, performance and robustness targets.

The area is unusual in that it relies on a small number of basic relationships from which to develop a useable design theory. They are summarised below for completeness:

Stability, Poles and Zeros: The ratio of the closed-loop characteristic polynomial $\rho_c(s)$ to the open-loop characteristic polynomial $\rho_o(s)$ is equal to the Return-difference Determinant

$$\rho_c(s)/\rho_o(s) = |I_m + G(s)K(s)| \qquad (2)$$

where

$$Q(s) = G(s)K(s) \qquad (3)$$

is the forward path TFM. The idea of poles moved easily to the multivariable case but zeros caused more of a problem. For design purposes, the relevant definition applies most usefully to the case when $m = l$ i.e. the zeros of a system with TFM $R(s)$ and characteristic polynomial $\rho(s)$ are the zeros of the zero polynomial

$$z(s) = \rho(s)|R(s)| \qquad (4)$$

Performance: The closed-loop TFM relating the plant output y to the demand signal r is computed as

$$H_c(s) = (I_m + Q(s))^{-1}Q(s) \qquad (5)$$

Interpreted element by element, this, in principle, provided complete data on the closed loop transient performance. Such an analysis is essential but the pre-occupation of the researchers in the period being discussed was on the reduction of input-output interaction as represented by the off-diagonal terms of $H_c(s)$. Stability and Robustness: The relevant TFM here was typically the Sensitivity TFM relating the tracking error $e = r - y$ to the signal r:

$$S(s) = (I_m + Q(s))^{-1} \qquad (6)$$

3. DIAGONAL DOMINANCE AND THE BEGINNINGS OF A THEORY

There is no doubt that the issue of interaction was seen as a barrier to real progress in multivariable design and consequently pre-occupied the minds of researchers. The mathematical reason for this is, in simple terms, the non-linear dependence of the stability relationship and the closed-loop TFM on the off-diagonal terms of $G(s)$. In a mathematical sense, if $m = l$, the removal of interaction simply relies on the construction of an inverse of the plant model i.e. the choice of the controller TFM

$$K(s) = G^{-1}(s)D(s) \qquad (7)$$

where $D(s)$ is diagonal ensures that interaction is removed from both the forward path and closed-loop systems. The complexities of the inverse and the perceived priorities to achieve a simple controller (for ease of construction and commissioning/retuning) made this ”solution” unacceptable to designers at the time.

The pioneering idea that exact removal of all interaction was unnecessary and unachievable in the presence of plant uncertainty and modelling errors is easy to state but non-trivial to prove.

The first pioneering result that provided a clear and rigorously defined situation where interaction is sufficiently small as to permit the design of the m control loops independently (ignoring the interaction!) was provided by (Rosenbrock, 1969). The importance of his contribution was that the criterion was based on classical frequency domain data and hence fitted well into the classical way of thinking. This meant that it was potentially easily understood by classically trained control engineers and hence more easily transferred into industrial use - benefits that were difficult to ignore!

The mathematical trick used by Rosenbrock was to demand that suitable TFMs have the property of being diagonally dominant.

Definition 1. An $m \times m$ matrix M is diagonally (row) dominant if, for each index i,

$$|M_{ji}| > \sum_{j \neq i} = r_i \tag{8}$$

Note 1. M is diagonally column dominant if M^T is row dominant.

Note 2. In both cases, diagonal matrices are diagonally dominant and diagonally dominant matrices are, in a precise sense, nearly diagonal.

The nice thing about this well-known result is that it permits a graphical interpretation i.e. the requirement that, for each index i, the Gershgorin circle of radius r_i and centre M_{ii} in the complex plane does not contain the origin.

The idea can be used in a number of ways (see e.g. (Owens, 1978*a*)) that will be illustrated in the presentation. The essence of the kind of result obtained can be summarised as verifications that the process of applying familiar Nyquist or inverse Nyquist stability criteria to diagonal elements of $Q(s)$ and/or its inverse will successfully predict closed-loop stability. This is achieved provided that uncertainty introduced by ignoring the off-diagonal (interaction) terms is removed by demanding that the diagonal term's Nyquist or inverse Nyquist loci with super-imposed Gershgorin circles creates a band in the complex plane that does not contain a specified fixed point (typically the origin or the classically familiar $(-1,0)$ point).

Rosenbrock's work, in effect was one of the first results on robust multivariable control where the nominal model is taken to be the model obtained by ignoring interaction. It had a powerful impact on thinking in research in multivariable control systems design because it demonstrated feasibility of a frequency domain approach and had a familiar ”feel” to the graphical designers.

The work lead to a multitude of papers on how to achieve the diagonal dominance condition using structured pre-compensator(s) and many applications (see for example (Patel and Munro, 1982)). It was very influential and the discerning reader will find echoes of the ideas in aspects of modern robust control and related areas. Applications were not restricted to linear systems as can be seen in (Cook, 1972).

4. LOOP ADDITION CONCEPTS

The search for methods that reduced to the analysis of simple single loop designs (where interaction was either suppressed or removed from the analysis) also lead to the practically motivated idea of closing the loops one-by-one - a technique that had possibly been used by practitioners for many years but which had little theoretical basis for design. The theoretical base was provided by (Mayne, 1973) (and also described in Owens (1978)) in the form of recursive Nyquist-like design processes. The methods had computational issues to address and optimistically assumed that the designer had an idea of the ”best” sequence of loop to close.

The ideas attracted great attention for its theoretical interest but was not, to my knowledge, used extensively in applications or further developed at the theoretical level.

5. EIGENVALUE BASED METHODS

Rosenbrock's work is an approximate eigenvalue method as Gershgorin's theorem, in its original form, states that the eigenvalues of a complex matrix M lie in the union of its Gershgorin circles. A diagonally (row or column) dominant matrix is hence non-singular.

Macfarlane pioneered the use of eigenvalues (as precise functions of frequency) within multivariable frequency domain design theory (MacFarlane (1979, 1980) in a series of papers with his students that identified a rigorous relationship between the eigenvalues of $Q(s)$ and closed-loop stability. His more heuristic earlier work was put on a rigorous foundation with (Postlethwaite and MacFarlane, 1979) and further developed as a computational design tool with others including Kouvaritakis and Edmunds (MacFarlane, 1980). The method was called the ”Characteristic Locus Design Method” (or CL Method) (see (Maciejowski, 1989); (MacFarlane, 1980);(Owens, 1978*a*) for various descriptions/interpretations)

The mathematical formulation of the stability criteria and ”encirclement conditions” required the use of complex variable theory on Riemann surfaces (where frequency dependent eigenvalues have well-defined functional characteristic to which complex variable theory can be applied). Computational aspects of the theory were subject to more uncertainty as design ultimately consisted of the systematic manipulation of the frequency dependence of eigenvalues to ensure stability, performance and reduced closed-loop interaction. This problem lead to a number of computational approaches, some of which are summarised in (MacFarlane, 1980).

It is difficult to do real justice to this major contribution in a few words. It certainly clarified the fundamental relationship between stability and eigenvalues and provided necessary and sufficient conditions for stability (in contrast to Rosenbrock's results which were clearly only sufficient). Many of Rosenbrock's results can be derived from this work using eigenvalue approximation theory. The computational aspects of design were subject to more uncertainty - eigenvalue manipulation is notoriously difficult and the need to manipulate frequency dependent eigenvalues was even more complex. Substantial progress was made however using the idea of commutative and approximately commutative controllers (MacFarlane 1980) where attempts were made to influence eigenvalues at frequency points and in the vicinity of those points. Interaction reduction was achieved typically at low and high frequencies by the choice of coefficient matrices of multivariable PI controllers. At other frequencies, interaction was typically reduced by the use of high loop gains.

The author was involved with multivariable control from a very special perspective that ultimately linked in closely to the ideas of Rosenbrock and MacFarlane. The intuition (with hindsight) is that systems with a "modal" structure may be particularly suited to analysis using eigenvalue-type methods. The author was working on the design of sector based control systems for xenon-induced spatial oscillations in thermal nuclear power reactors at AEE Winfrith, UK. The symmetrical placing of the sensors and actuators lead to symmetry properties of the TFM $G(s)$ which then lead to simple expressions for eigenvalues and symmetric control structures that retained this property and allowed systematic and exact control design without the problems that could arise in the use of the CL method. Details can be found in (Owens, 1973*b*) where the ideas also were easily extended to include analysis of failure situations.

The extension of this work was to the idea of dyadic transfer function matrices (Owens, 1973*a*) based on the idea that eigenvalues were not necessarily physical quantities of the system e.g. the choice of units for inputs and outputs influence the eigenvalues but do not change the underlying reality of the system dynamics! Whereas eigenvalues are obtained by similarity transformation of a TFM, the argument was that physical quantities would be accessible through equivalence transformations. In the context of the nuclear reactor application, this was expressed by noting that $G(s)$ has the dyadic TFM form

$$G(s) = PD(s)Q \tag{9}$$

where P and Q are constant nonsingular matrices and $D(s)$ is a diagonal TFM. For the nuclear reactor, $D(s)$ could be related to physical modes of the underlying partial differential equation model with P and Q reflecting mode shapes and the influences of modes on the output and input on modes respectively.

The constancy of P and Q made the method a special case where difficult issues arising in CL analysis were absent. In Owens (1974) (also reprinted in MacFarlane (1979) and described in Owens (1978)) the author extended the ideas to show that the same principles hold for an arbitrary square, invertible TFM. That is, a decomposition of the from of (9) can be derived at any selected frequency but where P and Q then depend on that frequency and $D(s)$ is only diagonal at that frequency. The resultant Method of Dyadic Expansion hence made it possible to exactly manipulate eigenvalues at any specified frequency using simple compensators. It also made it possible to manipulate (in an approximate sense) the behaviour of eigenvalues in the vicinity of that frequency using the diagonal terms of $D(s)$ as models and using Gershgorin circles to estimate the uncertainty.

6. WHY DIT IT CHANGE?

Eigenvalue methods were very much a UK phenomenon and caused great excitement. It is a matter of speculation and opinion as to why they were ultimately replaced by H-infinity methods (see Maciejowski, 1989). It can be argued that the computational complexities and uncertainties of achieving diagonal dominance conditions and/or shaping eigenvalues overwhelmed the community who then opted for the hands-off computational certainties of the optimisation-based H-infinity approach. No doubt much was achieved by this change. A personal view is that something may also have been lost e.g. how would the modal structures of dyadic TFMs have arisen from H-infinity; how would loop-by-loop methods be developed and more, generally, how is it possible to input physical information and insight to the process. No doubt many discussions in bars around the world will help to resolve the issue?

Much remains of the ideas in the form of the underlying methods of robust control (Skogestad and Postlethwaite, 1996) and issues of approximation (see (Owens and Chotai, 1986)) who consider the use of frequency domain design base on the use of experimental step response data) but the current accumulation of ideas and methods is now potentially much richer and computationally better understood. The open question is whether or not substantial further progress is possible!

7. MULTIVARIABLE ROOT LOCI

The development of a generalisation of the classical and well-known root-locus method in the mid seventies created several years of excitement with most results summarised in three texts, namely (MacFarlane 1980; Owens 1978b; Postlethwaite and MacFarlane, 1979) and a number of associated publications. The work yielded a deep understanding of both the complex variable theory of multivariable root loci and a clear and systematic method for controller design based of the manipulation of root locus asymptotes but was not used extensively for reasons that may boil down to those of mathematical complexity or simple a human response to the perception that the multivariable generalisation was unlikely to replicate the usefulness of the classical version.

The formulation of the problem was relatively simple but the subsequent analysis could be approached from several viewpoints. MacFarlane, Postlethwaite and co-workers undertook fundamental studies of the complex variable theory of root loci using the CL methodology of Riemann surfaces. This lead to a complete analytic function theory of the concept and the existence and parameterisation of series expansions for the closed loop poles about ”the point at infinity” i.e. infinite loop gain. The role and properties of zeros were retained in the multivariable case and, broadly speaking, the asymptotic analysis indicated that a multivariable root locus could be visualised as *m* superimposed root-loci of m single-input-single-output (SISO) systems. This is a crude picture but useful for visualisation and the interpretation of an $m \times m$ multivariable system as containing m separate SISO systems of different relative degree. These relative degrees could then be interpreted in terms of properties of oscillation and stability in a classical sense, although quite how these properties would appear in the system’s outputs could not be predicted without additional analysis.

MacFarlane and Postlethwaite’s work did however indicate that, in certain situations, the above picture fails - in the creation of what appeared to represent non-integer relative degrees. The author of this paper took a more computational viewpoint of the root locus question by providing algorithms for characterising the nature and parameters of the asymptotes through matrix computations on the first few terms in the Markov matrix sequence $\{CA^kB\}$ - the so-called method of dynamic transformation (see (Owens, 1978*a*)). This analysis confirmed the fact that the orders of the asymptotes generically took certain values (later proved to be related to the structural invariants of a group action on matrix triples (A,B,C,D) (Owens, 1978*b*)) and provided formulae for the asymptotic directions and pivots.

The computational approach was powerful in characterising and computing asymptotes but failed to address the form of the non-generic case with any ease. It did however provide methods for systematic manipulation (design) of parameters in the asymptotes (Owens, 1978b) and hence influencing the response speed and degree of oscillation (performance?) of the closed-loop system. The asymptotic nature of the analysis meant however that the methods provided little information on the position of roots at smaller gains and could have sensitivity problems that meant that high gain properties did not provide any guidance to lower gain behaviours. As this is likely to be the region of practical relevance, it was considered that multivariable root locus methods would have to be supported by other techniques such as the ones considered earlier in this paper.

8. FREQUENCY RESPONSE METHODS IN OTHER AREAS

The excitement that was associated with the above work spilled over into related areas.
(MacFarlane, 1979) provides a section on the application of the ideas to so-called-multi-dimensional systems characterised by TFMs dependent on two independent complex variables (representing dynamics in two orthogonal dimensions). These ideas arose from areas such as 2D digital signal processing. The author was involved with the early recognition (by Edwards and Owens) that some physical processes involving repetition where the output from one repetition influenced the output from the next are examples of 2D systems. Examples of application areas included long-wall coal cutting control, control of agricultural ploughing and metal rolling with more theoretical applications in areas of iterative algorithms and iterative learning control theory.

Some of these ideas were reported in the text by (Edwards and Owens, 1982) and have been substantially further developed by (Rogers and Owens, 1992). The basic stability result (Owens, 1976; Edwards and Owens, 1982; Rogers and Owens, 1992) was expressed as necessary and sufficient conditions for stability in terms of the Characteristic Loci (eigenvalues) of a carefully selected TFM. In this case no further improvement is possible as H-infinity methods are only sufficient in the 2D context.

9. SUMMARY AND CONCLUSIONS

The 1960s and 1970s development of extensions of classical frequency domain design methods to cope with multivariable systems saw a great change in the ability of the community to analyse and design multiloop control systems. The UK research community was a major player in these innovations based on the use of eigenvalue approximation in Rosenbrock’s inverse Nyquist array method, Mayne’s Sequential Return Difference Method, MacFarlane’s eigenvalue-based Characteristic Locus Method and the author’s contribution to CL theory in the form of the Method of Dyadic Expansion. The insight produced by these studies coupled with the technically more complex results for multivariable root loci identified solutions and unsolved problems, many of which still remain.

The move from these approaches to the more algorithmic use of H-infinity approaches can be explained in many ways. What is not in doubt is that H-infinity builds firmly on the foundations provided by this early work. Students and researchers new to the area may find a perusal through the references will provide the connections that they need to make further progress in what is still a vibrant area of relevant research.

10. A NOTE ON MODEL DESIGN

It has not been possible to be comprehensive within the limitations of this paper. The paper has focussed on the UK (where most of the developments began) but there was interest around the world. It has also focussed on the design process for a given model. The idea of dyadic transfer function matrices indicates what can be achieved if additional model structure is included in the theoretical analysis and development of the design process. Dyadic TFMs are, intuitively only one of a large number of modelling assumptions that could be made. In classical control, the canonical models chosen for simplified design and conceptualisation are the familiar first and second order models (formulae) omitted. The author introduced similar concepts for square $m \times m$ systems $G(s)$ (brought together in the text, Owens (1978b)) by defining first order MIMO lags as systems with inverse TFM of the form

$$G(s)^{-1} = sA_0 + A_1 \tag{10}$$

and second order MIMO systems via

$$G(s)^{-1} = s^2A_0 + sA_1 + A_2 \tag{11}$$

with A_0 non-singular. These definitions also cover the SISO case. Mixed first and second order systems are obtained for MIMO systems by allowing A_0 to be singular in (11)! The benefits of these assumptions are clear in design terms but have not been exploited fully in practice. My belief is that they offer an opportunity and have been sufficiently successful to merit practical and theoretical consideration of these and other MIMO structures to act as guides and benchmarks for industrial use, whatever the chosen design method!

11. REFERENCES

Cook, P. A. (1972). Modified multivariable circle theorems. In: *Proceedings of the IMA Conference on Recent Mathematical Developments in Control*. London. pp. 367–372.

Edwards, J. B. and D. H. Owens (1982). *Analysis and control of multipass processes*. Research Studies Press. Chichester.

MacFarlane, J. A. G. (1979). *Frequency response methods in control systems*. IEEE Press. New York.

MacFarlane, J. A. G. (1980). *Complex variable methods for linear multivariable feedback systems*. Taylor and Francis. London.

Maciejowski, J. M. (1989). *Multivariable feedback design*. Addison-Wesley. Wokingham.

Mayne, D. Q. (1973). The design of multivariable systems. *Automatica* **9**, 201–207.

Owens, D. H. (1973*a*). Dyadic approximation method for multivariable control systems analysis with a nuclear reactor application. *Proc. IEE* **120**, 801–809.

Owens, D. H. (1973*b*). Multivariable control systems design concepts in the failure analysis of nuclear reactor spatial control systems. *Proc. IEE* **120**, 119–125.

Owens, D. H. (1974). Dyadic expansion for the analysis of linear multivariable systems. *Proc. IEE* **121**, 713–716.

Owens, D. H. (1976). The stability of linear multipass processes. *Proc. IEE* **124**, 1079–1082.

Owens, D. H. (1978*a*). *Feedback and multivariable systems*. Peter Peregrinus Ltd.

Owens, D. H. (1978*b*). On structural invariants and the root-loci of linear multivariable systems. *Int J. Control* **28**, 187–196.

Owens, D. H. and A. Chotai (1986). Approximate models in multivariable process control an inverse nyquist array and robust tuning regulator interpretation. *Proc. IEE* **133**, 1–12.

Patel, R. V. and N. Munro (1982). *Multivariable system theory and design*. Chap. 9. Vol. 4 of *International Series on Systems and Control*. Pergamon Press. Oxford.

Postlethwaite, I. and A. G. J. MacFarlane (1979). *A complex variable approach to the analysis and design of linear multivariable feedback systems*. Vol. 12 of *Lecture Notes in Control and Information Sciences*. Springer-Verlag.

Rogers, E. and D. H. Owens (1992). *Stability analysis for linear repetitive processes*. Vol. 175 of *Lecture Notes in Control and Information Sciences*. Springer-Verlag.

Rosenbrock, H. H. (1969). Design of multivariable feedback systems using the inverse nyquist array. *Proc. IEE* **116**, 1929–1936.

Skogestad, S. and I. Postlethwaite (1997). *Multivariable feedback control*. Wiley.