ROBUST MULTIVARIABLE CONTROL USING H^∞ METHODS
Analysis, Design and Industrial Applications

Ian Postlethwaite *
Engineering Department
University of Leicester
Leicester LE1 7RH, UK

Sigurd Skogestad †
Chemical Engineering
University of Trondheim, NTH
N-7034 Trondheim, Norway

Abstract

The purpose of this paper is to introduce the reader to multivariable frequency domain methods including $\mathbf{H}^\infty$-design. These methods provide a direct generalization of the classical loop-shaping methods used for SISO systems. We also aim to provide a basic understanding of how robustness problems arise, and what analysis and design tools are available to identify and to avoid them.

As an introduction to the robustness problems in multivariable systems we discuss the control of a distillation column. Because of strong interactions in the plant, a decoupling control strategy is extremely sensitive to input gain uncertainty (caused by actuator uncertainty). These interactions are analyzed using singular value decomposition (SVD) and relative gain array (RGA) methods.

We then discuss possible sources of model uncertainty, and look at the traditional methods for obtaining robust designs, such as gain margin, phase margin and maximum peak criterions (M-circles). However, these measures are difficult to generalize to multivariable systems. In such cases a more detailed modelling of the uncertainty in terms of norm-bounded perturbations (Δ's) is used. The frequency-domain is particularly well suited for representing non-parametric (unstructured) uncertainty. To test for robust stability and performance in the presence of model uncertainty, the structured singular value, μ, provides a powerful tool.

The latter part of the paper is concerned with design and in particular multivariable loop shaping of singular values of appropriately specified transfer functions. One and two degrees of freedom controllers are considered, and the paper ends with a case study on advanced control of high performance helicopters.

*E-mail: ixp@le.ac.uk; phone: +44-533-522546; fax: +44-533-522619
†E-mail: skoge@kjemi.unit.no; phone: +47-7-594154; fax: +47-7-594080

Contents

1 Introduction

The main goal of this introduction is to answer the following question: Why use the frequency domain ($\mathbf{H}^\infty$-norm) for defining performance and describing uncertainty? We will also discuss the two main approaches to $\mathbf{H}^\infty$-design, namely the loop-shaping and the signal-oriented approaches.

We use $\|M\|_\infty$ to denote the $\mathbf{H}^\infty$-norm of a linear transfer function $M(s)$. For the scalar case, $\|M\|_\infty$ is simply equal to the peak magnitude $\sup_\omega |M(j\omega)|$, where *sup* denotes *the least upper bound*, which for all practical purposes is equal to the maximum value. For the multivariable case we, "sum up" the channels using the singular value and we have

$$\|M\|_\infty \stackrel{\text{def}}{=} \sup_\omega \bar{\sigma}(M(j\omega)) \tag{1}$$

1.1 The loop shaping approach

This approach to control system design could also be called the classical approach, the engineering approach, the frequency domain approach or the transfer function approach.

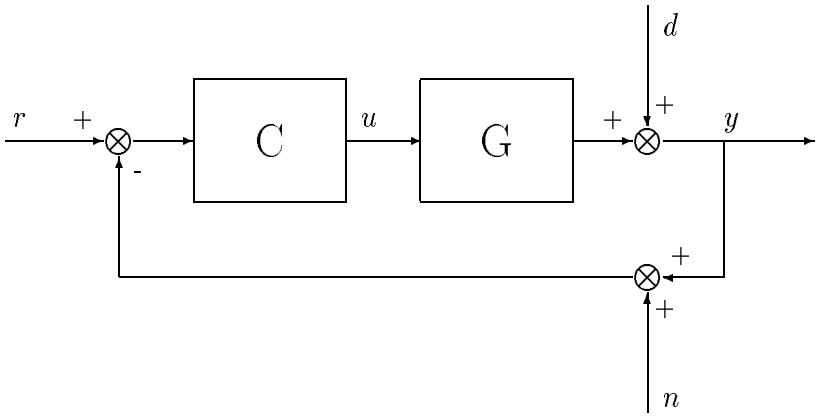

Figure 1.1: Conventional feedback control system.

Consider the conventional feedback system in Fig. 1.1 where $G(s)$ is the plant and $C(s)$ the feedback controller.

By "loop shaping" one traditionally means a design approach where one specifies directly the shape of the magnitude of the open-loop transfer function

$L = GC$. However, we shall use the term in a wider sense and also allow the specification of closed-loop transfer functions such as $S = (I + GC)^{-1}$. We use the following definition: *Loop-shaping is any design method that involves directly specifying the magnitude of one or more nominal transfer functions.* To distinguish between various approaches we will talk about "shaping L" or "shaping S", and so on.

Shaping L. For single-input single-output (SISO) systems the specifications in terms of the *open-loop* transfer function $L = GC$ typically include:

i) Crossover frequency, ω_c (defined as $|L(j\omega_c)| = 1$).

ii) System type (defined as number of integrators in $L(s)$).

iii) The shape of $L(j\omega)$, e.g., in terms of the slope of $|L(j\omega)|$ in certain frequency ranges.

iv) Phase margin, PM (given by the phase of $L(j\omega_c)$).

v) Gain margin, GM (given by the gain of L at the frequency where its phase is -180^o).

The first three specifications have to do with performance in terms of speed of response and allowable tracking error. The last two specifications are included to avoid some of the potential difficulties with feedback: 1) the closed-loop system may become unstable, 2) noise and disturbances in a certain frequency range close to the bandwidth frequency may get amplified.

Specifications directly on $L = GC$ make the design procedure simple as it is clear how changes in the controller affect $L(s)$, and this approach is well suited for non-formalized design procedure. Indeed, towards the end of this paper we shall discuss the MacFarlane-Glover loop-shaping procedure where the initial step in the controller design is to select a reasonable $L(s)$.

Shaping closed-loop transfer functions. In may cases one prefers to define specifications in terms of a *closed-loop* transfer function for the following reasons: 1) The final performance we want to evaluate is that of the closed-loop system. 2) The robustness specifications in terms of GM and PM are difficult to generalize to MIMO systems. 3) Synthesis is difficult if the specifications are in terms of L (one may have to resort to numerical procedures such as the Method of Inequalities (MOI) described towards the end of the paper).

Shaping S. The closed-loop sensitivity function, $S = (I+GC)^{-1}$, is a very good indicator of performance. Typical specifications in terms of S include:

i) Minimum bandwidth frequency ω_B (defined as the smallest frequency at which $\bar{\sigma}(S(j\omega)) = 0.707$)

ii) Allowable tracking error at selected frequencies.

iii) System type, or if system contains no integrators, the allowed static tracking error, A.

iv) The shape of S over selected frequency ranges.

v) Maximum allowed peak magnitude for S, $\|S(j\omega)\|_\infty = M_s$.

The peak specification prevents amplification of noise at high frequencies, and also introduces a margin of robustness; typically we select $M_s = 2$. Mathematically, these specifictions may be captured simply by an upper bound, $1/|w_P|$[1] on the magnitude of S, namely

$$\bar{\sigma}(S(j\omega)) \leq 1/|w_p(j\omega)|, \forall\omega \quad \Leftrightarrow \quad \|w_P S\|_\infty \leq 1 \tag{2}$$

A typical upper bound is shown in Fig. 1.2. The weight illustrated on that plot may be represented as

$$w_P(s) = \frac{s/M_s + \omega_B}{s + \omega_B A} \tag{3}$$

and we see that $|w_P(j\omega)|^{-1}$ is equal to $A \leq 1$ at low frequencies, is equal to $M_s \geq 1$ at high frequencies, and the asymptote crosses 1 at the bandwidth frequency, ω_B.

Figure 1.2: Upper bound on $\bar{\sigma}(S)$ given by weight $1/|w_P(j\omega)|$ in (3).

[1] Subscript P stands for performance since S is mainly used as a performance indicator

The loop shape $L = \omega_B/s$ yields an S which exactly matches the bound (3) at frequencies just below the bandwidth and easily satisfies the bound at other frequences. This L has a slope ("roll-off") in the frequency range below the bandwidth of about -1 on a log-log plot (-20 dB/decade). In many cases, in order to improve performance, we may want a steeper slope for L (and S) in some frequency range below the bandwidth, and a higher-order weight may be selected.

Stacked requirements. The specification $\|w_P S\|_\infty \leq 1$ does not allow us to specify an *upper* bound on the bandwidth or the "roll-off" of $L(s)$ above the bandwidth. To specify this one needs to make a specification on another transfer function, for example, on the complementary sensitivity $T = I - S = GCS$. Also, one may want to bound other transfer functions, to achieve robustness or to avoid too large input signals.

As an example, one may define an upper bound, $1/|w_T|$ on the magnitude of T to make sure the system behaves "nicely" at high frequencies, and an upper bound, $1/|w_u|$, on the magnitude of CS to avoid large input signals. To combine these specifications, the "stacking approach" is usually used, resulting in the following specification:

$$\|M\|_\infty = \sup_\omega \bar{\sigma}(M(j\omega)) \leq 1; \quad M = \begin{pmatrix} w_P S \\ w_T GCS \\ w_u CS \end{pmatrix} \tag{4}$$

The "mixed-sensitivity" specification with $M = \begin{pmatrix} w_P S \\ w_T GCS \end{pmatrix}$ is used, for example, by Chiang and Safonov (1988, 1992) in the Matlab Robust Control Toolbox Manual.

The "stacking procedure" is selected for mathematical convenience as it does not allow us to exactly specify the bounds on the individual transfer functions as was described above. For example, assume that $\phi_1(C)$ and $\phi_2(C)$ are two real scalars (here we could have $\phi_1(C) = \|w_P S\|_\infty$ and $\phi_2(C) = \|w_T GCS\|_\infty$) and that we want to achieve

$$\phi_1 \leq 1 \quad \text{and} \quad \phi_2 \leq 1 \tag{5}$$

This is not quite the same as the "stacked" requirement

$$\bar{\sigma} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \sqrt{\phi_1^2 + \phi_2^2} \leq 1 \tag{6}$$

The two requirements are quite similar when either ϕ_1 or ϕ_2 is small, but in the worst case when ϕ_1 and ϕ_2 are equal, we get from (6) that $\phi_1 \leq 0.707$ and $\phi_2 \leq 0.707$, that is, there is a possible "error" equal to a factor $\sqrt{2} \approx 3$ dB. In

general, with n stacked requirements the resulting error is at most $\sqrt{n}$. This inaccuracy in the specifications is something we are probably willing to sacrifice in the interests of mathematical convenience. In any case, the specifications are in general rather rough, and are effectively knobs for the engineer to select and adjust until a satisfactory design is reached.

The $\mathbf{H}^\infty$-optimal controller is obtained by solving the problem

$$\min_C \|M(C)\|_\infty \tag{7}$$

Provided M can be written as a linear fractional transformation (LFT) of C, $M(C) = N_{11} + N_{21}C(I - N_{22}C)^{-1}N_{12}$, the solution is easily obtained with standard software (e.g. the Robust Control or Mu Toolboxes in Matlab). In practice, to be able to write M as an LFT of C, one must be able to represent M by a block diagram with the input (or output) at only one location. For example, the M in (4) may be represented in a block diagram with a single input entering at the output of the plant as shown in Fig.1.3.

Figure 1.3: Block digram corresponding stacked requirement in (4).

Let $\gamma_0 = \min_C \|M(C)\|_\infty$ denote the optimal $\mathbf{H}^\infty$-norm for the problem in (7). An important property of $\mathbf{H}^\infty$-optimal controllers is that it will yield a flat frequency response, that is, it will yield $\bar{\sigma}(M(j\omega)) = \gamma_0$ at all frequencies. *The practical implication is that, except for at most a factor $\sqrt{n}$, the transfer functions resulting from solving (7) will be very close to γ_0 times the bounds selected by the designer.* This means, that the designer may almost exactly specify the final shape of, for example, $\bar{\sigma}(S)$, $\bar{\sigma}(GCS)$ or $\bar{\sigma}(CS)$.

Remark. For cases where M cannot be written as an LFT of C, which is a special case of the Hadamard-weighted $\mathbf{H}^\infty$-problem studied by van Diggelen and Glover (1991, 1992), the solution to the $\mathbf{H}^\infty$-problem in (7) remains intractable. Van Diggelen and Glover (1991,1992) do, however, present a solution for a similar problem where the Frobenius norm is used instead of the singular value to "sum up the channels" in the $\mathbf{H}^\infty$-norm.

Summary. The classical loop-shaping approach to controller design involves direct specifications of how the final solution should be in terms of the magnitude frequency response. It requires the engineer to be able to formulate bounds that lead to acceptable robustness and closed-loop performance. This approach is often preferred because it has few adjustable design parameters (knobs) and directly involves the engineer in the design. We shall return with a more detailed discussion on the physical significance of some transfer functions which the engineer may want to bound in Section 5.1.

1.2 The signal-oriented approach

The signal-oriented approach is very general, and may be more appropriate for multivariable problems in which a number of objectives must be taken into account simultaneously. Here we define the plant, including possibly the model uncertainty, define the class of external signals affecting the system and define the norm of the "error signals" we want to keep small. Direct bounds on selected transfer functions, such as the closed-loop bandwidth, cannot be specified in this case. On the other hand, one may argue that the concept of bandwidth is difficult to use for complex systems.

The "modern" state space methods from the 60's, such as LQG control, are based on the signal-oriented approach. Here the input signals are assumed to be stochastic (or alternatively impulses in a deterministic setting) and the output signals are measured in terms of the 2-norm. These methods may be generalized to include frequency dependent weights on the signals leading to what is called the Wiener-Hopf or H_2-norm design method.

Sinusoidal signals and the $\mathbf{H}^\infty$-norm. We may also consider the system response to persistent sinusoidal signals of varying frequency. This leads to the signal- oriented $\mathbf{H}^\infty$-norm approach used, for example, by Doyle et al. (1987) in their space-shuttle application. A signal-oriented problem specification with disturbances, commands and noise, and with bounds on both the input and outputs is shown in Fig. 1.4. The overall performance objective is that $\|E\|_\infty \leq 1$. For more details the reader is referred to Lundström et al. (1991).

The $\mathbf{H}^\infty$-norm may be interpreted in other ways, such as the induced 2-norm from inputs to outputs. In any event, as far as signals are concerned

Figure 1.4: Typical block diagram for signal-oriented approach to $\mathbf{H}^\infty$-performance.

there does not seem to be an overwhelming case for using the $\mathbf{H}^\infty$-norm rather than the more traditional H_2-norm (LQG). When we begin to consider issues of model uncertainty, however, the frequency domain approaches such as $\mathbf{H}^\infty$ are preferred.

Model uncertainty. The traditional method of dealing with robustness and model uncertainty within the framework of "optimal control" (LQG) has been to introduce uncertain signals (noise and disturbances). One particular approach is loop transfer recovery (LTR) where unrealistic noise is added specifically to obtain a robust design. Of course, one may say that model uncertainty generates some sort of disturbance, but this disturbance depends on the other signals in the system and thus introduces an element of feedback. Therefore there is a fundamental difference between these two sources of "uncertainty" (at least for linear systems): model uncertainty may introduce instability, whereas signal uncertainty can not.

A more direct way to handle robustness issues within the signal approach, is to model the uncertainty explicitly. It appears that the frequency domain is very well suited for describing model uncertainty, and in particular for describing non-parametric uncertainty, resulting, for example, by using a simplified low-order model of a high-order plant. Indeed, Owen and Zames (1991) make the following observation in a recent paper: *"The design of feedback controllers in the presence of nonparametric and unstructured uncertainty ... is the raison*

d'être for $\mathbf{H}^\infty$ *feedback optimization, for if disturbances and plant models are clearly parameterized* $\mathbf{H}^\infty$*-methods seem to offer no clear advantages over more conventional state-space and parametric methods."*

If the $\mathbf{H}^\infty$-norm is selected for the uncertainty, then, again for mathematical convenience, one may also want to select the $\mathbf{H}^\infty$-norm for performance. This leads to a robust performance (RP) problem, for which there exists a very efficient analysis tool, namely the structured singular value, μ. But, for controller synthesis there are difficulties: the μ-synthesis problem, in its full generality, is not yet solved mathematically; where solutions exist the controllers tend to be high order; the available algorithms may not always converge and design problems are sometimes difficult to specify directly.

1.3 Combined approaches

The loop shaping approach and the signal approach above may, of course, be combined. One such approach is "loop-shaping with uncertainty" as used, for example, by Skogestad et al (1988). Here performance is specified in terms of an upper bound on the sensitivity S, leading to the specification $\|w_P S\|_\infty \leq 1$. One then takes a "worst-case" approach and requires that this bound is satisfied for all plants as defined by the uncertainty description. For the case of input uncertainty of relative magnitude w_I this leads to the following robust performance analysis condition (see Eq.48)

$$RP \quad \text{iff} \quad \mu(N(j\omega)) < 1, \ \forall\omega; \quad N = \begin{pmatrix} w_I CSG & w_I CS \\ w_P SG & w_P S \end{pmatrix} \tag{8}$$

where μ is the structured singular value computed with respect to a special block-diagonal structure. This is similar to the $\mathbf{H}^\infty$-condition in (4), but in (8) the bounds on the transfer functions are specified indirectly, and it is not clear what nominal transfer functions CSG, CS, SG and S are allowed. Specifically, recall that the $\mathbf{H}^\infty$-optimal controller would essentially result in a controller which matches all the bounds at all frequencies (except for at most a factor $\sqrt{n}$). On the other hand, the μ-optimal design, may result in a design where, for example, $\bar{\sigma}(w_P S)$ is very small at some frequency and large at another frequency. It is therefore not clear from the specifications what the final (nominal) design will be.

However, the μ-approach does have definite advantages since we do know that the worst-case sensitivity function S_p (p stands for perturbed) will exactly satisfy our requirements, i.e., for all possible perturbed plants $\|w_P S_p\|_\infty \leq 1$. Whereas for the $\mathbf{H}^\infty$-problem we only specify nominal transfer functions and must make sure by specifying these carefully that robustness is achieved. When

applied to design, the approach (8) has the usual problems associated with μ-synthesis: computational difficulties, high-order controllers, and the indirect specification of individual transfer functions . From the above discussion then, it follows that μ-analysis may be very useful for checking the robustness of designs obtained, for example, by an $\mathbf{H}^\infty$ design procedure.

Another "combined approach" is the the "Glover-McFarlane loop-shaping procedure". In this, one first specifies the desired *open-loop* shape $L(s)$ for performance using simple pre- and post- compensators. One then "robustifies" this design by considering a particular robust stability condition, which involves solving an $\mathbf{H}^\infty$-problem. This procedure is further described in Section 5 and used in the design example in Section 6.

1.4 Summary

We have considered two alternative approaches to controller design: the loop-shaping approach and the signal approach. In both cases we find the frequency-domain to be the natural setting. The loop-shaping approach, with direct specifications on bandwidth etc, is directly based in the frequency domain so here there is no alternative. For the signal-oriented approach there are a variety of ways to define the signals. The reason why the frequency-domain ($\mathbf{H}^\infty$-norm) is again preferred is that it is very well suited to handling unstructured model uncertainty.

In a practical design situation, the above two approaches may be combined. For example, one may design the controller (Step B below) by some loop-shaping approach (involving $\mathbf{H}^\infty$-synthesis) and then analyze the solution (Step C below) using a signal-oriented approach with model uncertainty explicitly included (involving μ-analysis).

This paper is concerned with analysis and design of control systems for industrial plants. In this case the designer must usually go through the following steps:

Step A. Controllability analysis: This is where the plant is analysed and we discover what closed-loop performance can be expected, what the limitations are, how good the control might be.

Step B. Controller design: This is where the design problem is formulated and the controller synthesized.

Step C. Control system analysis: This is where the feedback control system is assessed by analysis and simulation to judge how well it might behave in practice.

With the above steps in mind, the following topics are covered in the remainder of this paper.

Section 2: Analysis of the plant - controllability

Section 3: Robustness problems - Introductory distillation example

Section 4: Tools for robustness analysis

Section 5: Robust control system design

Section 6: A case study.

1.5 Notation

G - nominal plant model
M - matrix used to test for robust stability (section 1-4) or coprime factor of G (section 5)
RGA - matrix of relative gains, $= G \times (G^{-1})^T$ where $\times$ represents element-by-element multiplication.
s - Laplace variable ($s = j\omega$ yields the frequency response)
$S = (I + GC)^{-1}$ - sensitivity function
$T = GC(I + GC)^{-1}$ - closed-loop transfer function
$T_I = CG(I + CG)^{-1}$ - closed-loop transfer function from the plant input
w and W_i - frequency-dependent weighting functions

Greek letters
Δ - overall perturbation block used to represent uncertainty
Δ_I - perturbation block for input uncertainty
$\gamma(A) = \bar{\sigma}(A)/\underline{\sigma}(A)$ - condition number of matrix A
$\mu(A)$ - structured singular value of matrix A
ω - frequency [rad/s or rad/min]
$\bar{\sigma}(A)$ - maximum singular value of matrix A
$\underline{\sigma}(A)$ - minimum singular value of matrix A

Subscripts
p - perturbed (with model uncertainty)
P - performance

2 Analysis of the plant - Controllability

Before attempting to start any controller design one should have some idea of how easy the plant actually is to control. Is it a difficult control problem? Indeed, does there even exist a controller which meets the required performance objectives? It appears that the frequency-domain is very well suited for answering such problems in a general setting. One reason for this is the very useful idea of "bandwidth" which is a purely frequency-domain concept. The concept of right half plane (RHP) zeros is also of fundamental importance in answering questions of the kind.

In this paper the term "controllability" (of a plant) has the meaning of "inherent control characteristics of the plant" or maybe better "achievable performance" (irrespective of the controller). This usage is in agreement with most persons intuitive feeling about the term, and was also how the term was used historically in the control literature. For example, Ziegler and Nichols (1943) define controllability as *"the ability of the process to achieve and maintain the desired equilibrium value"*.

Unfortunately, in the 60's Kalman defined the term "controllability" in the very narrow meaning of "state controllability". This concept is of interest for realizations and numerical calculations, but as long as we know that all the unstable modes are both controllable and observable, it has almost no practical significance.

It would be desirable to have a more precise definition of controllability, but on the other hand this is difficult and probably not useful. An exact definition would require selection of a certain norm to measure the control error, and would also require a detailed specification of all external signals such as noise, reference signals and disturbances. Indeed, Ziegler and Nichols (1943) note in their paper that although they "took the area under a recovery curve as one measure of controllability ... this is only one of many possible bases for comparison of control results". They also stress that it is difficult to narrow controllability down to one single attribute of the plant. They say: "Unfortunately, the authors are not able to give a formula for controllability. It appears that when such a factor is devised it will consist of several factors. One might be called the "recovery factor", the ability of the process to recover from the maximum change in demand or load. Another, a "load factor" must take into account the point in the process at which the disturbance occurs". Later in the paper they state that the total integrated control error, $\int |e(t)|dt$, is equal to: (Load Factor) $\cdot$ (Recovery Factor).

Essentially, the "recovery factor" depends on the process model, $g(s)$, and recovery is poor (and thus the recovery factor is large) if it contains large time

delays or if the plant gain is small. The "load factor" expresses the effect of the disturbances and thus depends on the disturbance model, $g_d(s)$. These concepts are very similar to the ideas summarized below in terms of upper and lower bandwidth limitations.

2.1 Summary of controllability results for SISO plants

Consider the control system in Fig. 1.1 for the case when all blocks are scalar. The control error $e = y - r$ may be written

$$e(s) = g(s)u(s) + g_d(s)d(s) - r(s) \tag{9}$$

We assume that the signals are persistent sinusoids, and assume that g and g_d are scaled, such that at each frequency the allowed input $|u(j\omega)| < 1$, the expected disturbance $|d(j\omega)| < 1$, and the outputs are scaled such that the expected reference signal $|r(j\omega)| < 1$.

Below we have given some "controllability" requirements which apply to the closed-loop bandwidth, ω_B. The requirements are fundamental, although the expressions given in terms of bounds on ω_B are not exact. However, in practice they must be fulfilled.

i) *Disturbances. Must require $\omega_B > \omega_d$.* Here ω_d is the frequency at which $|g_d(j\omega_d)|$ crosses 1 from above. Below this frequency the error will be unacceptable ($|e| > 1$) for a disturbance $d = 1$ unless control is used. More specifically, we must for feedback control require at frequencies lower than ω_d: $|gc(j\omega)| > |g_d(j\omega)|$.

ii) *Commands (setpoints). Specify directly minimum required ω_B.* This requirement comes in addition to the bandwidth requirement imposed by the disturbances, and is usually relatively easy to specify.

iii) *Open-loop unstable pole at $s = p$. Must require $\omega_B > 1/|p|$. We need fast control to stabilize the system, and the bandwidth must approximately be greater than $1/|p|$ where $|p|$ is the distance of the RHP-pole from the origin.*

iv) *Input constraints, must require $|g(j\omega)| > 1, \forall \omega < \omega_B$.* This is needed to avoid input constraints ($|u(j\omega)| < 1$) for perfect tracking of $r(j\omega) = 1$.

v) *Input constraints, must require $|g(j\omega)| > |g_d(j\omega)|$, $\forall \omega < \omega_d$.* This is needed to avoid input constraints for perfect rejection of disturbance $d(j\omega) = 1$.

The above two conditions are requirements that the plant must satisfy in order to be able to apply tight control in a certain frequency reange. They are independent of the controller, and can therefore be affected only by changing the plant $g(s)$.

In the frequency range up to the bandwidth ω_B there should not be any time delays, RHP-zeros and high-order plant dynamics that need to be counteracted. We get

vi) *Time delay θ. Must require $\omega_B < 1/\theta$.*

vii) *RHP-zero at $s = z$. Must require $\omega_B < |z|$.*

Note that RHP-zeros close to the origin are the worst. LHP-zeros pose no fundamental limitation, but a LHP-zero close to the origin yields an "overshoot" in the open-loop response which may be difficult to counteract. Therefore, to simplify controller design and avoid robustness problems, it is often best to have the LHP-zeros as far away from the origin as possible.

The above two constraints for time delays and RHP-zeros are fundamental, but the above relationships are rather approximate. Also, if there are combinations of both RHP-zeros and time delays then they must be considered combined, because they all make feedback control difficult (simply consider the overall phase lag).

viii) *In most practical cases: $\omega_B < \omega_{180}$.*

Here ω_{180} is the frequency at which the phase of $g(j\omega)$ is -180^o. This condition is not a fundamental limitation, but more of a practical limitation. In particular it applies if the phase drops rather quickly around the frequency ω_{180}. The condition follows since in most cases the plant is not known sufficiently accurately to place zeros to counteract the poles at high frequency.

2.2 Controllability analysis for multivariable plants

We do not have space to go into detail about the controllability analysis of multivariable plants, but most of the ideas presented above may be generalized, e.g., see Wolff et al. (1992). Instead we will summarize some of the main tools which are used. All of them are based on the plant model $G(s)$ and the disturbance model $G_d(s)$.

i) Compute the multivariable RHP-poles and RHP-zeros and their associated directions. Test for functional controllability (the rank of G should equal the number of outputs).

ii) Perform a frequency-dependent SVD-analysis to understand the multivariable directions.

iii) Perform a frequency-dependent RGA-analysis to check for fundamental limitations due to inherently coupled outputs. Compute the plant condition number.

iv) Evaluate disturbance sensitivity. For decentralized control the use of the CLDG-matrix, $G_{diag}G^{-1}G_d$, directly generalizes the SISO results. Here G_{diag} is a diagonal matrix consisting of the diagonal elements of G. For the general case it is more complicated, but an SVD-analysis of G_d and $G^{-1}G_d$ yields useful information about which disturbances are difficult, and the bandwidth requirement in certain directions.

The above tools for controllability analysis are simple indicators which are easy to compute, and help the engineer to obtain insight into what the control problems are for the plant in question. In some cases a more detailed analysis which includes finding the optimal controller may be desirable. A suitable tool is the structured singular value μ (which must be minimized over all controllers to find the achievable performance for the problem). However, the use of such methods requires a careful definition of the performance specifications and model uncertainty which is often not available or which requires a significant effort to obtain.

Although, there has been good progress during the last few years, the area of controllability analysis is still a very interesting area for future research.

3 Robustness - Introductory distillation column example

An idealized distillation column example will be used to introduce the reader to the adverse effects of model uncertainty, in particular for multivariable plants. The example is taken mainly from Skogestad et al. (1988)

Before considering the example a short introduction to robustness and uncertainty seems in order.

3.1 Robustness and model uncertainty

A control system is robust if it is insensitive to differences between the actual system and the model of the system which was used to design the controller. Robustness problems are usually attributed to differences between the plant model and the actual plant (usually called model/plant mismatch or simply model uncertainty). Uncertainty in the plant model may have several origins:

i) There are always parameters in the linear model which are only known approximately or are simply in error.

ii) Measurement devices have imperfections. This may even give rise to uncertainty on the manipulated inputs, since the actual input is often measured and adjusted in a cascade manner. For example, this is often the case with valves where we measure the flow. In other cases limited valve resolution may cause input uncertainty.

iii) At high frequencies even the structure and the model order is unknown, and the uncertainty will exceed 100% at some frequency.

iv) The parameters in the linear model may vary due to nonlinearities or changes in the operating conditions.

v) In addition, the controller implemented may differ from the one obtained by solving the synthesis problem, and one may include uncertainty to allow for controller order reduction and implementation inaccuracies.

Other considerations for robustness include measurement and actuator failures, constraints, changes in control objectives, opening or closing other loops, etc. Furthermore, if a control design is based on an optimization then robustness problems may also be caused by the mathematical objective function, that is, how well this function describes the real control problem.

In the somewhat narrow use of the term used in this paper, we shall consider robustness with respect to model uncertainty, and assume that a fixed (linear) controller is used. Intuitively, to be able to cope with large changes in the process, this controller has to be detuned away from the best response we might have achieved if the process model was exact.

To consider the effect of model uncertainty, the uncertainty needs first to be quantified in some way. There are several ways of doing this. One powerful method is the frequency domain (so-called H-infinity uncertainty description) in terms of norm-bounded perturbations (Δ's). With this approach one can also take into account unknown or neglected high-frequency dynamics.

The following terms are useful:

- Nominal stability (NS). The system is stable with no model uncertainty.
- Nominal Performance (NP). The system satisfies the performance specifications with no model uncertainty.
- Robust stability (RS). The system is stable for all perturbed plants about the nominal model up to the worst-case model uncertainty.
- Robust performance (RP). The system satisfies the performance specifications for all perturbed plants about the nominal model up to the worst-case model uncertainty.

3.2 The distillation column model

We consider two-point (dual) composition control of a distillation column. The overhead composition of a distillation column is to be controlled at $y_D = 0.99$ (output 1) and the bottom composition at $x_B = 0.01$ (output 2), with reflux L (input 1) and boilup V (input 2) as manipulated inputs for composition control, i.e.,

$$y = \begin{pmatrix} \Delta y_D \\ \Delta x_B \end{pmatrix}, \quad u = \begin{pmatrix} \Delta L \\ \Delta V \end{pmatrix}$$

By linearizing the steady-state model and assuming that the dynamics may be approximated by a first order response with time constant $\tau = 75$ min, we derive the following linear model in terms of deviation variables

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = G \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \quad G(s) = \frac{1}{\tau s + 1} \begin{pmatrix} 87.8 & -86.4 \\ 108.2 & -109.6 \end{pmatrix} \tag{10}$$

Here we have scaled the inputs and outputs to be less than 1 in magnitude (this corresponds to the outputs in 0.01 mole fraction units, and the inputs scaled relative to the feed rate). The gains are much larger than 1 indicating no problems with input constraints, but this is somewhat deceiving as the gain in the the low-gain direction (corresponding to the smallest singular value) is actually just above 1.

This is admittedly a very crude model of a distillation column. Specifically, a) the parameters may vary drastically with operating point, b) there should be a high-order lag in the transfer function from u_1 to y_2 to represent the liquid flow down to the column, and c) higher-order composition dynamics should also be included. However, the model is simple and displays important features of the distillation column behavior. The RGA-matrix for this model is at all frequencies

$$RGA(G) = \begin{pmatrix} 35.1 & -34.1 \\ -34.1 & 35.1 \end{pmatrix} \tag{11}$$

The large elements in this matrix indicate that this process is fundamentally difficult to control (see section 4.2).

3.2.1 Interactions and ill-conditioned plants

From (10) we get

$$y_1(s) = \frac{87.8}{75s+1} u_1(s)$$

Thus an increase in u_1 by only 0.01 (with u_2 constant) yields a steady- state change in y_1 of 0.878, that is, the outputs are very sensitive to changes in u_1. Similarly, an increase in u_2 by only 0.01 (with u_1 constant) yields $y_1 = -0.864$. Again, this is a very large change, but in the opposite direction of that for the increase in u_1.

We therefore see that changes in u_1 and u_2 counteract each other, and if we increase u_1 and u_2 simultaneously by 0.01, then the overall steady- state change in y_1 is only $0.878 - 0.864 = 0.014$. Physically, the reason for this small change is that the compositions in the column are only weakly dependent on changes in the *internal flows* (i.e., simultaneous changes in the internal flows L and V).

Summary: Since both u_1 and u_2 affect both outputs, y_1 and y_2, we say that the process is *interactive*. This is quantified by relatively large off-diagonal elements in $G(s)$. Furthermore, the process is *ill-conditioned*, that is, some combinations of u_1 and u_2 have a strong effect on the outputs, whereas other combinations of u_1 and u_2 (corresponding to $u_1 \approx u_2$) have a weak effect on the outputs. This is quantified by the condition number; the ratio between the gains in the strong and weak directions; which is large for this process (as seen below it is 141.7).

3.2.2 Singular Value Analysis of the Model

The above discussion shows that this distillation column is an ill- conditioned plant, where the effect (the gain) of the inputs on the outputs depends strongly on the *direction* of the inputs. To see this better, consider the SVD of the steady-state gain matrix

$$G = U\Sigma V^T \tag{12}$$

or equivalently since $V^T = V^{-1}$

$$G\bar{v} = \bar{\sigma}(G)\bar{u}, \quad G\underline{v} = \underline{\sigma}(G)\underline{u}$$

where

$$\Sigma = diag\{\bar{\sigma}, \underline{\sigma}\} = diag\{197.2, 1.39\}$$

$$V = (\bar{v}\ \underline{v}) = \begin{pmatrix} 0.707 & 0.708 \\ -0.708 & 0.707 \end{pmatrix}$$

$$U = (\bar{u}\ \underline{u}) = \begin{pmatrix} 0.625 & 0.781 \\ 0.781 & -0.625 \end{pmatrix}$$

The large plant gain, $\bar{\sigma}(G) = 197.2$, is obtained when the inputs are in the direction $\binom{u_1}{u_2} = \bar{v} = \binom{0.707}{-0.708}$. From the direction of the output vector $\bar{u} = \binom{0.625}{0.781}$, we see that these inputs cause the outputs to move in the same direction, that is, they mainly affect the average output $\frac{y_1+y_2}{2}$. The low plant gain, $\underline{\sigma}(G) = 1.39$, is obtained for inputs in the direction $\binom{u_1}{u_2} = \underline{v} = \binom{0.708}{0.707}$. From the output vector $\underline{u} = \binom{0.781}{-0.625}$ we see that the effect then is to move the outputs in different directions, that is, to change $y_1 - y_2$. Thus, it takes a large control action to move the compositions in different directions, that is, to make both products purer simultaneously. Indeed, we see that in this direction it may be possible that one could be limited by input constraints (corresponding to $|u| > 1$). The condition number of the plant, which is the ratio of the high and low plant gain, is

$$\gamma(G) = \bar{\sigma}(G)/\underline{\sigma}(G) = 141.7 \tag{13}$$

The RGA is another indicator of ill-conditionedness, which is generally better than the condition number, because it is scaling independent. The sum of the absolute value of the elements in the RGA (denoted $\|RGA\|_{sum} = \Sigma|RGA_{ij}|$) is approximately equal to the minimized (with respect to input and output scaling) condition number, $\gamma^*(G) = \min_{D_1,D_2} \gamma(D_1 G D_2)$ where D_1 and D_2 are real diagonal "sacling" matrices. In our case we have $\|RGA\|_{sum} = 138.275$ and $\gamma^*(G) = 138.268$. (We note that the minimized condition number is quite similar to the condition number in this case, but this does not hold in general.)

3.3 Control of the column

3.3.1 Decoupling control

For "tight control" of ill-conditioned plants the controller should compensate for the strong directions by applying large input signals in the directions where the plant gain is low, that is, a "decoupling" controller similar to G^{-1} in directionality is desired. However, because of uncertainty, the direction of the large inputs may not correspond exactly to the low plant-gain direction, and the amplification of these large input signals may be much larger than expected. As shown in the simulations below, this will result in large values of the controlled variables y, leading to poor performance or even instability. Consider the following decoupling controller (or equivalently a steady-state decoupler combined

with a PI controller):

$$C_1(s) = \frac{k_1}{s} G^{-1}(s) = \frac{k_1(1+75s)}{s} \begin{pmatrix} 0.39942 & -0.31487 \\ 0.39432 & -0.31997 \end{pmatrix}, \quad k_1 = 0.7\text{min}^{-1} \tag{14}$$

We have $GC = 0.7/sI$. In theory, this controller should counteract all the directions of the plant and give rise to two decoupled first-order responses with time constant $1/0.7 = 1.43$ min. This is indeed confirmed by the solid line in Fig.3.1 which shows the simulated response to a setpoint change in y_1. *We thus conclude that the decoupling controller satisfies the nominal performance (NP) requirement.*

Figure 3.1: Response for decoupling controller to a unit setpoint change in y_1 with time constant 5 min, i.e, $r_1 = 1/(5s+1)$. Solid line: Nominal response with no uncertainty. Dotted line: 20% gain uncertainty as defined by Equation 15.

3.3.2 Robustness of decoupling control

We also note that this simple design yields an infinite gain margin (GM) and a phase margin (PM) of 90^o in both channels. For multivariable systems such margins are however misleading as we shall see in the following.

To be specific consider the case with 20% error (uncertainty) in the gain in each input channel ("diagonal input uncertainty"):

$$u_1 = 1.2u_{1c}, \quad u_2 = 0.8u_{2c} \tag{15}$$

Note that this expression is in terms of deviation variables. Here u_1 and u_2 are the actual changes in the manipulated flow rates, while u_{1c} and u_{2c} are the desired changes (what we believe the inputs are) as specified by the controller. It is important to stress that this diagonal input uncertainty, which stems from our inability to know the exact values of the manipulated inputs, is *always* present. Note that the uncertainty is on the *change* in the inputs (flow rates), and not on their absolute values. A 20% error is reasonable for process control applications (some reduction may be possible, for example, by use of cascade control using flow measurements, but there will still be uncertainty because of errors in measurement sensitivity). Anyway, the main objective of this paper is to demonstrate the effect of uncertainty, and its exact magnitude is of less importance.

It is straightforward to see that the uncertainty in (15) does not by itself yield instability, thus *we have robust stability (RS) for the decoupling controller.* However, whereas for SISO systems we generally have that NP and RS imply robust performance (RP) this is often not the case for MIMO systems.

This is clearly shown from the dotted lines in Fig.3.1 which shows the response with the uncertainty in (15). It differs drastically from the nominal response represented by the solid line, and even though it is stable the response is clearly not acceptable; it is no longer decoupled, and $y_1(t)$ and $y_2(t)$ reach a value of about 2.5 before settling at their desired values of 1 and 0. *Thus RP is not satisfied for the decoupling controller.*

There is a simple reason for the observed poor response to the setpoint change in y_1. To accomplish this change, which occurs mostly in the "bad" direction corresponding to the low plant gains, the inverse-based controller generates large changes in u_1 and u_2, while trying to keep the $u_1 - u_2$ very small. However, uncertainty with respect to the actual values of u_1 and u_2 makes it impossible to make them both large while at the same time keeping their difference small – the result is an undesired large change in the actual value of $u_1 - u_2$, which subsequently results in large changes in y_1 and y_2 because of the large plant gain in this direction.

Remark. The system satisfied RS because the uncertainty only occurs at the input to the plant. In practice, with for example a small time delay added to one of the outputs, this controller would give an unstable response.

3.3.3 A robust controller: Single-loop PID

Unless special care is taken, most multivariable design methods (MPC, DMC, QDMC, LQG, LQG/LTR, DNA/INA, IMC, etc.) yield similar inverse-based controllers, and do not generally yield acceptable designs for ill-conditioned plants. This follows since they do not explicity take uncertainty into account, and the optimal solution is then to use a controller which tries to remove the interactions by inverting the plant model.

Figure 3.2: Response for PID controller.

The simplest way to make the closed-loop system insensitive to input uncertainty is to use a *simple* controller, for example two single-loop PID controllers, which does not try to make use of the details of the directions in the plant model. The problem with such a controller is that little or no correction is made for the strong interactions in the plant, and then even the nominal response (with no uncertainty) is relatively poor. This is shown in Fig.3.2 where we have used the following PID controllers (Lundström et al., 1991)

$$y_1 - u_1 : \quad K_c = 1.62; \tau_I = 41 \text{ min}; \tau_D = 0.38 \text{ min} \tag{16}$$

$$y_2 - u_2 : \quad K_c = -0.39; \tau_I = 0.83 \text{ min}; \tau_D = 0.29 \text{ min} \tag{17}$$

The controller tunings yield a relatively fast response for y_2, and a slower response for y_1. As seen from the dotted line in Fig.3.2 the response is not very

much changed by introducing the model error in Eq.15.

Figure 3.3: Response for μ-optimal controller.

In Fig.3.3 we show the response with the μ-optimal controller (see Lundström et al., 1991) which is designed to optimize the worst-case response (robust performance) as discussed towards the end of this paper. Although this is a multivariable controller, we note that the response is not too different from that with the simple PID controllers, although the response settles faster to the new steady-state.

3.3.4 Limitations with the example: Real columns

It should be stressed again that the column model used above is not representative of a real column. In a real column the liquid lag, θ_L, from the top to the bottom, makes the initial response less interactive and the column is easier to control than found above. It turns out that the important parameter to consider for controllability is *not* the RGA at steady-state (with exception of the sign), but rather the RGA at frequencies corresponding to the closed-loop bandwidth. For a model of a real distillation column the RGA is large at low frequencies (steady-state), but it drops at high frequencies and the RGA-matrix becomes close to the identity matrix at frequencies greater than $1/\theta_L$.

Thus, since the interactions are much less at high frequencies, control is

simple, even with single-loop PI or PID controllers, if we are able to achieve very tight control of the column. However, if there are significant measurement delays (these are typically 5 min or larger), then we are forced to operate at a low bandwidth, and the responses in Figs.3.1-3.3 are more representative. Furthermore, it holds in general that one should *not* use a steady-state decoupler if the steady-state RGA-elements are large (typically larger than 5).

4 Tools for robustness analysis

In this section we will first introduce some simple tools, such as the frequency-dependent RGA, to understand the poor responses observed in the distillation example in the last section. Then we consider more general methods, which allow for a detailed description of the model uncertainty. This leads into a discussion of the structured singular value, μ, as an analysis tool for evaluating whether a system satisfies robust stability (RS) and robust performance (RP). Readers who want to learn more about μ are referred to Doyle (1982), Doyle at al. (1982), Skogestad et al. (1987), or to the texts by Morari and Zafiriou (1989) and Maciejowski (1989).

4.1 Simple tools for robustness analysis

4.1.1 SISO systems

For single-input-single-output (SISO) systems one has traditionally used gain margin (GM) and phase margin (PM) to avoid problems with model uncertainty. Consider a system with open-loop transfer function $g(s)c(s)$, and let $gc(j\omega)$ denote the frequency response. The GM tells by what factor the loop gain $|gc(j\omega)|$ may be increased before the system becomes unstable. The GM is thus a direct safeguard against steady-state gain uncertainty (error). Typically we require GM > 1.5.

The phase margin tells how much negative phase we can add to $gc(s)$ before the system becomes unstable. The PM is a direct safeguard against time delay uncertainty: If the system has a crossover frequency equal to ω_c, (defined as $|gc(j\omega_c)| = 1$), then the system becomes unstable if we add a time delay of $\theta = PM/\omega_c$. For example, if PM $= 30^o$ and $\omega_c = 1$ rad/min, then the allowed time delay error is $\theta = (30/57.3)$[rad]$/1$[rad/min] $= 0.52$ min.

Maximum peak criterions. In practice, we do not have pure gain and phase errors. For example, in a distillation column the time constant will usually increase when the steady-state gain increases. A more general way to specify stability margins is to require the Nyquist locus of $gc(j\omega)$, to stay outside some

region of the -1 point (the "critical point") in the complex plane. Usually this is done by considering the maximum peak, M_t of the closed-loop transfer function T or the peak M_s of the sensitivity function. The reader may be familiar with M-circles drawn in the Nyquist plot or in the Nichols chart. Typically, we require that M_s and M_t are less than 2 (6 dB). $1/M_s$ is simply the minimum distance between $gc(j\omega)$ and the -1 point. In most cases the values of M_t and M_s are closely related, especially when the peak is large, There is a close relationship between M_t/M_s and PM and GM. Specifically, for a given M_s we are guaranteed

$$GM \geq \frac{M_s}{M_s - 1}; \qquad PM \geq 2\arcsin(\frac{1}{2M_s}) \geq \frac{1}{M_s}[\text{rad}] \tag{18}$$

For example, with $M_s = 2$ we have GM ≥ 2 and PM $\geq 29.0^o > 1/M_s$ [rad] $= 28.6^o$. Similarly, for a given value of M_t we are guaranteed GM $\geq 1 + \frac{1}{M_t}$ and PM $\geq 2\arcsin(\frac{1}{2M_t}) > \frac{1}{M_t}$.

4.1.2 MIMO systems

It is difficult to generalize GM and PM to MIMO systems. On the other hand, the maximum peak criterions may be generalized easily. The most common generalization is to replace the absolute value by the maximum singular value, for example, by considering

$$M_t = \max_{\omega} \bar{\sigma}(T(j\omega)); \quad T = GC(I + GC)^{-1} \tag{19}$$

Even though we may easily generalize the maximum peak criterion to multivariable systems, it is often not useful for the following three reasons:

1) In contrast to the SISO case, it may be not sufficient to look at only the transfer function T. Specifically, for SISO systems $GC = CG$, but this does not hold for MIMO systems. This means that although the peak of T (in terms of $\bar{\sigma}(T(j\omega))$) is low, the peak of $T_I = CG(I + CG)^{-1}$ may be large.

2) The singular value may be a poor generalization of the absolute value. There may be cases where the maximum peak criterion , eg. in terms of $\bar{\sigma}(T)$, is *not* satisfied, but in reality the system may be robustly stable. The reason is that the uncertainty generally has "structure", whereas the use of the singular value assumes unstructured uncertainty. As shown below one should rather use the structured singular value, i.e. $\mu(T)$.

3) In contrast to the SISO case, the response with model error may be poor (RP not satisfied), even though the stability margins are good (RS is satisfied) and the response without model error is good (NP satisfied). For example, recall the distillation example above where for the decoupling controller $GC(s) = CG(s) = 0.7/sI$, and the values of M_t and M_s are both 1. Yet, the response

with only 20% gain error in each input channel is extremely poor. To handle such effects in general one has to define the model uncertainty and compute the structured singular value for RP.

The conclusion of this section is that most of the tools developed for SISO systems, and also their direct generalizations such as the peak criterions, are not sufficient for MIMO systems.

4.2 The RGA as a simple tool to detect robustness problems

4.2.1 RGA and input uncertainty

We have seen that a decoupler performed very poorly for the distillation model. To understand this better consider the loop gain GC. The loop gain is an important quantity because it determines the feedback properties of the system. For example, the transfer function from setpoints, r, to control error, $e = y - r$, is given by $e = -Sr = -(I + GC)^{-1}r$. We therefore see that large changes in GC due to model uncertainty will lead to large changes in the feedback response. Consider the case with diagonal input uncertainty, Δ_I. Let Δ_1 and Δ_2 represent the relative uncertainty on the gain in each input channel. Then the actual ("perturbed") plant is

$$G_p(s) = G(s)(I + \Delta_I); \quad \Delta_I = \begin{pmatrix} \Delta_1 & 0 \\ 0 & \Delta_2 \end{pmatrix} \tag{20}$$

Note that Δ_i is *not* normalized to be less than 1 in this case. The perturbed loop gain with model uncertainty becomes

$$G_pC = G(I + \Delta_I)C = GC + G\Delta_I C \tag{21}$$

If a diagonal controller $C(s)$ (eg., two PI's) is used then we simply get (since Δ_I is also diagonal) $G_pC = GC(I + \Delta)$ and there is no particular sensitivity to this uncertainty. On the other hand, with a perfect decoupler (inverse- based controller) we have

$$C(s) = k(s)G^{-1}(s) \tag{22}$$

where $k(s)$ is a scalar transfer function, for example, $k(s) = 0.7/s$, and we have $GC = k(s)I$ where I is the identity matrix, and the perturbed loop gain becomes

$$G_pC = G(I + \Delta_I)C = k(s)(I + G\Delta_I G^{-1}) \tag{23}$$

For the distillation model (10) studied above the error term becomes

$$G\Delta_I(G)^{-1} = \begin{pmatrix} 35.1\Delta_1 - 34.1\Delta_2 & -27.7\Delta_1 + 27.7\Delta_2 \\ 43.2\Delta_1 - 43.2\Delta_2 & -34.1\Delta_1 + 35.1\Delta_2 \end{pmatrix} \tag{24}$$

This error term is worse (largest) when Δ_1 and Δ_2 have opposite signs. With $\Delta_1 = 0.2$ and $\Delta_2 = -0.2$ as used in the simulations (Eq.15) we find

$$G\Delta_I G^{-1} = \begin{pmatrix} 13.8 & -11.1 \\ 17.2 & -13.8 \end{pmatrix} \tag{25}$$

The elements in this matrix are much larger than one, and the observed poor response with uncertainty is not surprising.

The observant reader may have noted that the RGA-elements appear on the diagonal in the matrix $G\Delta_I G^{-1}$ in (24). This turns out to be true in general as diagonal elements of the error term prove to be a direct function of the RGA (Skogestad and Morari, 1987)

$$(G\Delta G^{-1})_{ii} = \Sigma_{j=1}^{n} \lambda_{ij}(G)\Delta_j \tag{26}$$

Thus, if the plant has large RGA elements and an inverse-based controller is used, the overall system will be extremely sensitive to input uncertainty.

Control implications. Consider a plant with large RGA-elements in the frequency-range corresponding to the closed-loop time constant. A diagonal controller (eg., single-loop PI's) is robust (insensitive) with respect to input uncertainty, but will be unable to compensate for the strong couplings (as expressed by the large RGA- elements) and will yield poor performance (even nominally). On the other hand, an inverse-based controller which corrects for the interactions may yield excellent nominal performance, but will be very sensitive to input uncertainty and will not yield robust performance. In summary, plants with large RGA-elements around the crossover-frequency are fundamentally difficult to control, and decouplers or other inverse-based controllers should never be used for such plants (The rule is never to use a *decoupling controller* for a plant with large RGA-elements). However, one-way decouplers may work satisfactorily.

4.2.2 RGA and element uncertainty/identification

Above we introduced the RGA as a sensitivity measure with respect to input gain uncertainty. In fact, the RGA is an even better sensitivity measure with respect to element-by-element uncertainty in the matrix.

Consider any complex matrix G and let λ_{ij} denote the ij'th element in it's RGA-matrix. The following result holds (Hovd and Skogestad, 1992):

The (complex) matrix G becomes singular if we make a relative change $-1/\lambda_{ij}$ in its ij-th element, that is, if a single element in G is perturbed from g_{ij} to $g_{pij} = g_{ij}(1 - \frac{1}{\lambda_{ij}})$.

Thus, the RGA-matrix is a direct measure of sensitivity to element-by-element uncertainty and matrices with large RGA-values become singular for small relative errors in the elements.

Example. The matrix G in (10) is non-singular. The 1,2-element of the RGA is $\lambda_{12}(G) = -34.1$. Thus the matrix G becomes singular if $g_{12} = -86.4$ is perturbed to $g_{p12} = -86.4(1 - 1/(-34.1)) = -88.9$.

The result above is primarily an important algebraic property of the RGA, but it also has some important control implications:

1) Consider a plant with transfer matrix $G(s)$. If the relative uncertainty in an element at a given frequency is larger than $|1/\lambda_{ij}(j\omega)|$ then the plant may be singular at this frequency. This is of course detrimental for control performance. However, the assumption of element-by-element uncertainty is often poor from a physical point of view because the elements are usually always *coupled* in some way. In particular, this is the case for distillation columns: We know that the elements are coupled such that the model will not become singular due to small individual changes in the elements. The importance of the result above as a "proof" of why large RGA-elements imply control problems is therefore not as obvious as it may first seem.

2) However, for process identification the result is definitely useful: Models of multivariable plants, $G(s)$, are often obtained by identifying one element at the time, for example, by using step or impulse responses. From the result above it is clear this method will most likely give meaningless results (eg., the wrong sign of the steady-state RGA) if there are large RGA- elements within the bandwidth where the model is intended to be used. Consequently, identification must be combined with first principles modelling if a good multivariable model is desired in such cases.

Example. Assume the true plant model is

$$G = \begin{pmatrix} 87.8 & -86.4 \\ 108.2 & -109.6 \end{pmatrix}$$

By extremely careful identification we obtain the following model:

$$G_p = \begin{pmatrix} 87 & -88 \\ 109 & -108 \end{pmatrix}$$

This model seems to be very good, but is actually useless for control purposes since the RGA-elements have the wrong sign (the 1,1-element in the RGA is -47.9 instead of $+35.1$). A controller with integral action based on G_p would yield an unstable system.

To learn more about the RGA the reader is referred to Hovd and Skogestad (1992) where additional references can be found.

4.3 Advanced tools for robustness analysis: μ

So far in this paper we have pointed out the special robustness problems encountered for MIMO plants, and we have used the RGA as our main tool to detect these robustness problems. We found that plants with *large* RGA- elements are 1) fundamentally difficult to control because of sensitivity to input gain uncertainty, and decouplers should not be used, and 2) very difficult to identify because of sensitivity to element-by-element uncertainty.

We have not yet addressed the problem of analyzing the robustness of a given system with plant $G(s)$ and controller $C(s)$. In the beginning of this section we mentioned that the peak criterions in terms of M were useful for robustness analysis for SISO systems both in terms of stability (RS) and performance (RP). However, for MIMO systems things are not as simple. We shall first consider uncertainty descriptions and robust stability and then move on to performance. The calculations and plots in the remainder of this paper refer to the simple distillation model (10), using as a controller a steady-state decoupler plus PI-control.

4.3.1 Uncertainty modelling

Before considering how to analyze uncertain systems, we will consider the $\mathbf{H}^\infty$-approach to modelling plant uncertainty.

Linear Fractional Transformations (LFT) provide a general framework for modelling uncertainty (Doyle, 1984). A LFT may be written in the following form (see Fig. 4.1)

$$z = F_u(P, \Delta)w = (P_{22} + P_{21}\Delta(I - P_{11}\Delta)^{-1}P_{12})w \tag{27}$$

Figure 4.1: Uncertainty represented as linear fractional transformation (LFT).

Here P_{22} is the nominal mapping from w to z and Δ is a $\mathbf{H}^\infty$-norm bounded

perturbation,

$$\|\Delta\|_\infty = \sup_\omega \bar{\sigma}(\Delta(j\omega)) \leq 1 \tag{28}$$

Several sources of uncertainty may be combined and then $\Delta = diag\{\Delta_i\}$ is a block-diagonal matrix with perturbation blocks Δ_1, Δ_2 etc. These blocks may represent parametric uncertainty, in which case they are scalars Δ_i , possibly repeated, or they may represent unstructured uncertainty in which case they may be matrix-valued.

Each of these perturbations is bounded in terms of its $\mathbf{H}^\infty$-norm. For parametric uncertainty this is actually not very convenient as it would allow for complex variations in the parameter, $|\Delta_i| \leq 1$. Therefore for parametric uncertainty we generally restrict Δ_i to be real. Thus, it clear that the frequency domain does not offer any advantage for parametric uncertainty. On the other hand, the frequency bounds come in nicely when handling non-parametric uncertainty such as neglected dynamics. Also, it is very convenient for lumping several sources of uncertainty, although this must be done with some care to avoid being too conservative (when the uncertainty description allows unrealistic plants).

For unstructured uncertainty we have to make a choice of where to place the perturbation representing the uncertainty in question. Some alternatives are shown in Fig.4.2. These may all be represented by the LFT in Eq. (27).

There is no definite rule on which unstructured uncertainty to use, but the following may be useful: 1) Use the multiplicative (relative) uncertainty to represent neglected and uncertain dynamics occuring *between* the plant and the controller (e.g., neglected or uncertain actuator and measurement dynamics). 2) Use the "feedforward" (additive) forms when the zero uncertainty is large (in particular if a zero may go from the LHP to RHP) 3) Use the "feedback" forms when the pole uncertainty is large (in particular if a pole may cross the $j\omega$-axis). One particular combination of the feedforward and feedback forms, which appears to be useful, is the coprime uncertainty used in the Glover-McFarlane loop shaping procedure described in the next section.

However, care must be taken when representing uncertainty in an unstructured form. For example, for our distillation column example, it may be tempting to add some unstructured additive uncertainty to the plant. It turns out that this uncertainty description would be extremely conservative for this plant as the sign of the plant (represented by the sign of $\det G(s)$ or by the signs in the RGA-matrix) is extremely sensitive to such changes. In practice, as noted earlier, this kind of uncertainty does not occur for distillation columns as there are strong couplings between the elements in $G(s)$.

Two examples illustrate the usefulness of the general uncertainty description

Figure 4.2: Alternative ways of representing unstructured uncertainty. (a) Additive uncertainty, (b) Multiplicative input uncertainty, (c) Multiplicative output uncertainty, (d) Inverse additive uncertainty, (e) Inverse multiplicative input uncertainty, (f) Inverse multiplicative output uncertainty.

given above.

Neglected dynamics. Assume that the real set of plants is something like

$$\text{Real plant}: \quad g' = k'e^{-\theta s}, \quad k' \in [0.8k, 1.2k] \tag{29}$$

where k is the nominal ("average") gain, and we allow for gain variations of $\pm 20\%$. To simplify the controller design we want to use a simple nominal model with no delay, i.e.,

$$\text{Nominal model}: \quad g = k \tag{30}$$

The uncertainty in the gain may be handled directly as parametric uncertainty, but the neglected delay must clearly be represented in a non-parametric manner. In order to simplify the uncertainty description we choose to lump together gain variations and the neglected delay as unstructured multiplicative (relative) uncertainty:

$$\text{Set of possible models}: g_p(s) = k(1 + w_I\Delta_I); \quad |\Delta_I(j\omega)| \leq 1, \quad \forall\omega \tag{31}$$

Here Δ_I is a *complex* scalar. The modelled set g_p must include the real set of plants $g'(s)$. Let r_k represent the relative uncertainty in the gain. Then the following approximation for the weight is derived using a first-order Pade-approximation

$$(1+r_k)e^{-\theta s} - 1 \approx (1+r_k)\frac{1-\frac{\theta}{2}s}{1+\frac{\theta}{2}s} - 1 \tag{32}$$

Since it is only the magnitude that matters we make this expression minimum phase and derive the following simple weight

$$w_I(s) = r_k \frac{1+(\frac{1}{r_k}+\frac{1}{2})\theta s}{1+\frac{1}{2}\theta s} \tag{33}$$

The weight is somewhat optimistic (too small) at intermediate frequencies. In our case with $r_k = 0.2$ the magnitude of the weight is $r_k = 0.2$ at low frequencies, crosses 1 at about frequency $1/\theta$ and approaches $2(1 + r_k/2) = 2.2$ at high frequencies.

Note that even though the uncertainty weight only has 1 state it will allow for an infinite number of plants of arbitrary high order. On the other hand, (31) is *not* an exact representation of the original set of plants $g'(s)$ and may be conservative for that reason. For a scalar case it is probably not very conservative as the delay is generally the "worst case". However, in the multivariable case this may not always be true.

Pole variations represented as parametric uncertainty. Consider the set of plants $g' = 1/(s+a')$ where $-1 \leq a' \leq 3$. This may be exactly represented as

$$g_p = \frac{1}{s+a+2\Delta}; \quad , a = 1,\ |\Delta| \leq 1 \tag{34}$$

where Δ is a *real* scalar perturbation. This is in fact an inverse additive uncertainty (see Fig.4.2) with nominal model $g(s) = 1/(s+a)$ and $w_{ia} = 2$. Note also that poles crossing from the left to the right half plane may be modelled tightly with this uncertainty.

4.3.2 Conditions for Robust stability

By Robust Stability (RS) we mean that the system is stable for all possible plants as defined by the uncertainty set (using the Δ's as discussed above). This is a "worst case" approach, and for this reason one must be careful about not including unrealistic or impossible parameter variations. With this caution in mind, it turns out that the $\mathbf{H}^\infty$-norm (for completely unstructured uncertainty) and the structured singular value (for diagonally structured uncertainty) provide an exact way of analyzing robust stability,

Figure 4.3: Multiplicative input uncertainty.

As an example, consider the the case with multiplicative input uncertainty shown in Fig. 4.3. We assume that the system without uncertainty ($\Delta = 0$) is stable (we have NS). Instability may then only be caused by the "new" feedback paths caused by the Δ-block. Therefore, to test for robust stability (RS) we rearrange the feedback system with uncertainty into the standard form in Fig.4.4 with the two blocks Δ and M. Here the *interconnection matrix* M is the transfer function from the output,u_Δ, to the input, y_Δ, of the Δ-block. For the case of multiplicative input uncertainty we have $\Delta = \Delta_I$ and obtain $M = wC(I + GC)^{-1}G = w_I T_I = w_I CSG$ (the negative sign has been dropped as it does not matter). To test for stability we make use of the "small gain theorem". Since the Δ-block is normalized to be less than 1 at all frequencies, this theorem says that the system is stable if the M-block is less than 1 at all frequencies. Robust stability is then satisfied if

$$\bar{\sigma}(M) = \bar{\sigma}(w_I T_I(j\omega)) < 1, \quad \forall\omega \tag{35}$$

Figure 4.4: General block diagram for studying robust stability and robust performance.

Unstructured uncertainty. One crucial point is that this condition is also necessary (it is clearly sufficient) for RS provided we allow for *all* Δ's satisfying

$\bar{\sigma}(\Delta) \leq 1, \forall\omega$. That is, we have for the general block diagram in Fig.4.4:

$$RS \ \forall \ \|\Delta\|_\infty \leq 1 \quad \text{iff} \quad \|M\|_\infty < 1 \tag{36}$$

The same robust stability condition applies for each of the six forms of unstructured uncertainty shown in Fig. 4.2 when we use

$$M_a = W_A CS, \quad M_b = W_I CSG, \quad M_c = W_O GCS \tag{37}$$

$$M_d = W_{iA} SG, \quad M_e = W_{iI}(I + CG)^{-1}, \quad M_f = W_{iO} S \tag{38}$$

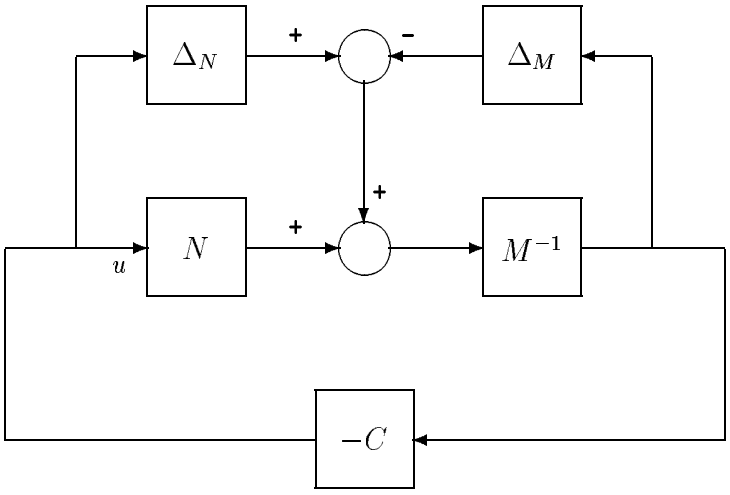

Figure 4.5: Coprime uncertainty description

However, even though (36) is mathematically correct it will generally be conservative for the following two reasons: 1) It allows for Δ to be complex, 2) It allows for Δ to be a full matrix.

It is actually the second point which is the main problem in most cases. However, before discussing it we shall introduce the coprime uncertainty description which will be used in the next section.

Coprime uncertainty description. Consider the uncertainty description in Fig.4.5 (note that the M in that figure denotes one coprime factor of the plant and not the interconnection matrix). This uncertainty description is rather general, as it allows for both zeros and poles crossing into the right half plane,

and has proved to be very useful in applications. To test for RS we rearrange the block diagram to match Fig.4.4 with

$$\Delta = \begin{pmatrix} \Delta_N \\ \Delta_M \end{pmatrix}; \quad M_{RS} = \begin{pmatrix} C \\ I \end{pmatrix} (I + GC)^{-1} M^{-1} \tag{39}$$

where M_{RS} is the interconnection matrix. We get

$$RS \ \forall \left\| \begin{array}{c} \Delta_N \\ \Delta_M \end{array} \right\|_\infty \leq 1 \quad \text{iff} \quad \|M_{RS}\|_\infty < 1 \tag{40}$$

The reason why we get a tight condition in terms of the $\mathbf{H}^\infty$-norm even though we have two uncertainty blocks is that the blocks enter into the same point in the block diagram.

Structured uncertainty. We will now consider the general case where (36) does not provide a tight bound because we have several Δ-blocks caused by individual sources of uncertainty.

For example, if the input uncertainty represents neglected dynamics in the the individual channels then the set of possible plants is given by

$$G_p(s) = G(I + w_I \Delta_I); \quad \Delta_I = \begin{pmatrix} \Delta_1 & 0 \\ 0 & \Delta_2 \end{pmatrix} \tag{41}$$

where Δ_i represents the independent uncertainty in each input channel such that the overall Δ_I is a diagonal matrix (it has "structure"). ((41) is identical to Eq.(20), except that w_I yields the magnitude, since Δ_i is now normalized to be less than 1.)

Also, for multivariable plants it makes a difference whether the uncertainty is at the input or the output of the plant. Thus, we may want to consider combined input and output uncertainty. This may be represented in the general form in Fig.4.4 with M as 2×2 block matrix and $\Delta = diag\{\Delta_I, \Delta_O\}$. Again, we note that Δ has a diagonal structure and (36) is conservative.

To improve the tightness of condition (36) we first note that the issue of stability should be independent of scaling. We then have the improved condition

$$RS \quad \text{if} \quad \min_{D(\omega)} \bar{\sigma}(DMD^{-1}) < 1, \forall \omega \tag{42}$$

where D is a real block-diagonal scaling matrix with structure corresponding to that of Δ, such that $\Delta D = D\Delta$. A further refinement of this idea led to the introduction of the structured singular value, $\mu(M)$ (Doyle, 1982). We have (essentially this is the definition of μ)

$$RS \ \forall \text{ structured } \ \Delta \quad \text{iff} \quad \mu_\Delta(M) < 1 \tag{43}$$

This is a tight condition provided the uncertainty description is tight. Note that for computing μ we have to specify the block-*structure* of Δ and also if Δ is real or complex. Today there exists very good software for computing μ when Δ is complex. The most common method is to approximate μ by a "scaled" singular value as introduced in (42):

$$\mu_{\Delta}(M) \leq \min_{D} \bar{\sigma}(DMD^{-1}) \tag{44}$$

This upper bound is exact when Δ has three or fewer "blocks", and the largest deviation found so far for more blocks is 10-15% (Doyle, 1982).

Figure 4.6: μ-plots for distillation example with decoupling controller.

Distillation example revisited. Consider the distillation example from the previous section and consider multiplicative input uncertainty in each of the two input channels

$$w_I(s) = 0.2 + \frac{0.9s}{0.5s+1} = 0.2\frac{5s+1}{0.5s+1} \tag{45}$$

With reference to (33) we see that this corresponds to 20% gain error and a neglected time delay of about 0.9 min. The weight levels off at 2 (200% uncertainty) at high frequency. The dotted line in Fig.4.6 shows $\mu(M) = \mu(w_I T_I)$

for RS with this uncertainty using the decoupling controller. The μ-plot for RS shows the inverse of the margin we have with respect to our stability requirement. For example, the peak value of $\mu_{\Delta_I}(M)$ as a function of frequency is about 0.53. This means that we may increase the uncertainty by a factor $1/\mu = 1.89$ before the worst-case model yields instability. This means that we tolerate about 38% gain uncertainty and a time delay of about 1.7 min before we get instability.

Remark: For the decoupling controller we have $GC = \frac{0.7}{s}I$, and $T_I = T = \frac{1}{1.43s+1}I$. For this particular case it turns out that the structure of Δ does not matter, and we get $\mu_{\Delta}(M) = \bar{\sigma}(w_I T_I) = |0.2\frac{5s+1}{(0.5s+1)(1.43s+1)}|$. However, in other cases it may be critical to use the right structure, e.g., see Fig. 16 in Skogestad et al. (1988).

4.3.3 Conditions for Robust Performance

An additional bonus of using the $\mathbf{H}^\infty$-norm both for uncertainty and performance is that the robust performance (RP) problem may be recast as a special case of the RS-problem (Doyle et al, 1982) with the performance specification represented as a fake uncertainty block, Δ_P. To test for RP one then considers the interconnection matrix M_{RP} from the the outputs to the inputs of *all* the Δ-blocks, including the Δ_P-block for performance. Note that Δ_P is a "full" matrix (no diagonal structure). This follows since performance is defined using the singular value and we have $\bar{\sigma}(A) = \mu_{\Delta}(A)$ when Δ is a full matrix. M_{RP} depends on the plant G, the controller C and the weights used to define uncertainty and performance. The condition for robust performance within the $\mathbf{H}^\infty$-framework then becomes

$$RP \quad \text{iff} \quad \mu_{\tilde{\Delta}}(M_{RP}) < 1; \quad \forall\omega, \quad \tilde{\Delta} = \begin{pmatrix} \Delta & 0 \\ 0 & \Delta_P \end{pmatrix} \tag{46}$$

Distillation example revisited. Let us now check if RP is satisfied for the distillation example. To do this performance must first be defined.

Nominal performance. NP is defined such that at each frequency the value of the weighted sensitivity, $\bar{\sigma}(w_P S)$, should be less than 1. We select the weight

$$w_P(s) = \frac{s/2 + 0.05}{s} \tag{47}$$

With reference to Eq.(3) we see that this requires integral action, a closed-loop bandwidth of about 0.05 [rad/min] (which of course is relatively slow when the allowed time delay is only about 0.9 min) and a maximum peak for $\bar{\sigma}(S)$ of $M_s = 2$.

As discussed above we may define μ for NP as $\mu_{\Delta_P}(w_P S) = \bar{\sigma}(w_P S)$ where Δ_P is a full matrix. As expected, we see from the dashed line in fig.4.6 that the NP-condition is easily satisfied with the decoupling controller: $\bar{\sigma}(w_P S)$ is small at low frequencies and approaches $1/M_s = 0.5$ at high frequency because of the maximum peak requirement on $\bar{\sigma}(S)$.

Robust Performance. RP means that the performance specification is satisfied for the worst-case uncertainty. The most efficient way to test for RP is to compute μ for RP. If this μ-value is less than 1 at all frequencies then the performance objective is satisfied for the *worst case*. Although our system has good robustness margins and excellent nominal performance we know from the simulations in Fig.3.1 that the performance with uncertainty (RP) may be extremely poor. This is indeed confirmed by the μ-curve for RP in Fig.4.6 which has a peak value of about 6. This means that even with 6 times less uncertainty, the performance will be about 6 times poorer than what we require. μ for robust performance was computed as $\mu_{\tilde{\Delta}}(M_{RP})$ where the matrix $\tilde{\Delta}$ in this case has a block-diagonal structure with Δ_I (the true uncertainty) and Δ_P (the fake uncertainty stemming from the performance specification) along the main diagonal, and

$$M_{RP} = \begin{pmatrix} w_I CSG & w_I CS \\ w_P SG & w_P S \end{pmatrix} \tag{48}$$

The derivation of M_{RP} follows by representing the performance as an inverse multiplicative perturbation similar to that in Fig.4.2d, and rearranging the block to match Fig.4.4 (see Skogestad et al, 1988).

The μ-optimal controller is the controller which minimizes μ for RP. The present approach to designing the μ-optimal controller ('D-K iteration') is a rather tedious procedure which involves solving a number of scaled $\mathbf{H}^\infty$-problems. The iterations are not guaranteed to converge and generally result in high-order controllers.

For our example Lundström et al. (1991) obtained a μ-optimal controller with 22 states which yields an essentially flat μ-curve with a "peak" of μ of 0.978. The simulation in Fig.3.3 shows that the response even with this controller is relatively poor (taken into account that the only obvious limitation is a delay of about 1 min). The reason is that the combined effect of large interactions (as seen from the large RGA-values) and input uncertainty makes this plant fundamentally difficult to control.

Comment: In the time domain our RP-problem specification may be formulated *approximately* as follows: Let the plant be

$$G_p(s) = G(s) \begin{pmatrix} k_1 e^{-\theta_1 s} & 0 \\ 0 & k_2 e^{-\theta_2 s} \end{pmatrix} \tag{49}$$

where $G(s)$ is given in (10). Let $0.8 \leq k_1 \leq 1.2$, $0.8 \leq k_2 \leq 1.2$, $0 \leq \theta_1 \leq 0.9$ [min], and $0 \leq \theta_2 \leq 0.9$ [min]. The response to a step change in setpoint should have a closed-loop time constant less than about 20 minutes. Specifically, the error of each output to a unit setpoint change should be less than 0.37 after 20 minutes, less than 0.13 after 40 minutes, and less than 0.02 after 80 minutes, and with no large overshoot or oscillations in the response.

Conclusion. The structured singular value, μ, provides an excellent tool for analyzing the robustness of control systems. Within the $\mathbf{H}^\infty$-framework it is possible to consider most sources of model uncertainty, including parametric and unstructured uncertainty, and with help of μ one can essentially directly pick out the worst-case plant and see if it satisfies the specifications for RS or RP. However, for a number of reasons μ seems to be best suited for analysis, i.e, to answer "what if" questions. It may also be suited for evaluating the upper bound on achievable performance, i.e., as a kind of ultimate controllability tool. However, for actual controller design it seems like simpler methods, as the ones described in the next section, are more appropriate.

5 Robust Control System Design

In this section, we will focus on a loop shaping methodology for the design of robust multivariable control systems.

The classical loop shaping approach to control system design has been applied to industrial systems over several decades. For single-input single-output systems and loosely coupled systems, the approach has worked well. But for truly multivariable systems it has only been in the last decade that a reliable generalization of the approach has emerged. Multivariable loop shaping is based on the idea that a satisfactory definition of gain (range of gain) for a matrix transfer function is given by the singular values of the transfer function. By multivariable loop shaping, therefore, we mean the shaping of singular values of appropriately specified transfer functions.

5.1 Trade-offs in multivariable feedback design

In February 1981, the IEEE Transactions on Automatic Control published a Special Issue on Linear Multivariable Control Systems, the first six papers of which were on the use of singular values in the analysis and design of multivariable feedback systems. The paper by Doyle and Stein (1981) was particularly influential: it was primarily concerned with the fundamental question of how to achieve the benefits of feedback in the presence of unstructured uncertainty, and through the use of singular values it showed how the classical loop shaping

ideas of feedback design could be generalized to multivariable systems. To see how this was done, consider the one degree of freedom configuration shown in figure 5.1.

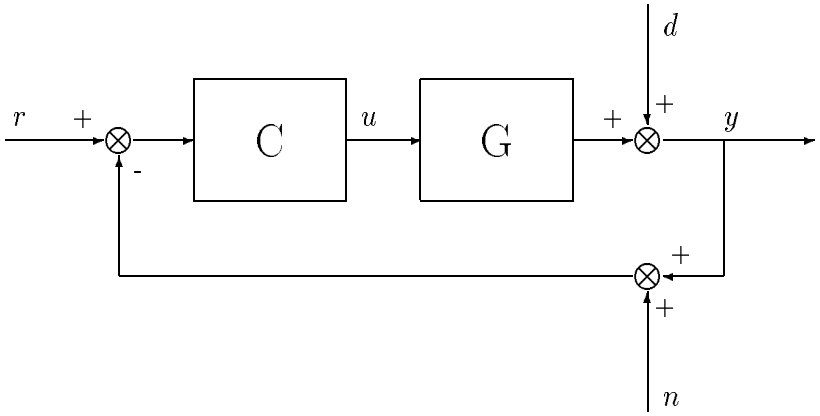

Figure 5.1: One degree of freedom feedback configuration

The plant G and controller C interconnection is driven by reference commands r, output disturbances d, and measurement noise n. y are the outputs to be controlled, and u are the control signals. In terms of the sensitivity function $S = (I + GC)^{-1}$ and the closed-loop transfer function $T = GC(I + GC)^{-1} = I - S$, we have the following important relationships:

$$y(s) = T(s)r(s) + S(s)d(s) - T(s)n(s) \tag{50}$$

$$u(s) = C(s)S(s)[r(s) - n(s) - d(s)] \tag{51}$$

These relationships determine several closed-loop objectives, in addition to the requirement that C stabilizes G; namely:

1. For *disturbance rejection* make $\bar{\sigma}(S)$ small.
2. For *noise attenuation* make $\bar{\sigma}(T)$ small.
3. For *reference tracking* make $\bar{\sigma}(T) \cong \underline{\sigma}(T) \cong 1$.
4. For *control energy reduction* make $\bar{\sigma}(CS)$ small.

If the unstructured uncertainty in the plane model G is represented by an additive perturbation i.e. $G_p = G + \Delta$, then a further closed-loop objective is (recall (37)):

5. For *robust stability* make $\bar{\sigma}(CS)$ small.

Alternatively, if the uncertainty is modelled by a multiplicative output perturbation such that $G_p = (I + \Delta)G$, then we have:

6. For *robust stability* make $\bar{\sigma}(T)$ small.

The closed-loop requirements 1 to 6 cannot all be satisfied simultaneously. Feedback design is therefore a trade-off over frequency of conflicting objectives. This is not always as difficult as it sounds because the frequency ranges over which the objectives are important can be quite different. For example, disturbance rejection is typically a low frequency requirement while noise mitigation is often only relevant at higher frequencies.

In classical loop-shaping, it is the magnitude of the open-loop transfer function GC which is shaped, whereas the above design requirements are all in terms of closed-loop transfer functions. However, it is relatively easy to convert the closed-loop requirements into the following open-loop objectives:

1. For *disturbance rejection* make $\underline{\sigma}(GC)$ large.

2. For *noise attenuation* make $\bar{\sigma}(GC)$ small.

3. For *reference tracking* make $\underline{\sigma}(GC)$ large.

4. For *control energy reduction* make $\bar{\sigma}(C)$ small.

5 & 6. For *robust stability* make $\bar{\sigma}(C)$ small.

Typically, requirements 1 and 3 are important at low frequencies, while 2, 4, 5 and 6 are high frequency conditions as illustrated in Figure 5.2.

To shape the gains (singular values) of GC by selecting C is a relatively easy task but to do this in a way which also guarantees closed-loop stability is in general non-trivial. Doyle and Stein (1981) suggested that an LQG controller could be used in which the regulator is designed using a "sensitivity recovery" procedure of Kwakernaak (1969) to give desirable properties (gain and phase margins) in GC. They also gave a dual "robustness recovery" procedure for designing the filter in an LQG controller to give desirable properties in CG. Recall that CG is not in general equal to GC which implies that stability margins vary from one break point to another in a multivariable system. Both these loop transfer recovery procedures had problems:

Figure 5.2: Design tradeoffs for the multivariable loop transfer function GC

- they were unsuitable for directly achieving specified loop shapes
- the gauranteed stability margins were only gauranteed as limiting properties in the design
- in the limit the controllers effectively inverted the plant and so the procedure broke down for nonminimum phase systems.

It was not until 1990, that a satisfactory loop shaping design procedure was developed by McFarlane and Glover (1990). This will be described in section 5.3, but first it will be necessary to consider a related robust stablization problem.

5.2 Robust Stabilization

As previously discussed in this paper, gain and phase margins are unreliable indicators of robust stability for multivariable systems because they do not take account of the coupling between loops. In section 4, several uncertainty descriptions were presented in which the uncertainty was captured by a norm bounded perturbation. Robustness levels could then be quantified in terms of the maximum singular values of various closed-loop transfer functions.

For example, in the feedback configuration of figure 5.1, if G is replaced by $G_p = G + \Delta$, where $\bar{\sigma}[\Delta(jw)] < \varepsilon(w)$, then the closed-loop remains stable if

$\bar{\sigma}[C(jw)S(jw)] < \varepsilon^{-1}(w)$ for all w. A design objective, for robust stabilization, might therefore be to find a C which stabilizes G and minimizes $\| CS \|_\infty$. A more general uncertainty description, which allows for both poles and zeros crossing into the RHP, is the coprime uncertainty description used by Glover and McFarlane (1989). This leads to an attractive robust stabilization problem formulated in an H^∞ framework. The main results are summarized below.

5.2.1 Normalized coprime factorization

The plant model

$$G = M^{-1}N, \tag{52}$$

is a normalized left coprime factorization (LCF) of G if $M, N \in RH_\infty$ (the set of stable real rational transfer function matrices) and $MM^* + NN^* = I$ where for a real rational function of s, X^* denotes $X^T(-s)$.

With the notation

$$G(s) = D + C(sI - A)^{-1}B \stackrel{s}{=} \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \tag{53}$$

a state-space realization of a normalized coprime factorization of G is given (Vidyasagar, 1988) by

$$\begin{bmatrix} N & M \end{bmatrix} \stackrel{s}{=} \left[\begin{array}{c|cc} A + HC & B + HD & H \\ \hline R^{-1/2}C & R^{-1/2}D & R^{-1/2} \end{array}\right] \tag{54}$$

where

$$H = -(BD^T + ZC^T)R^{-1} \tag{55}$$

$$R = I + DD^T \tag{56}$$

and the matrix $Z \geq 0$ is the unique stabilizing solution to the algebraic Riccati equation (ARE)

$$(A - BS^{-1}D^TC)Z + Z(A - BS^{-1}D^TC)^T - ZC^TR^{-1}CZ + BS^{-1}B^T = 0 \tag{57}$$

where

$$S = I + D^TD. \tag{58}$$

5.2.2 Perturbed plant model

A perturbed model G_p can be defined as

$$G_P = (M + \Delta_M)^{-1}(N + \Delta_N) \tag{59}$$

where $\Delta_M, \Delta_N \in RH_\infty$ and $\left\| \begin{matrix} \Delta_N \\ \Delta_M \end{matrix} \right\|_\infty < 1$, as illustrated in figure 5.3.

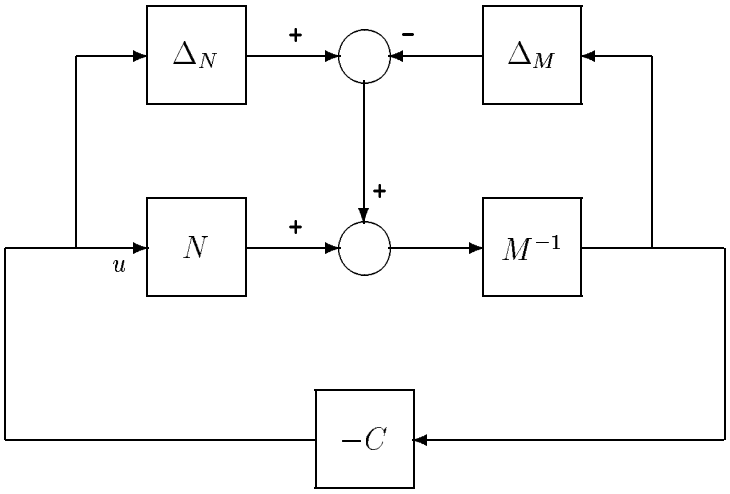

Figure 5.3: Perturbed plant model and controller

5.2.3 Robust stabilization

The robust stability condition for the class of perturbed models defined by (59) was derived previously in (40). For $\left\| \begin{matrix} \Delta_N \\ \Delta_M \end{matrix} \right\|_\infty < 1$ we have

$$RS \quad \text{iff} \quad \gamma \stackrel{\text{def}}{=} \left\| \begin{bmatrix} C \\ I \end{bmatrix} (I + GC)^{-1} M^{-1} \right\|_\infty \leq 1 \tag{60}$$

A reasonable objective is therefore to find the stabilizing controller that minimizes γ and thus allows for the largest perturbations. This is the problem of robust stabilization of normalised coprime factor plant descriptions as introduced by Glover and McFarlane (1989). The minimum value of γ for all stabilizing controllers C is

$$\gamma_0 = \inf_{C \text{ stabilising}} \left\| \begin{bmatrix} C \\ I \end{bmatrix} (I + GC)^{-1} M^{-1} \right\|_\infty \tag{61}$$

and is given in Glover and McFarlane (1989) by

$$\gamma_0 = \left(1 - \| [N, M] \|_H^2\right)^{-1/2}, \tag{62}$$

where $\| \cdot \|_H$ denotes the Hankel norm. From (Glover and McFarlane, 1989)

$$\| [N, M] \|_H^2 = \lambda_{\max} \left(ZX(I + ZX)^{-1}\right), \tag{63}$$

where $\lambda_{\max}(.)$ represents the maximum eigenvalue, and $X \geq 0$ is the unique stabilizing solution of the ARE

$$(A-BS^{-1}D^TC)^TX+X(A-BS^{-1}D^TC)-XBS^{-1}B^TX+C^TR^{-1}C=0. \quad (64)$$

Hence, it can be shown that

$$\gamma_0 = (1+\lambda_{\max}(ZX))^{1/2}. \quad (65)$$

A controller which achieves γ_0 is given in (McFarlane and Glover, 1990) by

$$C \stackrel{s}{=} \left[\begin{array}{c|c} A+BF+\gamma_0^2(Q^T)^{-1}ZC^T(C+DF) & \gamma_0^2(Q^T)^{-1}ZC^T \\ \hline B^TX & -D^T \end{array}\right], \quad (66)$$

where

$$F = -S^{-1}(D^TC+B^TX), \quad (67)$$

and

$$Q = (1-\gamma_0^2)I + XZ. \quad (68)$$

The above results on robust stabilization are particularly attractive because the optimal γ and the corresponding optimal controller can be found without an iterative search on γ which is normally required to solve H^∞ problems.

In the next section, it is shown how the robust stabilization problem can be used in conjunction with the ideas of Doyle and Stein on singular value loop shaping to arrive at a reliable multivariable loop shaping design procedure.

5.3 Loop shaping design

Robust stabilization alone is not much used in practice because the designer is not able to specify the desired performance requirements. To do this McFarlane and Glover (1990) proposed pre- and post-compensating the plant to shape the open-loop singular values prior to robust stabilization of the "shaped"plant.

If W_1 and W_2 are the pre- and post-compensators respectively, then the shaped plant G_S is given by

$$G_S = W_2GW_1 \quad (69)$$

as shown in figure 5.4. The controller C is synthesised by solving the robust stabilization problem of section 5.2 for the shaped plant G_S with a normalized left coprime factorization $G_s = M_s^{-1}N_s$. The feedback controller for the plant G is then $C = -W_1C_sW_2$.

The above procedure contains all the essential ingredients of classical loop shaping, and can be easily implemented using reliable algorithms in, for example,

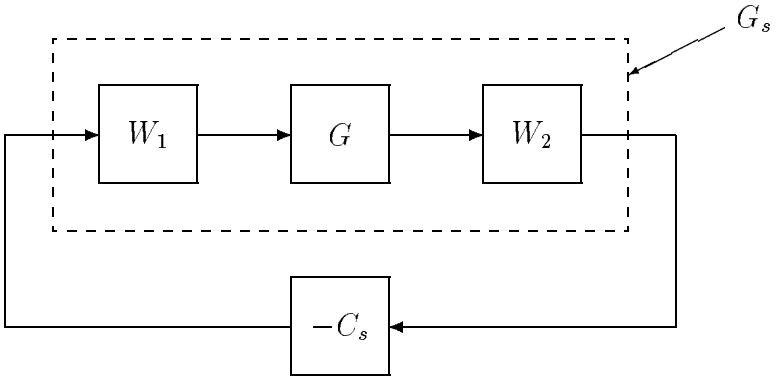

Figure 5.4: The shaped plant and controller

Matlab. Skill is required in the selection of weights, but experience on real applications has shown that robust controllers can be designed with relatively little effort by following a few simple guidelines. Hyde (1991) offers a step by step procedure for weights selection developed during his Ph.D work with Glover on the robust control of VSTOL aircraft. These guidelines are summarised below in subsection 5.3.1. Successful application of the procedure has also been reported by Postlethwaite and Walker (1992) in their work on advanced control of high performance helicopters some of which will be described in section 6.

5.3.1 A loop shaping design procedure

The following procedure is a summary of that found in (Hyde, 1991):

1. Scale the outputs so that the same amount of cross coupling into each of the outputs is equally undesirable.

2. Scale the inputs to reflect the relative actuator capabilities or expected usage. This may involve a few iterations based on the control signals which result from successive designs.

3. The inputs and outputs should be ordered so that the plant is as diagonal as possible. The relative gain array can be useful here.

4. Select the elements of diagonal pre- and post-compensator weights W_1 and W_2 so that the roll off rates of the singular values are approximately 20

dB/decade at the desired bandwidths. Some trial and error is involved here.

W_2 is often chosen as a constant reflecting the relative importance of the outputs to be controlled while W_1 contains the dynamic shaping.

Integral action (for steady-state accuracy) and high frequency roll off (for noise attenuation and robustness) should be placed in W_1 if desired.

5. Sometimes it is found useful to "align "the singular values at the desired bandwidth using a further constant weight W_A cascaded with W_1. This is effectively a decoupler and should not be used if the plant is ill-conditioned.

6. Robust stabilization of the shaped plant is carried out as described in section 5.2. If the optimal gamma, γ_0, is less than about 4, then the design is usually successful. That is, the shape of the open-loop singular values will not have changed much after robust stabilization. A large value of γ indicates that the specified singular value shapes are incompatible with robust stability requirements.

7. Analysis of the design may prompt further modifications of the weights if all the specifications are not met.

8. When implementing the controller, the configuration shown in figure 5.5 has been found useful when compared with the conventional set up in figure 5.1. This is because the references do not directly excite the dynamics of C_s which can result in large amounts of overshoot (classical derivative kick). The prefilter ensures a steady state gain of 1 between r and y.

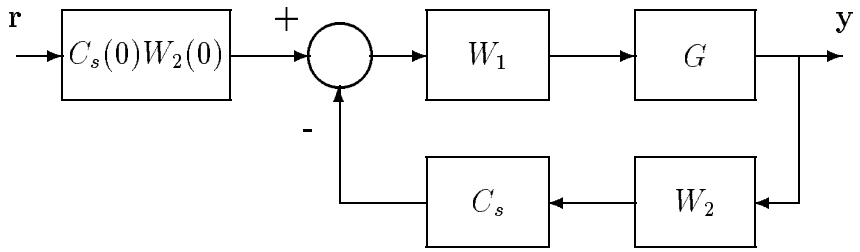

Figure 5.5: A practical implementation of the loop shaping controller

5.3.2 Loop shaping design and the method of inequalities

In Whidborne et al (1992), it has recently been shown how the method of inequalities (Zakian and Naib, 1973) can be used in the loop shaping design procedure to select the weights W_1 and W_2 to satisfy a given set of performance inequalities. Although computationally demanding the technique has proved useful when stringent performance specifications are required to be met.

The method of inequalities (MOI) introduced by Zakian (Zakian and Naib, 1973) is a computer-aided multi-objective design approach, where desired performance is represented by a set of algebraic inequalities, and the aim of the design is to simultaneously satisfy these inequalities. The design problem is expressed as

$$\phi_i(p) \leq \varepsilon_i \quad \text{for} \quad i = 1 \ldots n \tag{70}$$

where ε_i are real numbers, $p \in \mathcal{P}$ is a real vector $(p_1, p_2, \ldots, p_q)$ chosen from a given set $\mathcal{P}$ and ϕ_i are real functions of p. The functions ϕ_i are performance indices, the components of p represent the design parameters and ε_i are chosen by the designer and represent tolerable values of ϕ_i. The aim is the satisfaction of the set of inequalities in order that an acceptable design is reached.

The functions $\phi_i(p)$ may be functionals of the system step response, for example the rise time, overshoot or the integral absolute error, or functionals of the frequency response, such as the bandwidth. They can also represent measures of the system stability. The actual solution to the set of inequalities (70) may be obtained by means of numerical search algorithms, such as the moving boundaries process (Zakian and Naib, 1973).

In some previous applications of the MOI, the design parameter has parameterized a controller with a particular structure. For example, $p = (p_1, p_2)$ could parameterize a PI controller $p_1 + p_2/s$. This has meant that the designer has had to choose the structure of the control scheme and the order of the controllers. In general, the smaller the size of the design vector p, the easier it is for the numerical search algorithm to find a solution, if one exists. While this does give the designer some flexibility and leads to simple controllers, and is of particular value when the structure of the controller is constrained in some way, it does mean that better solutions may exist with more complicated and higher order controllers. A further limitation of using the MOI in this way is that a stability point must be located as a pre-requisite to searching the parameter space to improve the index set ϕ, that is a point such that $\phi_i < \infty$ for $i = 1, 2, \ldots, n$ must be found initially.

Two aspects of design using the loop shaping design procedure (LSDP) make it amenable to combine this approach with the MOI. Firstly, unlike most H_∞-optimization problems, the H_∞-optimal controller for the weighted plant can

be synthesised from the solution of just two ARE's and does not require time-consuming γ-iteration. Secondly, in the LSDP, the weighting functions are chosen by considering the open-loop response of the weighted plant, so effectively the weights W_1 and W_2 are the design parameters. This means that the design problem can be formulated as in the method of inequalities, with the weighting parameters used as the design parameters p to satisfy some set of *closed-loop* performance inequalities.

Such an approach to the MOI overcomes the limitations to the MOI described earlier. The designer does not have to choose the order or structure of the controller, but instead chooses the structure and order of the weighting functions. With low-order weighting functions, high order controllers are synthesised which often leads to significantly better performance or robustness than if simple low order controllers were used. Additionally, the problem of finding a stability point is simply a case of choosing sensible initial weighting functions; an easy matter if the open-loop singular value plots are studied.

For more details of loop shaping design and the method of inequalities see Whidborne et al (1992).

5.4 Two degrees of freedom controllers

Most control design problems naturally possess two degrees of freedom (DOF). In general this arises from the existence of, on the one hand, measurement or feedback signals and on the other, commands or references. Quite often, one degree of freedom is forsaken in the design, and the controller is driven by, for example, an error signal (i.e. the difference between command and output). Other *ad hoc* means may also be used to arrive at a 1-DOF implementation. A general 2-DOF feedback control scheme is depicted in figure 5.6. The commands and feedbacks enter the controller separately where they are independently processed.

5.4.1 An extended loop shaping design procedure

Limebeer et al (1993) have recently proposed an extension of McFarlane and Glover's loop shaping design procedure (LSDP) which uses a 2-DOF scheme to enhance the model matching properties of the closed-loop system. The feedback part of the controller is designed to meet robust stability and disturbance rejection requirements in a manner identical to the 1-DOF LSDP. That is, weights are first selected to produce a shaped plant with desirable singular values. An additional prefilter part of the controller is then introduced to force the response of the closed loop system to follow that of a specified model T_0.

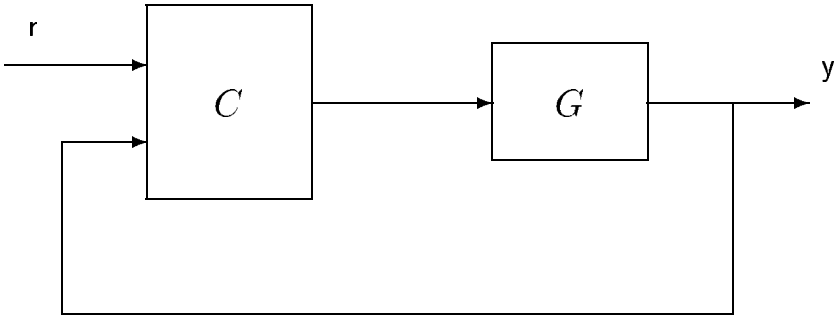

Figure 5.6: General 2-DOF feedback control scheme

The design problem to be solved is illustrated in figure 5.7. The scalar parameter ρ is used to adjust the emphasis that is placed on model matching in the optimization. For $\rho = 0$, the problem reverts to the standard LSDP. As ρ is increased, more emphasis is placed on model following.

The H^∞ optimization problem to be solved is that of finding a controller C which stabilizes G and which minimizes the H^∞ norm of the transfer function between the signals $(r^T \phi^T)^T$ and $(u^T y^T z^T)^T$ as defined in figure 5.7. Note that the robust stabilization problem alone involves minimizing the H^∞ norm of the transfer function between ϕ and $(u^T y^T)^T$. The two degrees of freedom design problem is easily solved using standard routines, in for example Matlab, but as a standard H^∞ optimization problem an iterative approach is required. In practice, sub-optimal controllers are often used which satisfy a given bound γ on the transfer function being minimized.

The two degrees of freedom approach will usually also involve a few iterations on ρ to achieve the desired balance between robust stabilization and model matching.

5.4.2 A further extension using the method of inequalities

In section 5.3, we saw how the one degree of freedom loop shaping design procedure could be enhanced by using the method of inequalities to select the weights W_1 and W_2. It is tempting therefore to think that the same could be done for the 2-DOF approach. However, the latter requires γ-iteration for its solution which makes it too slow computationally to be effectively combined with the MOI. Whidborne et al (1993), therefore, proposed an alternative strategyy based on fixing the structure of the prefilter part of a 2-DOF controller.

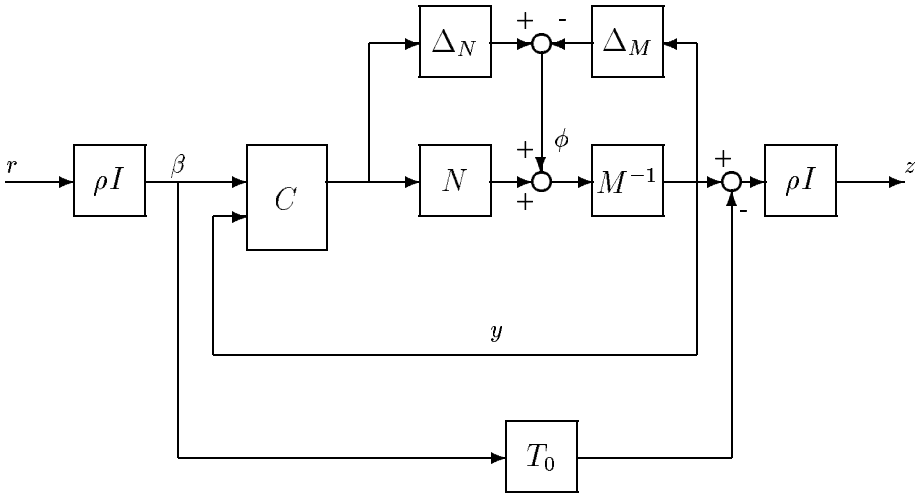

Figure 5.7: Two degrees-of-freedom design problem

The proposed approach involves adding a prefilter C_p to a feedback controller C designed, using the 1-DOF LSDP, as illustrated in figure 5.8. C_p is parameterised with a subset of design parameters while C is the solution to the LSDP with weights W_1 and W_2 parameterised with the remaining design parameters.

Functional constraints

$$\phi_i(W_1, W_2, C_p) \leq \varepsilon_i \quad \text{for} \quad i = 1 \ldots n \tag{71}$$

can then be defined to represent performance requirements and a MOI approach used to find the parameters of W_1, W_2 and C_p. Given W_1 and W_2, C_p follows straight from a 1-DOF LSDP with no iterations.

For more details see Whidborne et al (1993), where the MOI approach has been successfully used to design a 2-DOF controller for the distillation column benchmark example.The results of this case study will be presented at the mini-course on Robust Multivariable Control using H^∞ Methods (2nd European Control Conference, Groningen, 1993) for which this paper has been prepared. In the next section, a second case study on helicopter control will be considered. It is a straightforward application of the loop shaping 2 DOF design methodology without MOI.

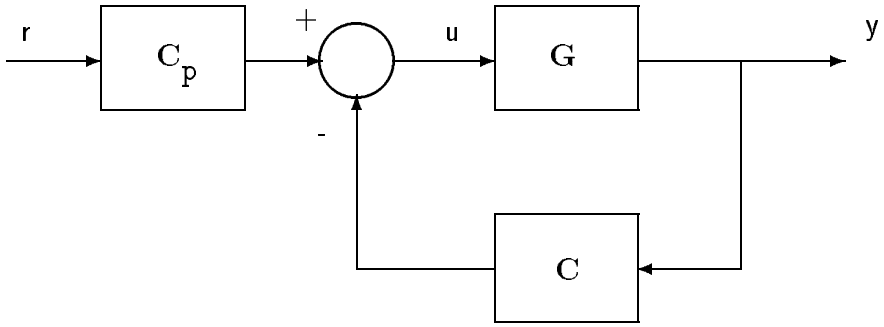

Figure 5.8: A 2 degrees of freedom controller configuration

6 Advanced control of high performance helicopters: a case study

To fly a high performance helicopter in low-level or nap-of-the-earth flight currently demands a high pilot workload; so high in fact as to limit the potential of the aircraft to a level below what is theoretically possible. Thus in order to enable the next generation of helicopters to fulfil the challenging specifications that are emerging, an automatic flight control system will be an essential ingredient. In this section, we will report on the findings of an on-going study into the role of advanced multivariable control theory in the design of full-authority full-flight-envelope control laws.

6.1 Background

The Control Systems Research Group at Leicester University has several years experience in the design of advanced control laws for future generation helicopters. For the past three years a major research project has been undertaken, funded by the UK Defence Research Agency (DRA) Bedford, formerly the Royal Aerospace Establishment, to investigate the role of advanced multivariable control theory in the design of full-authority full-flight envelope control laws.

This section outlines some of the salient features and results arising out of this DRA-funded research. The work has enabled an in-depth study using computer simulation to help assess the impact that advanced control systems might play in improving the handling qualities of future military helicopters.

The main achievements have been to extend and to improve upon the results of earlier work (Yue and Postlethwaite, 1990) and to demonstrate that multivariable design methods using H^∞-optimization, and in particular the loop shaping methodology of section 5, provide a valuable way forward in the design of robust full-flight-envelope control systems.

In May 1992, an important goal of the research project was achieved, with the successful piloted simulation, using the Large Motion System simulator at DRA Bedford, of a multivariable control system designed for wide envelope use. The system was tested over a period of three days by two experienced helicopter test pilots, one from the Royal Navy, the other from the Army. The testing consisted of two phases: the first, a familiarisation phase, during which the pilots could accustom themselves to the response types available from the control system and generally gain a feel for how to fly the aircraft via the controller; the second, the test phase, during which the pilots were asked to perform a set of specified tasks, each designed to highlight certain characteristics of the aircraft's response. Each pilot completed an in-cockpit assessment of the system's response using the Cooper-Harper pilot rating scale (Cooper and Harper, 1969). With this the pilot can classify the desirable and unsatisfactory handling aspects on a points system, scaled from 1 to 10, where 1 represents the most satisfactory qualities. A rating of 10 represents major and unacceptable system deficiences, where control may be lost during part of the flight envelope. Cooper-Harper ratings of 1 to 3 are said to conform to a level 1 handling qualities rating and are a goal of any helicopter flight control system.

The control system tested in May 1992 received level 1 ratings in a large majority of the runs made. The controller was also tested over a wide range of speeds, from hover to well in excess of 100 knots. During the three days of tests, it became clear that in simulation at least, the multivariable controller was able to provide robust stability and decoupled performance. Both pilots agreed that, in spite of certain deficiencies in the primary yaw response and collective to yaw coupling, the control law provided excellent stability and control.

Following the May trials a redesign was undertaken to increase the yaw axis bandwidth and to slightly reduce the heave axis bandwidth. This was a very simple matter with the 2-DOF loop shaping design technique being used. The new design was fully tested in piloted simulation in December 1992. At these trials, the previously identified deficiencies were no longer present and all the mission task elements performed were given level 1 Cooper-Harper ratings.

6.2 The helicopter model

The aircraft model used in our work is representative of the Westland LYNX, a twin-engined multi-purpose military helicopter, approximately 9000 lbs gross weight, with a four-blade semi-rigid main rotor. The unaugmented aircraft is unstable, and exhibits many of the cross-couplings characteristic of a single main-rotor helicopter. These characteristics have been captured by a computer model known as the Rationalized Helicopter Model (RHM) (Padfield, 1981) that was used in our study. This model has been developed at DRA Bedford over a number of years and is a mature and fairly accurate (though by no means definitive) nonlinear model of the Lynx. In addition to the basic rigid body, engine and actuator components, it also includes second order rotor flapping and coning modes for off-line use. The model has the advantage that essentially the same code can be used for the real-time piloted simulation as for the workstation-based off-line handling qualities assessment.

The equations governing the motion of the helicopter are complex and difficult to formulate with high levels of precision. For example, the rotor dynamics are particularly difficult to model. A robust design methodology is therefore essential for high performance helicopter control.

The starting point for our designs was a set of five eighth-order linear differential equations modelling the small-perturbation rigid body motion of the aircraft about five trimmed conditions of straight-and-level flight in the range 0 to 80 knots. The controller designs were first evaluated on the eighth-order models used in the design, then on twenty-one state linear models, and finally using the full nonlinear model. The robust design methodology used in the controller design did turn out to provide excellent robustness with respect to nonlinearities and time delays which, although simulated, were not explicitly included in the linear design process.

6.3 Design objectives

The main objectives were to design a full-authority control system that:

- Robustly stabilized the aircraft with respect to changes in flight condition, and model uncertainty and non-linearity.
- Provided high levels of decoupling between primary controlled variables.
- Achieved compliance with the Level 1 criteria given in the US Army Aeronautical Design Standards, (ADS-33C, 1989).

6.4 Design method

The two degrees of freedom loop shaping design procedure of section 5.4 was used to robustly stabilize the aircraft over a wide range of flight conditions, whilst simultaneously forcing the closed-loop system to approximate the behaviour of a specified transfer function model T_0. The overall control law was actually comprised of five controllers, designed at a range of flight conditions between 0 and 80 knots, each one having a Kalman filter-like structure. The latter is a property of the LSDP and is very useful when scheduling between controllers over the flight envelope. As the dynamics of the open-loop aircraft vary with speed, so too do the controllers obtained at each operating point. Therefore, these controllers can be scheduled with forward speed if required, to give wide-envelope performance.

The aim of the design was to synthesize a full-authority controller that robustly stabilized the aircraft and provided a decoupled Attitude-Command/ Attitude-Hold (ACAH) response type that closely approximated the behaviour of a simple transfer-function model.

The outputs to be directly controlled were:

- Heave velocity
- Pitch attitude
- Roll attitude
- Heading rate

With a full authority control law such as that proposed here, the controller has total control over the blade angles, and is interposed between the pilot and the actuation system. The pilot flies the aircraft by isuing appropriate demands to the controller. These demands, together with the sensor feedback signals, are fed to the flight control computer which generates appropriate blade angle demands. Other than that we make no assumptions about the implementation details.

The controller was designed to operate on six feedback measurements: the four controlled outputs listed above and the body-axis pitch and roll rate signals. The other inputs to the controller consisted of the 4 pilot inceptor inputs.

The control law output consisted of four blade angle demands:

- Main rotor collective
- Longitudinal cyclic
- Lateral cyclic

- Tail rotor collective

These demands were passed directly to the actuator model.

6.5 Weighting function selection, the design parameter ρ, and the desired transfer function T_0

The same loop shaping weights W_1 and W_2, and the same desired transfer function T_0, were used for all 5 operating point designs. It was found that the ADS-33C bandwidth requirements impact directly on the "cross over "frequency of the weight W_1 which was chosen to have the first order diagonal form

$$W_1 = diag\left\{\frac{s+a}{s}, \frac{s+b}{s}, \frac{s+c}{s}, \frac{s+d}{s}\right\}. \tag{72}$$

A static diagonal W_2 was chosen as

$$W_2 = diag\{1, 1, 1, 1, 0.1, 0.1\} \tag{73}$$

That is, the roll rate and pitch rate signals are weighted less than the other four outputs which are required to be controlled. The rates are included as extra measurements because it is well known that they will make the control problem easier.

The desired closed-loop transfer function T_0 is chosen to be diagonal with second order transfer functions on each of the four channels: heave velocity, pitch and roll attitudes and heading rate. The damping and natural frequencies of these transfer functions were selected to give what were considered to be adequate responses. The selection of ρ is a compromise between robust stability

ρ	0	0.1	0.2	0.4	0.75	1.0	1.5	2.0	3.0
γ_o	2.89	2.90	2.92	2.99	3.23	3.46	3.98	4.59	6.35

Table 1: Relationship between ρ and γ_o for a hover design

and model matching. Table 1, shows the relationship between ρ and γ_0 for a hover design, where γ_0 is the minimum H^∞-norm of the transfer function being minimised. The reciprocal of γ_0 is roughly proportional to the multivariable stability margin. For the hover design in question a value of $\rho = 1.5$ was used, together with a suboptimal value of $\gamma = 4.2$.

6.6 Controller scheduling

The controller was designed to run in either of two modes: (i) fixed gain, (ii) interpolated. In fixed gain mode, the closest controller for the given flight condition would be switched in to provide control. This controller would remain operative until the mode was de-selected. If the interpolated mode was engaged, the controllers would be interpolated smoothly as a function of air-speed to compensate for variation in dynamics. To implement for real would require an accurate measurement (or estimate) of forward air-speed.

6.7 Outer-loop modes

To enhance the handling qualities provided by the basic ACAH response of the inner loop H^∞ controller, three outer loop modes were also implemented.

- Turn coordination: this was provided by augmenting the heading rate demand as a function of bank angle at moderate/high speed. This enabled a coordinated turn to be effected as a single axis task.
- Automatic trimming: this was achieved using a trim-map to offset the linear inner loop controller with the appropriate trim attitude.
- Hover acquisition/hold: this mode enabled the pilot to acquire and hold hover automatically. Longitudinal and lateral velocty state estimates were needed to achieve this.

During the piloted trials, the first two modes were used continuously, but insufficient time was available to evaluate the hover acquisition utility.

6.8 Step response analysis

The response of the closed-loop system (comprising controller and full nonlinear model) to step input demands on pitch and roll channels are shown in figures 6.1 and 6.2. These show, respectively, an acceleration from hover and the commencement of a coordinated turn at 60 knots. In both cases there is seen to be minimal cross-coupling.

Figure 6.1: Pitch axis step response: outputs and actuators

Figure 6.2: Roll axis step response: outputs and actuators

6.9 Handling qualities assessment: off-line analysis

ADS-33C (1989) details the latest requirements' specification for military helicopters which is intended to ensure that mission effectiveness will not be compromised by deficient handling qualities. The requirements are stated in terms of three limiting " levels "of acceptability of one or more given parameters. The levels indicate performance attributes that equate with pilot ratings on the Cooper-Harper scale. A Matlab Handling Qualities Toolbox (Howitt, 1991) was used as a supplement to existing computer aided control system design packages in order to integrate handling qualities assessment into the complete design and analysis cycle. The dynamics of the closed loop vehicle were assessed against the dynamic response requirements specified in sections 3.3 and 3.4 of ADS-33C using the off-line simulation model. A selection of the results are given in Walker et al (1993). In summary, the performance provided by the control law led to level 1 handling quality ratings for almost all of the mission tasks performed.

6.10 Handling qualities assessment: piloted simulation on the DRA Bedford large motion simulator

The simulation model was written in Fortran and run on an Encore Concept-32 computer with an integration step of 20 mS. A Lynx-like single seat cockpit was used, mounted on the large motion system which provides ± 30 degrees of pitch, roll and yaw, ± 4 metres of sway and ± 5 metres of heave motion. Also, the pilot's seat was dynamically driven to give vibration and sustained normal acceleration cues. The visual display was generated by a Link-Miles Image IV CGI system and gave approximately 48 degrees field of view (FOV) in pitch and 120 degrees FOV in azimuth with full daylight texturing. A three axis side-stick was used to control pitch, roll and yaw together with a conventional collective for heave.

Handling qualities were assessed for three hover/low speed mission task elements (sidestep, quick-hop, bob-up) and three moderate/high speed tasks (lateral jinking, hurdles, yaw pointing) using CGI databases developed by DRA Bedford.

Two DRA test pilots took part in the trials (of May and December 1992), both with significant experience of Lynx and the simulator. For each task in turn, the pilot performed two or three familiarisation runs before performing a definitive evaluation run, at the end of which the simulation was paused so that comments and handling qualities ratings could be recorded. The six tasks are briefly described below.

Sidestep: With reference to figure 6.3a, the objective was to translate side-

ways through 150 ft from a hover at a height of 30 ft above ground level in front of one diamond and square sighting arrangement, to acquire and maintain a stable hover in front of the next sighting system.

Quick-hop: The quick-hop task (figure 6.3b) is the corresponding longitudinal task to the sidestep, requiring a re-position from hover over a distance of 500 ft. Again, similar levels of initial pitch attitude were used to determine the task aggression. The task was flown down a walled alley to give suitable height and lateral position cues.

Bob-up: The bob-up task was performed in front of one of the V-notch hurdles (figure 6.3c). From a hover aligned with the bottom of the V-notch, the pilot had to acquire and maintain a new height denoted by a mark on the notch.

Lateral jinking: The lateral jinking task concerned a series of 'S' turns through slalom gates followed by a corresponding line tracking phase (figure

6.3d). The task had to be flown whilst maintaining a speed of 60 knots and a height of 25 ft.

Hurdles: Using the same V-notch hurdles as seen for the bob-up task, a collective-only flight path re-positioning task was flown at 60, 75 and 90 knots to represent increasing task aggression. From an initial height aligned with the bottom of the V-notch, the pilot had to pass through each hurdle at the height denoted by a mark on the notch and then regain the original speed and height as quickly as possible.

Yaw pointing: Whilst translating down the runway centre line at 60 knots, the pilot was required to yaw to acquire and track one of a number of offset posts.

Table 2 is a detailed compilation of one of the pilot's questionnaires based on the May 1992 trials. The primary response in heave, pitch and roll was excellent. But the primary response in yaw was sluggish and there was some undesirable collective-to-yaw coupling.

A redesign was undertaken to increase the yaw axis bandwidth and to slightly reduce the heave axis bandwidth. This was done very simply in the two degrees of freedom loop shaping design procedure by simply modifying the desired closed-loop transfer function T_0.

The new design was tested in December 1992 and achieved pilot ratings of level 1 for all six tasks.

6.11 Conclusion

The results of this case study have demonstrated that multivariable design techniques can play a significant role in the design of control systems for high performance helicopters.

7 Conclusions

The paper has provided an introduction to frequency domain methods for the analysis and design of multivariable control systems. Particular attention was given to H^∞ methods and to problems of robustness which arise when plant models are uncertain, which is always the case. The additional problems associated with the control of ill-conditioned plants were also considered.

The relative gain array, the singular value decomposition and the structured singular value were shown to be invaluable tools for analysis.

For multivariable design, emphasis was given to the shaping of the singular values of the loop transfer function. The technique of McFarlane and Glover and its extension to two degrees of freedom controllers were considered in detail.

The power of the approach was demonstrated by its application to the design of a full-authority wide-envelope control system for a high performance helicopter.

8 Acknowledgements

The authors are grateful to Mr Petter Lundström (Trondheim University) and Mr Neale Foster (Leicester University) for their comments and assistance in preparing these notes.

The helicopter controller design was largely completed by Dr. Daniel Walker (Leicester University) and section 6 draws heavily on the papers by Postlethwaite and Walker (1992) and Walker et al (1993).

The helicopter case study was conducted with the support of the UK Ministry of Defence through Extramural Research Agreement No. 2206/32/RAE(B).

9 References

Anon., Aeronautical Design Standard ADS-33C "Handling Qualities Requirements for Military Rotorcraft", *US Army AVSCOM* , August, 1989.

Chiang And Safonov, *Robust Control toolbox for Matlab. User's guide.* The Math-Works, South Natick, MA, USA (1988, 1992).

C.E.Cooper and R.P.Harper, "The use of pilot rating scale in the evaluation of aircraft handling qualities", NASA TM-D-5133, 1969.

D.J.Walker, I.Postlethwaite, J.Howitt, N.P.Foster, "Rotorcraft Flying Qualities Improvement Using Advanced Control", *Proc. American Helicopter Society/NASA Conf.* , San Francisco, January, 1993.

J.C. Doyle, "Analysis of Feedback Systems with Structured Uncertainties", *IEE Proc*, **129** (D), 242-250 (1982).

J.C.Doyle and G.Stein, "Multivariable Feedback Design: Concepts for a Classical/Modern Synthesis", *IEEE Trans. AC* , **26, 1** , 4-16, 1981.

J.C. Doyle, J.E. Wall and G. Stein, "Performance and Robustness Analysis for Structured Uncertainty", *Proc. IEEE Conf. on Decision and Control*, Orlando, Florida, Dec. 1992.

J.C. Doyle, Lecture Notes *ONR/Honeywell workshop on Adcances in Multivariable Control*, Minneapolis, MN (1984).

J.C. Doyle, K.Lenz, and A.K. Packard, "Design examples using μ-synthesis: Space shuttle lateral axis FCS during reentry", in NATO ASI Series, **F34**, *Modelling, Robustness and Sensitivity Reduction in Control Systems*, R.F. Curtain (Ed.), Springer-Verlag (1987).

K.Glover and D.C.McFarlane, "Robust stabilization of normalised coprime factor plant descriptions with H^∞-bounded uncertainty", *IEEE Trans. AC*, **34, 8** , 821-830, 1989.

M. Hovd and S. Skogestad, "Simple Frequency-Dependent Tools for Control System Analysis, Structure Selection and Design", *Automatica*, **28**, 5, 989-996 (1992).

J.Howitt, "Matlab toolbox for handling qualities assessment of flight control laws", *Proc. IEE Control '91*, Scotland, 1991.

R.A. Hyde, *The Application of Robust Control to VSTOL Aircraft*, Ph.D. Thesis, Cambridge University, 1991.

H.Kwakernaak, "Optimal Low-Sensitivity Linear Feedback Systems" *Automatica* , **5**, 279, 1969.

D.J.N.Limebeer, E.M.Kasenally, and J.D.Perkins, "On the Design of Robust Two Degree of Freedom Controllers", *Automatica* , **29, 1**, 157-168, 1993.

P. Lundström, S. Skogestad and Z-Q. Wang, "Performance Weight Selection for H-infinity and mu-control metods", *Trans. Inst. of Measurement and Control*, **13**, 5, 241-252, 1991.

P. Lundström, S. Skogestad and Z.Q. Wang, "Uncertainty Weight Selection for H-infinity and Mu-Control Methods", Proc. IEEE Conf. on Decision and Control (CDC), 1537-1542, Brighton, UK, Dec. 1991b.

J.M. Maciejowski, *Multivariable Feedback Design*, Addison-Wesley (1989).

D.C.McFarlane and K.Glover, *Robust Controller Design Using Normalised Coprime Factor Plant Descriptions*, Springer-Verlag , Berlin, 1990.

M. Morari and E. Zafiriou, *Robust Process Control*, Prentice-Hall (1989).

Owen and Zames, "Unstructured uncertainty in $\mathbf{H}^\infty$", In: *Control of uncertain dynamic systems*, Bhattacharyya and Keel (Eds.), CRC Press, Boca Raton, FL , 3-20 (1991).

G.D.Padfield, "Theoretical model of helicopter flight mechanics for application to piloted simulation", *RAE* , TR 81048, 1981.

I.Postlethwaite and D.J.Walker, "Advanced Control of High Performance Rotorcraft", *Proc. IMA Conf. on Aerospace Vehicle Dynamics and Control*, Cranfield Inst. of Technology, September, 1992.

S. Skogestad and M. Morari, "Implication of Large RGA-Elements on Control Performance", *Ind. Eng. Chem. Res.*, **26**, 11, 2323-2330 (1987).

S. Skogestad, M. Morari and J.C. Doyle, "Robust Control of Ill Conditioned Plants: High-Purity Distillation", *IEEE Trans. Autom. Control*, **33**, 12, 1092-1105 (1988). (Also see *correction* to μ-optimal controller in **34**, 6, 672 (1989)).

F. van Diggelsen and K. Glover, "Element-by-element weighted $\mathbf{H}^\infty$-Frobenius and H_2 norm problems", it Proc. 30 th IEEE Conf. on Decision and Control (CDC), Brighton, England, 923-924, 1991.

F. van Diggelsen and K. Glover, "A Hadamard weighted loop shaping design procedure", it Proc. 31 th IEEE Conf. on Decision and Control (CDC), Tuscon, Arizona, 2193-2198, 1992.

M.Vidyasager, "Normalised coprime factorizations for non strictly proper systems", *IEEE Trans. AC*, **33**,300-301, 1988.

J.F.Whidborne, I.Postlethwaite, D.W.Gu, "Robust controller design using H^∞ loop shaping and the method of inequalities", *Leicester University Engineering Department (LUED)*, Report No. 92-33, 1992.

E.A. Wolff, S. Skogestad, M. Hovd and K.W. Mathisen, "A procedure for controllability analysis", *Preprints IFAC workshop on Interactions between process design and process control, London, Sept. 1992*, Edited by J.D. Perkins, Pergamon Press, 1992, 127-132.

C.C. Yu and W.L. Luyben, "Robustness with Respect to Integral Controllability", *Ind. Eng. Chem. Res.*, **26**, 1043-1045 (1987).

A.Yue and I.Postlethwaite, "Improvement of helicopter handling qualities using H^∞ optimization". *IEE Proc.*, Part D, 137, 115-129, 1990.

V. Zakian and U.Al-Naib, "Design of dynamical and control systems by the method of inequalities", *Proc. IEE*, **120, 11** , 1421-1427, 1973.

Ziegler and Nichols, "Process lags in automatic-control circuits", *Trans. of the A.S.M.E.*, **65**, 433-444 (1943).

Task	Level of Aggression	Comments	HQR	Level
Side Step	Low	Loads of spare capacity	2	1
	Mid	Task workload still minimal, response perfect.	2	1
	High	Increased level of aggression does not increase workload. Very easy	2(low)	1
Quick Hop	Low	Desired performance easily achieved. Slight right drift. 3-axis task. A lot of inertia in model. Control law good.	2	1
	Mid	Easier at higher aggression because less anticipation required. No problems.	2	1
Hurdles	Low	Desired performance achieved satisfactorily. Yaw coupling only problem, but spare capacity.	3	1
	High	At top of hurdle, control activity high and little spare capacity. $> 10°$ coupling into heading.	5	2
Lateral jinking	Low	Stacks of spare capacity. Minimal control activity. Single axis task. No cross-coupling.	2	1
	Mid	As above	2	1
	High	As above	3(low)	1
Yaw Pointing	V.Low	Adequate performance achieved with difficulty. Control activity high. Not much spare capacity.Precision difficult.	5	2
	Low	PIO problems. Very high yaw inertia. Low sensitivity, possibly some lag. Maximum rate O.K. but needs to be tighter	7	3

Table 2: Pilot comment from the May 1992 trials