IFAC Workshop on interactions between process design and process control
Imperial College, London, Sept. 6-8, 1992

Controllability Analysis for Unstable Processes

Morten Hovd Sigurd Skogestad*
Chemical Engineering, University of Trondheim-NTH
N-7034 Trondheim, Norway

Abstract

We study the controllability of unstable processes, with emphasis on selection of measurements and manipulated variables, and the pairing problem for decentralized control. The well known pairing criteria based on the Niederlinski Index and the steady state Relative Gain Array (RGA) have been generalized to open loop unstable plants. Right Half Plane (RHP) zeros in individual transfer function elements may make it practically impossible to stabilize the individual loops, and in such cases these pairing criteria are not very helpful. We have found RGA $\approx I$ in the bandwidth region to indicate good pairings also when paired elements have significant RHP zeros, but have found the Direct Nyquist Array (DNA) to perform poorly as an indicator of good pairings. In some cases it is preferable to avoid pairings giving narrow Gershgorin bands.

1 Introduction

In engineering practice, a system is called controllable if it is possible to achieve the specified aims of the control, whatever these may be ([19], p. 171). Unfortunately, in standard state-space control theory the term "controllability" has a rather limited definition in terms of Kalman's *state* controllability, which mainly has to do with realization theory. State controllability will only be considered briefly in this paper, and it will be clear from the context whenever the term controllability refers to a state and not to a plant.

For an unstable plant we must use feedback for stabilization. Thus, whereas the presence of RHP-zeros put an upper bound on the the allowed bandwidth, the presence of RHP-poles put a lower bound. It is also clear that it may be difficult to stabilize an unstable plant if there are RHP-zeros or time delays - "the system goes unstable before we are able to observe what is happening". This qualitative statement is quantifies by the results on sensitivity relationships below.

In most of the paper we assume that a decentralized controller is used, as such controllers are very common in the chemical process industry. A significant amount of work has been done on the choice of pairings for decentralized control of stable plants, e.g. [1], [4], [13], [18] and [20], to reference a few. The concept of Decentralized Integral Controllability (DIC) [20], is not relevant for unstable plants, as the system must necessarily become unstable when the controller gains are sufficiently reduced. In the paper we show how stability based pairing criteria fail or must be interpreted differently when the number of RHP poles in $G(s)$ and $\tilde{G}(s)$ differ ($\tilde{G}(s)$ consists of the diagonal elements of $G(s)$). In contrast, we demonstrate the success of the frequency dependent RGA for unstable plants.

Notation. $G(s)$ is the $n \times n$ transfer function of the process, with the ij'th element denoted $g_{ij}(s)$. $\tilde{G}(s)$ is the transfer function matrix consisting of the diagonal elements of $G(s)$. The controller is denoted $C(s)$, with individual elements $c_i(s)$ for the case of decentralized control.

*Author to whom correspondence should be addressed. E-mail: SKOGE@KJEMI.UNIT.NO, phone: 47-7-594154, fax: 47-7-594080

In the following, the Laplace variable s will be dropped where it is not needed for clarity. The sensitivity function is given by $S = (I + GC)^{-1}$ and the complementary sensitivity function by $H = I - S = GC(I + GC)^{-1}$. Similarly, for decentralized control the sensitivity functions and complementary sensitivity functions of the individual loops can be collected in diagonal matrices, giving $\tilde{S} = (I + \tilde{G}C)^{-1}$ and $\tilde{H} = I - \tilde{S} = \tilde{G}C(I + \tilde{G}C)^{-1}$. The following relationships are also used

$$(I + GC) = (I + E_H\tilde{H})(I + \tilde{G}C) \quad (1)$$

$$E_H = (G - \tilde{G})\tilde{G}^{-1} \quad (2)$$

2 Controllability measures

In this section we review some proposed controllability measures, and generalize some of the controllability measures to unstable plants. In the paper we make use of the multivariable Nyquist theorem, and we will therefore state it here

Theorem 1 *Let the map of the Nyquist D contour under* $\det(I+G(s)C(s))$ *encircle the origin* n_C *times in the clockwise direction. Let the number of open loop unstable poles of* $G(s)C(s)$ *be* n_U*. Then the closed loop system is stable if and only if* $n_C = -n_U$.

Proof: The theorem has been proved several times, see [12].

2.1 State controllability and observability.

It is well known (e.g. [11]) that only states that are both observable and controllable can be stabilized by feedback control. It is therefore necessary that the selection of manipulated and measured variables result in all unstable states being controllable and observable.

We believe that in most cases an engineer with a good understanding of the process will intuitively choose measurements and manipulated variables such that this requirement will be fulfilled. Nevertheless, it does make sense to check that all unstable states are controllable and observable. Preferably, the controllability and observability of the unstable states should not rely on one single manipulated variable or measurement, as the system will then necessarily become unstable if this manipulated variable/measurement fails.

2.2 Sensitivity relationships and RHP-zeros.

Assume a plant $G(s)$ with a RHP-pole at $s = p$ is stabilized by feedback control such that $S(s)$ and $H(s)$ are stable. Then $S(s)$ must have a RHP zero at $s = p$ (follows since $H(s) = G(s)C(s)S(s)$ is stable while $G(s)$ has a RHP pole at $s = p$). This shows that the presence of an RHP-pole imposes restrictions on the closed-loop system in addition to the requirement of stabilization. The approach of first stabilizing the plant with some simple controller, and then proceed as if the RHP-pole never existed, as is proposed by some authors (e.g., [15], is therefore flawed.

To quantify the effect of the RHP-poles on the closed-loop system we shall consider the sensitivity integral relationships of Freudenberg and Looze [2, 3] which extend the Bode integral relationship to plants with RHP-poles and RHP-zeros. Let us first consider Single Input Single Output (SISO) systems, and suppose that $G(s)C(s)$ is rational and has at least two more poles than zeros. Let $G(s)C(s)$ have n_U poles (including multiplicities) in the open right half plane, at locations p_i, $i = 1, 2, \ldots, n_U$. Then, if the closed loop system is stable, the sensitivity function must satisfy:
No RHP zero:

$$\int_0^{\infty} \ln |S(j\omega)| d\omega = \pi \sum_{i=1}^{n_U} Re[p_i] \tag{3}$$

One real RHP zero at $s = z$:

$$\int_0^{\infty} \ln |S(j\omega)| W(z,\omega) d\omega = \pi \sum_{i=1}^{n_U} \ln \left| \frac{\bar{p}_i + z}{p_i - z} \right| \tag{4}$$

$$W(z,\omega) = \frac{2z}{z^2 + \omega^2}$$

$\bar{p}_i$ denotes the complex conjugate of p_i. Eq. (3) and Eq. (4) show that we need $|S(j\omega)| > 1$ over some range of frequencies, which means that the effect of disturbances is actually amplified at these frequencies. A RHP transmission zero of $G(s)$ limits the achievable bandwidth of the plant regardless of the type of controller used (e.g, [16]). This is confirmed by (4) where the shape of the weight $W(z,\omega)$ (equals $2/z$ at low frequencies and falls off with a -2 slope from $\omega = z$) implies that essentially all the positive area for $\ln |S(j\omega)|$ has to be at frequencies lower than the RHP-zero, and there will have to be a peak $|S(j\omega)| > 1$ which will become increasingly large as the crossover frequency approaches z.
Since with no RHP poles the integrals in (3) and (4) equal zero, wee see that the presence of RHP poles increase the area for which $|S| > 1$. We also see that the peak of $|S|$ will approach infinity if $p \to z$. In practice, this means that $|p_i|$ must be smaller than $|z|$ in order to stabilize the plant. In contrast, in the absence of RHP zeros, RHP poles do not impose any practical performance limitations (in terms of peaks in $|S|$), as the frequency range where $|S| > 1$ as required by Eq. (3) may be arbitrarily large and is only limited by high frequency roll-off considerations.

For Multiple Input Multiple Output (MIMO) systems the situation is not so clear, although some useful insight exists. For a $n \times n$ system with no RHP-zeros we get

$$\sum_{i=1}^{n} \int_0^{\infty} \ln \sigma_i[S(j\omega)] d\omega = \pi \sum_{i=1}^{n_U} Re[p_i] \tag{5}$$

The difference compared to SISO systems is that the relationship involves the sum of the log magnitudes of the singular values, suggesting that it may be possible to trade off sensitivity properties in different directions. For MIMO systems with a RHP transmission zero at z, we have

$$\int_0^{\infty} \ln \bar{\sigma}[S(j\omega)] W(z,\omega) d\omega \geq 0 \tag{6}$$

Note that this result depends on the presence of RHP-zeros only, and does not show the combined effects of RHP poles and zeros similar to (4) for SISO systems. The reason results for MIMO systems are so much harder to find is the issue of plant directionality, as illustrated by the following simple examples

$$G_1(s) = f(s) \begin{bmatrix} 1 & 0 \\ 0 & \frac{s-z}{s-p} \end{bmatrix}$$

$$G_2(s) = f(s) \begin{bmatrix} s-z & 0 \\ 0 & \frac{1}{s-p} \end{bmatrix}$$

where $f(s)$ is an arbitrary stable minimum phase rational transfer function with at least three more poles than zeros. G_1 and G_2 both have one RHP transmission zero and one RHP pole at the same frequencies. However, in G_1 the RHP pole and transmission zero lie in the same direction and may pose a serious performance if p approaches z (recall (4)), whereas in G_2 the RHP pole and transmission zero lie in directions at right angles to each other and only the RHP zero itself poses a limitation.
Jacobsen and Skogestad [7] have studied the combined effect of RHP zeros (time delays) and RHP poles for distillation columns.

Implications of RHP zeros on the selection of controlled and manipulated variables. One should attempt to choose inputs and outputs such that RHP transmission zeros are avoided. If an RHP transmission zero cannot be avoided, it should preferably be at as high a frequency as possible, and lie in a plant direction such that it interferes as little as possible with the control of any RHP poles.
Similar considerations apply for decentralized control when a paired *element* g_{ii} has a RHP zero that is not a transmission zero of the plant. In many cases such RHP zeros will disappear when the other loops are closed, and there is then no fundamental bandwidth restriction in channel i. However, we know that the zeros of $1 + g_{ii}c_i$ will approach the zeros of g_{ii} for high gain feedback. A choice will therefore have to be made between individual loop stability and system performance. This dilemma can only be avoided if the selection of controlled and manipulated variables makes it possible to choose a pairing for which none of the paired elements have RHP zeros within the desired bandwidth.

2.3 Decentralized fixed modes.

Wang and Davison [22] showed that it may be impossible to move some modes by decentralized feedback, even if all states are both controllable from the inputs and observable from the outputs. Wang and Davison termed such modes "decentralized fixed modes". Lunze [11] gives a good explanation of a necessary and sufficient condition for the existence of decentralized fixed modes. Consider a system described by

$$\dot{x} = Ax + Bu; \quad y = Cx + Du$$

The simplest way to prove that an eigenvalue of A does not correspond to a decentralized fixed mode is to try with an arbitrary constant feedback matrix K with the structure of the pairing in consideration. A mode which is fixed for any constant decentralized feedback is also fixed for dynamic decentralized feedback [22]. When selecting input and outputs for decentralized control, one must clearly ensure that there exists at least one pairing for which all decentralized fixed modes are stable. This will usually not be difficult, as it suffices to ensure that the unstable state is both controllable and observable in one individual channel [11].
Example: Consider the plant

$$\dot{x} = \begin{bmatrix} -10 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -8 \end{bmatrix} x + \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} x \tag{7}$$

which yields

$$G(s) = \begin{bmatrix} \frac{1}{s+10} & \frac{2(s+4)}{(s+10)(s2)} \\ 0 & \frac{1}{s+8} \end{bmatrix} \quad (8)$$

We first try the pairing $y_1 - u_1$, $y_2 - u_2$, which by inspection of Eq. (8) seems to be a poor choice. This is confirmed by using $K = \text{diag}\{k_1, k_2\}$ and computing the closed loop autotransition matrix

$$A + BKC = \begin{bmatrix} -10 + k_1 & k_1 & k_2 \\ 0 & 2 & k_2 \\ 0 & 0 & -8 + k_2 \end{bmatrix} \quad (9)$$

The eigenvalue at 2 is unaffected by feedback, and it is therefore impossible to stabilize the system with this pairing. With the opposite pairing, $y_1 - u_2$, $y_2 - u_1$, we have no decentralized fixed modes, and the system can be stabilized by decentralized feedback.

2.4 The Niederlinski Index

The Niederlinski Index [18] is defined as

$$N_I = \frac{\det G(0)}{\det \tilde{G}(0)} \quad (10)$$

For *stable* plants it has been shown that if all the individual loops are stable and have integral action, a necessary condition for the stability of the overall system is that $N_I > 0$ [4]. This result also holds for unstable plants if we assume that the number of RHP poles in G and $\tilde{G}$ are the same, but this assumption generally does not hold. However, we present a generalized Niederlinski index criterion for unstable plants:

Theorem 2 *Assume:*
1. The transfer function GC is strictly proper.
2. The controller C is diagonal, has integral action in all channels and is otherwise stable.
3. The number of open loop unstable poles in G is n_U.
4. The map of the Nyquist D contour under $\det(I + \tilde{G}C) = \det\tilde{S}^{-1}$ *encircles the origin* $\tilde{n}_C$ *times in the clockwise direction.*
Then, a necessary condition for the overall system to be stable is

$$\text{sign}\{N_I\} = \text{sign}\{(-1)^{-n_U - \tilde{n}_C}\} \quad (11)$$

Remark: The individual loops (i.e. $\tilde{S}$) may or may not be stable. If we require the individual loops to be stable, then $\tilde{n}_C = -\tilde{n}_U$, where $\tilde{n}_U$ is the number of open loop unstable poles in $\tilde{G}$.

Proof: Let the map of the Nyquist D contour under $\det(I + E_H\tilde{H})$ encircle the origin n_E times in the clockwise direction. Thus, from Eq. (1) we have $n_C = \tilde{n}_C + n_E$ and we get from Thm. 1 that the overall system is stable if and only if $n_E = -n_U - \tilde{n}_C$. $E_H\tilde{H}$ is strictly proper, as GC is strictly proper. $\tilde{H}(0) = I$ because of the requirement for integral action in all channels, regardless of whether $\tilde{H}(s)$ is stable. We therefore have (see Corollary 1.1 in [5])

$$\lim_{s\to\infty} (I + E_H(s)\tilde{H}(s)) = I \quad (12)$$

$$\lim_{s\to 0} (I + E_H(s)\tilde{H}(s)) = \lim_{s\to 0} G(s)\tilde{G}^{-1}(s) \quad (13)$$

Then the map of $\det(I + E_H\tilde{H})$ starts at N_I (for $s = 0$) and ends at 1 (for $s = \infty$). For stability this map must have $n_E = -n_U - \tilde{n}_C$ encirclements of the origin, and we must require N_I to be positive if n_E is even and N_I to be negative if n_E is odd.

2.5 The Relative Gain Array

The Relative Gain Array (RGA) was first introduced by Bristol [1]. It is defined at any frequency as

$$\Lambda = G \times [G^{-1}]^T \quad (14)$$

where the sign $\times$ denotes element by element multiplication (Hadamard product). The ij'th element of Λ is denoted[1] λ_{ij}.
To simplify notation, we will in the following consider loop 1, without loss of generality (the generalization to loop k is trivial). Introduce $G' = \text{diag}\{g_{11}, G^{11}\}$, where G^{11} is obtained from G by removing row 1 and column 1. Let G' have n'_U RHP poles. Note that n'_U may be different for different loops.

Corollary 1 *Under assumptions 1-3 of Thm. 2, a necessary condition for simultaneously obtaining*
a) Stability of the closed loop system
b) Stability of loop 1 by itself
c) Stability of the system with loop 1 removed
is that

$$\text{sign}\{\lambda_{11}(0)\} = \text{sign}\{(-1)^{-n_U + n'_U}\} \quad (15)$$

Proof: Follows by substituting G' for $\tilde{G}$ in the proof of Thm. 2, and noting that $\det GG'^{-1} = 1/\lambda_{11}$. This corollary generalizes the widely used RGA pairing criteria ([1], [4]) to unstable plants.
For 2×2 plants $N_I = 1/\lambda_{11}(0)$ but for larger systems these measures contain different information.

2.6 The Direct Nyquist Array

The Direct Nyquist Array (DNA) [19] is simply an array of polar plots of the elements $g_{ij}(s)$ of the plant transfer function matrix $G(s)$. The usefulness of the DNA technique stems from Gershgorins theorem, which states that the eigenvalues λ_i of a matrix $G(s)$ must lie within the union of circles $|\lambda_i - g_{ii}(s)| < r_i(s)$ where $r_i(s) = \sum_{i=1, i\neq j}^{n} |g_{ij}(s)|$. The circles $r_i(s)$ are superimposed on the plots of the diagonal elements of G. As the frequency changes, the circles will move (and their radius will change), thus forming n bands of circles, known as *Gershgorin bands*. Jensen et al. [8] suggest to use the DNA also to obtain pairings for decentralized control, preferring pairings giving Gershgorin bands which are narrow compared to the magnitude of the diagonal elements.
The following derivation shows why a plot of the Gershgorin bands of $G(S)$ is useful for closed loop stability: Consider first the eigenvalues of $(I + GC)$, which are located within the Gershgorin bands of $(I + GC)$. Since $\det(I + GC) = \prod_i \lambda_i(I + GC)$ it follows from Thm. 1 that *if the number of encirclements of the Gershgorin bands of* $(I + GC)$ *equals the number of unstable poles,* n_U, *and none of these bands include the origin, the closed loop system is stable.* For decentralized control, the centers of the circles which make up the Gershgorin bands of $(I + GC)$ are given by $(I + \tilde{G}C)$. When G and $\tilde{G}$ have different numbers of RHP poles and the individual loops are stable, the numbers of encirclement of the origin of the Nyquist D contours under $\det(I + GC)$ and $\det(I + \tilde{G}C)$ must differ. Since the centers of the Gershgorin bands of $I + GC$ are given by $(I + \tilde{G}C)$, some of the Gershgorin bands of $(I + GC)$ *must include* the origin for the overall system to be stable. Furthermore, it will then be undesirable that the Gershgorin bands of G are narrow, as this will make the stability margins for the individual loops and the overall system small.

[1] The ij'th relative gain λ_{ij} should not be confused with λ_i which denotes the i'th eigenvalue.

To derive the relationship between the Gershgorin bands of $(I + GC)$ and G normalize the elements in G with the diagonal elements, and get

$$G = \begin{bmatrix} 1 & \frac{g_{12}}{g_{22}} & \frac{g_{13}}{g_{33}} & \cdots \\ \frac{g_{21}}{g_{11}} & 1 & \frac{g_{23}}{g_{33}} & \cdots \\ \vdots & & \ddots & \vdots \end{bmatrix} \operatorname{diag}\{g_{ii}\} \quad (16)$$

This should be compared to

$$(I + GC) = (I + E_H\tilde{H})(I + \tilde{G}C) = \quad (17)$$

$$\begin{bmatrix} 1 & \tilde{h}_2\frac{g_{12}}{g_{22}} & \tilde{h}_3\frac{g_{13}}{g_{33}} & \cdots \\ \tilde{h}_1\frac{g_{21}}{g_{11}} & 1 & \tilde{h}_3\frac{g_{23}}{g_{33}} & \cdots \\ \vdots & & \ddots & \vdots \end{bmatrix} \operatorname{diag}\{1 + g_{ii}c_i\}$$

We see that the width of the Gershgorin bands (relative to the magnitude of the diagonal elements) will be the same for G and $(I + GC)$ at frequencies below the bandwidth (where $\tilde{h}_i \approx 1$). At frequencies beyond the bandwidth (where $\tilde{h}_i < 1$) $(I + GC)$ will have narrower Gershgorin bands than G. Only in the bandwidth region, where peaks in $\tilde{h}_i$ may occur, can the Gershgorin bands of $(I+GC)$ be wider than the Gershgorin bands of G. Thus, the widths of the Gershgorin bands of G (relative to the magnitude of the diagonal elements) can be used as estimates of the widths of the Gershgorin bands of $(I+GC)$. In this paper, we use the diagonal similarity transform of Mees [14] in order to reduce the conservativeness associated with the location of the eigenvalues of G. The diagonal elements of this transformation matrix are the elements of the left eigenvalue corresponding to the Perron root of $G\tilde{G}^{-1}$.

2.7 The SSV Interaction Measure.

The Structured Singular Value Interaction Measure (SSV-IM)[5] is the structured singular value (μ) of E_H with respect to a stable perturbation matrix with the same structure as $\tilde{H}$. The map of $\det(I + E_H\tilde{H})$ cannot encircle the origin if $\rho(E_H\tilde{H}) < 1 \quad \forall\omega$, which is satisfied if $\bar{\sigma}(\tilde{H})\mu(E_H) < 1 \quad \forall\omega$. The last relationship follows since the least conservative way to "split up" $\rho(E_H\tilde{H})$ is to use the structured singular value. Since these relationships are useful only when we require that $\det(I+E_H\tilde{H})$ should *not* encircle the origin (i.e. $-n_U - \tilde{n}_C = 0$), we see that they generally apply only when the individual loops are stable and G and $\tilde{G}$ have the same number of RHP poles.

2.8 The Performance RGA

The Performance Relative Gain Array (PRGA) has recently been introduced [21, 6]. It is derived from simple manipulations with the transfer function from setpoints r to offset $e = y - r = -Sr$: At low frequencies ($\omega < \omega_B$) where $\tilde{H} \approx I$, we have

$$(I + E_H\tilde{H}) \approx I + E_H = G\tilde{G}^{-1} = \Gamma^{-1}; \quad \omega < \omega_B$$

and we derive from Eq. (1) that

$$e \approx -\tilde{S}\Gamma r; \qquad \omega < \omega_B \quad (18)$$

where the PRGA matrix is

$$\Gamma = \tilde{G}G^{-1} \quad (19)$$

Despite their different use this definition points to the similarity between the PRGA and the Inverse Nyquist Array. The PRGA is dependent on scaling of the outputs, and must be recomputed for new pairings. Note, however, that the diagonal elements of the PRGA matrix are equal to the diagonal elements of the RGA matrix, and are hence independent of scaling. If G is scaled such that the maximum acceptable magnitude of the individual offsets e_i is unity, the PRGA matrix gives approximate bandwidth requirements and loop gain requirements at frequencies below the bandwidth, as the loop gain in loop i should be larger in magnitude than the magnitude of any element in row i of the PRGA matrix, and small PRGA elements are therefore preferred. The use and the limitations of the PRGA is discussed more thoroughly in [6].

2.9 High frequency RGA and stability

The encirclements of the origin of $\det(I + GC)$ caused by the RHP poles of G, will occur in the frequency region corresponding to the RHP poles of G. For practical systems the bandwidth region will lie at frequencies significantly higher than the RHP poles of G, and we therefore do not want any additional encirclements in the bandwidth region. The poles of $\tilde{G}$ lie at the same locations as the poles of G, and we therefore want $\det(I + \tilde{G}C)$ also to avoid encirclements of the origin in the bandwidth region. We thus want $\det(I + GC)$ and $\det(I + \tilde{G}C)$ to behave similarly in the bandwidth region. If G is triangular then $\det(I+E_H\tilde{H}) = 1$ and $\det(I+GC) = \det(I+\tilde{G}C)$, and in this case we also have that the RGA matrix $\Lambda = I$ and the PRGA matrix Γ is triangular with diagonal elements equal to 1. We therefore prefer pairings which give $\Lambda \approx I$ in the bandwidth region. This agrees with the results of Nett and coworkers ([17]). In summary, we want the PRGA elements to be small at low frequencies, and in the bandwidth region we desire the PRGA close to triangular with diagonal elements close to 1.

3 Examples.

3.1 Example 1.

Consider the system $G(s) = C(sI - A)^{-1}B + D$, with

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix}; \quad B = \begin{bmatrix} 5 & -8 \\ 4 & 10 \\ 2 & -8 \end{bmatrix}$$

$$C = \begin{bmatrix} -1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}; \quad D = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad (20)$$

The transfer function matrix G has <u>one</u> unstable pole at frequency 1 [rad/min], and we must therefore expect the closed loop bandwidth to be at least 1 [rad/min]. Note that $\tilde{G}$ has <u>two</u> unstable poles for both pairings, because the RHP pole appears in all four elements of G. Neither $G(s)$ nor any of its elements have RHP zeros. $G(0)$ is given by

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & -18 \\ -6 & 12 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \quad (21)$$

The pairing ($y_1 - u_1$, $y_2 - u_2$) indicated by Eq. (21) we term "pairing 1", and the opposite pairing we term "pairing 2". The Niederlinski Index is -8 for pairing 1 and 0.89 for pairing 2. Thm. 2 tells us that a negative Niederlinski Index is necessary if we require both loops in addition to the overall system being stable and we therefore have to choose pairing 1. If we use the steady state RGA in accordance with Cor. 1, we would arrive at the same conclusion.

The DNA of pairing 1 and pairing 2 are shown in Fig. 1, with Gershgorin bands superimposed on the 1,1 and 2,2 elements for pairing 1, and on elements 1,2 and 2,1 for pairing 2. We see that all the Gershgorin bands include the origin. However, pairing 1 gives much wider Gershgorin bands relative to the magnitude of the diagonal elements than pairing 2. We explained above that when G and $\tilde{G}$ have different number of RHP poles and stability of the individual loops is desired, narrow Gershgorin bands are undesirable. The DNA thus gives a weak indication that pairing 1 is preferable to pairing 2.

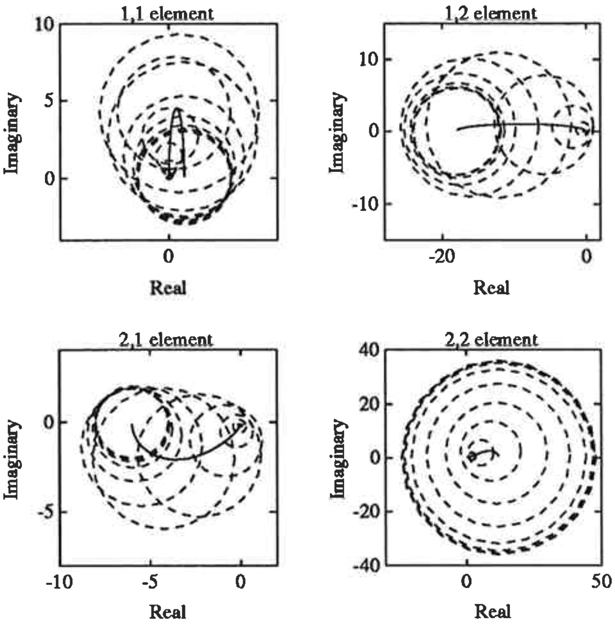

Figure 1: DNA with Gershgorin bands for Example 1.

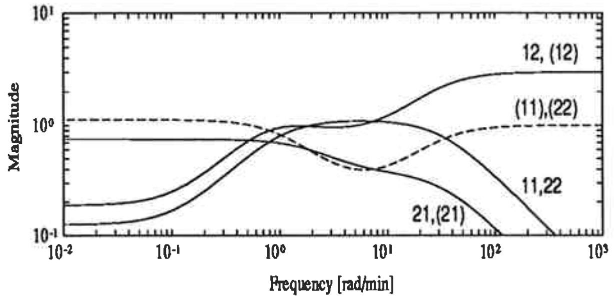

Figure 2: PRGA for Example 1. Solid lines denote pairing 1, and dashed lines and labels in parenthesis denote pairing 2

The PRGA for pairing 1 and pairing 2 are shown in Fig. 2[2]. With a closed loop bandwidth in the region $1-30$ [rad/min] (approximately) the PRGA indicate that pairing 1 is preferable (The PRGA is close to triangular and the RGA$\approx I$). For pairing 1, using the controllers $c_1(s) = -\frac{s+1}{s}$, $c_2(s) = -\frac{(s+1)(0.1s+1)}{s(0.01s+1)}$ we find that both loops and the overall system is stable, and the predictions using the Niederlinski Index and the PRGA are shown to hold. This example also demonstrates that for unstable systems there is no reason to avoid pairings corresponding to negative steady state RGA values.

3.2 Example 2: Polypropylene reactor.

This example illustrates the problems encountered when we are not able to stabilize the individual loops. We consider the polypropylene reactor control example studied by Lie [9, 10]. A schematic outline of the process is shown in Fig. 3. The monomer feed enters into a stirred tank reactor containing a slurry of monomer, catalyst, cocatalyst, polymer and some impurities. The reaction is exothermic, causing some of the slurry components to vaporize. The vapor leaving the reactor is transferred to an accumulator

[2]We assume the outputs in Eq. (20) to be properly scaled. For the PRGA plot for pairing 2, $G(s)$ has been rearranged to bring the paired elements on the diagonal.

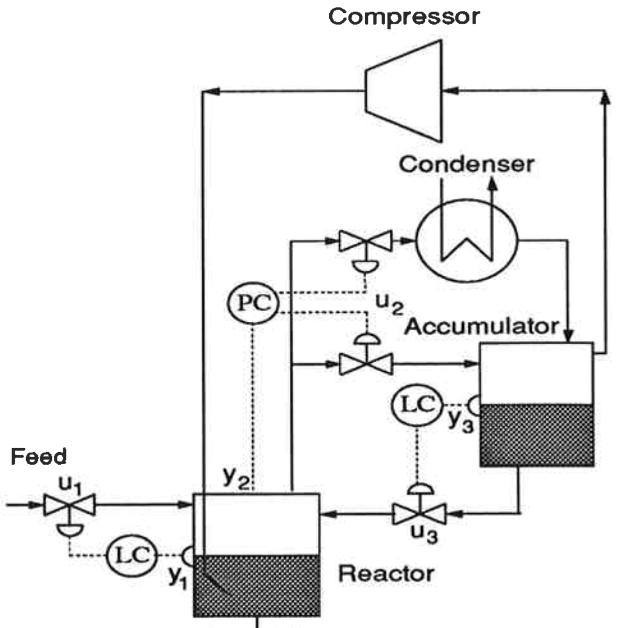

Figure 3: Schematic outline of the process in Example 2, with pairing 1 selected for control.

vessel. Heat is removed from the system by condensing parts of the vapor leaving the reactor, before it enters the accumulator. Heat removal is adjusted by adjusting a split range valve which determines what fraction of the vapor leaving the reactor is passed through the condenser. The liquid in the accumulator is returned to the reactor, and the vapor from the accumulator is compressed and bubbled through the reactor slurry. This results in a 3×3 plant model $G(s)$ with seven states as given in [9, 10]. The inputs and outputs are

y_1 - reactor slurry level $(0-1)$
y_2 - reactor pressure (gauge pressure in atmospheres)
y_3 - accumulator liquid level $(0-1)$.
u_1 - monomer feed flowrate (kg/h).
u_2 - split range valve position $(0-1)$.
u_3 - accumulator to reactor liquid flowrate (kg/h).

We scale the outputs such that a magnitude of 1 for the scaled outputs correspond to deviations: $y_1 = 0.05$, $y_2 = 1.0$atm and $y_3 = 0.10$. The plant G has a pair of RHP poles at $s = 0.685 \pm 0.688j$, one pure integrator (the accumulator level) and no RHP transmission zeros. The reactor holdup becomes unstable because of the overhead condensation loop. All elements of G except g_{11} have RHP zeros at frequencies close to the RHP poles (e.g. g_{12} has a RHP zero at 2.16, g_{22} have a pair of complex RHP zeros at $0.10 \pm 0.90j$).

As we are considering the use of low-order controllers (typically PI or PID controllers), we cannot expect loops with a RHP zero at frequencies around or lower than the frequency of the RHP pole to be stable. This makes it difficult to apply the Niederlinski Index (Thm. 2) or the steady state RGA (Cor. 1) for this example. Although six different pairings is possible for a 3×3 system, we will here only consider the two pairings that pair y_1 with u_1.[3] There is a strong incentive for pairing y_1 and u_1; as g_{11} has no RHP zero we are able to stabilize the system with only this loop closed. Readers are referred to [9] for a more thorough discussion of all possible pairings. Thus, for this example, the

[3]For this example the steady state RGA is positive and close to I for another pairing, $y_1 - u_2$, $y_2 - u_3$, $y_3 - u_1$. This pairing defies common sense. For example, it means controlling the accumulator level with the monomer feed flowrate to the reactor. However, the desired bandwidth is about 10rad/h, and in this frequency range the RGA indicates that pairing 1 is preferable.

pairing $y_1 - u_1$, $y_2 - u_2$, $y_3 - u_3$ is termed pairing 1, and the pairing $y_1 - u_1$, $y_2 - u_3$, $y_3 - u_2$ is termed pairing 2. Pairing 1 is shown in Fig. 3. The RGA plot is not shown here,

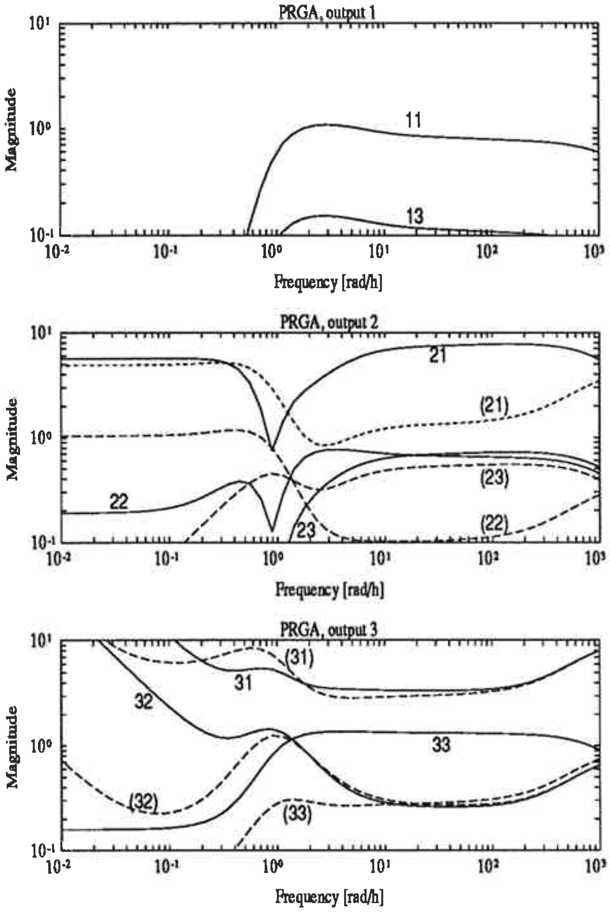

Figure 4: PRGA for Example 2. Solid lines denote pairing 1, and dashed lines and labels in parenthesis denote pairing 2.

but the diagonal elements equal the diagonal elements of the PRGA shown in Fig. 4. The figure clearly shows that pairing 1 is preferable, as the PRGA values for the paired elements are closer to one in the bandwidth region for pairing 1 than for pairing 2[4]. The PRGA also shows that for both pairings, the interaction leads to increased loop gain requirements at low frequencies for loops 2 and 3.

For this problem, we can only expect loop 1 to be stable by itself, the map of the Nyquist D contour under $1+g_{11}c_1$ will therefore encircle the origin twice in the anticlockwise direction. For pairing 1, g_{22} and g_{33} both have two RHP poles and two RHP zeros within the desired bandwidth. The two RHP zeros will most likely result in two unstable poles of $1+g_{22}c_2$ and $1+g_{33}c_3$. $1+g_{22}c_2$ and $1+g_{33}c_3$ must therefore be expected to add no net encirclements of the origin, and therefore we would prefer narrow Gershgorin bands for pairing 1. For pairing 2, g_{23} and g_{32} both have two RHP poles and one RHP zero within the desired bandwidth. $1+g_{23}c_2$ and $1+g_{32}c_3$ must therefore be expected to each add one anticlockwise encirclement of the origin. For pairing 2, narrow Gershgorin bands are therefore not desirable. The DNA plot is not shown for this example due to space limitations, but the Gershgorin bands are very wide for both pairings. The DNA is therefore inconclusive for both pairings.

Based on the plots of the frequency-dependent RGA, we choose pairing 1. This conclusion is consistent with the simulations of Lie [10] and with industrial practice.

[4]For pairing 2 $G(s)$ has been rearranged to bring the paired elements for pairing 2 on the diagonal.

References

[1] Bristol, E. H., (1966). On a New Measure of Interaction for Multivariable Process Control, *IEEE Trans. Autom. Control*, **AC-11**, 133-134.

[2] Freudenberg, J. S. and Looze, D. P. (1985). Right Half Plane Poles and Zeroes and Design Tradeoffs in Feedback Systems. *IEEE Trans. Autom. Control*, **AC-30**, 6, 555-565.

[3] Freudenberg, J. S. and Looze, D. P. (1988). *Frequency Domain Properties of Scalar and Multivariable Feedback Systems.* Lecture Notes in Control and Information Sciences 104, Springer Verlag, Berlin, Germany.

[4] Grosdidier, P., Morari, M. and Holt, B. R. (1985). Closed-Loop Properties from Steady-State Gain Information. *Ind. Eng. Chem. Fundam.*, **24**, 221-235.

[5] Grosdidier, P. and Morari, M. (1986). Interaction Measures for Systems Under Decentralized Control. *Automatica*, **22**, 3, 309-319.

[6] Hovd, M. and Skogestad, S. (1992). Simple Frequency-Dependent Tools for Control System Analysis, Structure Selection and Design. To appear in *Automatica.*

[7] Jacobsen, E. W. and Skogestad, S. (1991). Control of Unstable Distillation Columns. *Proc. American Control Conference*, Boston, Mass., 773-778.

[8] Jensen, N., Fisher, D. G. and Shah, S. L. (1986). Interaction Analysis in Multivariable Control Systems. *AIChE Journal*, **32**, 6, 959-970.

[9] Lie, B. (1990). Control Structures for Polymerization Processes Applied to Polypropylene Manufacturing. *Dr. Ing. Thesis*, University of Trondheim-NTH.

[10] Lie, B. and Balchen, J. G. (1992). A Comparison of Strategies for Control of a Polypropylene Reactor, *Preprints IFAC Symposium DYCORD+'92*, College Park, Maryland, Apr. 1992, 265-270.

[11] Lunze, J. (1992). *Feedback Control of Large-Scale Systems*, Prentice-Hall, New York.

[12] Maciejowski, J. M. (1989). *Multivariable Feedback Design.* Addison-Wesley, Wokingham, England.

[13] McAvoy, T. J., (1983). *Interaction Analysis*, ISA Monograph, Research Triangle Park, North Carolina.

[14] Mees, A. J. (1981). Achieving Diagonal Dominance. *Systems & Control Letters*, **1**, 3, 155-158.

[15] Morari, M., Grimm, W., Oglesby, M.J. and Prosser, I.D. (1985). Design of Resilient Processing Plants VII. Design of Energy Management System for Unstable Reactors - New Insights, *Chem. Eng. Sci.*, **40**, 187-198.

[16] Morari, M. and Zafiriou, E., (1989). *Robust Process Control*, Prentice-Hall, Englewood Cliffs, New Jersey.

[17] Minto, K.D. and Nett, C.N. (1989). A Quantitative Approach to the Selection and Partitioning of Measurements and Manipulators for the Control of Complex Systems. Presentation at *1989 American Control Conference*, Minneapolis.

[18] Niederlinski, A. (1971). A Heuristic Approach to the Design of Linear Multivariable Interacting Control Systems. *Automatica*, **7**, 691-701.

[19] Rosenbrock, H. H. (1970). *State-Space and Multivariable Theory.* Nelson, London, England.

[20] Skogestad, S. and Morari, M. (1988). Variable Selection for Decentralized Control. *AIChE Annual Meeting*, Washington DC, paper 128c. Also to appear in *Modelling, Identification and Control* (1992).

[21] Skogestad, S., and Hovd, M. (1990). Use of Frequency-Dependent RGA for Control Structure Selection. *Proc. American Control Conference*, San Diego, California, 2133-2139.

[22] Wang, S.-H. and Davison, E. J. (1973). On the Stabilization of Decentralized Control Systems. *IEEE Trans. Autom. Control*, **AC-18**, 5, 473-478.