Robust Pole Assignment in Linear State Feedback

J. Kautsky*, N. K. Nichols† and P. Van Dooren#

Numerical Analysis Report 5/84

* School of Mathematical Sciences,
Flinders University of
South Australia.

Philips Research Laboratory,
Brussels, Belgium.

Abstract

Numerical methods are described for determining robust, or well-conditioned, solutions to the problem of pole assignment by state feedback. The solutions obtained are such that the sensitivity of the assigned poles to perturbations in the system and gain matrices is minimized. It is shown that for these solutions, upper bounds on the norm of the feedback matrix and on the transient response are also minimized and a lower bound on the stability margin is maximized. A measure is derived which indicates the optimal conditioning that may be expected for a particular system with a given set of closed loop poles, and hence the suitability of the given poles for assignment.

Key words: automatic control, multivariable system, pole assignment, state feedback, numerical algorithm, inverse eigenvalue problem.

1.0 INTRODUCTION

The state feedback pole assignment problem in control system design is essentially an inverse eigenvalue problem. The solution is, in general, under-determined, with many degrees of freedom. A desirable property of any system design is that the poles should be insensitive to perturbations in the coefficient matrices of the system equations. This criterion may be used to restrict the degrees of freedom in the assignment problem, and to produce a well conditioned or robust solution to the inverse eigenproblem.

A number of constructive methods for pole assignment by state feedback are described in the literature [1], [4] - [6], [13]-[18], [20] - [22], [27], [29], but many of these are computationally unstable (see [8]). A few numerically reliable techniques are available [17], [18], [27], and in the single input case, where at most one solution to the pole assignment problem exists, these methods accurately compute the required feedback. In the multi-input case, however, where the feedback is under-determined, these methods do not generally lead to robust solutions to the problem.

In this paper we describe four algorithms for computing robust solutions to the multi-input state-feedback pole assignment problem. Two of the methods are complementary. In all cases the feedback matrix is obtained by assigning linearly independent eigenvectors corresponding to the required eigenvalues (or poles), such that the matrix of eigenvectors is as well-conditioned as possible [28]. The assigned poles are then as insensitive to perturbations as possible and the resulting feedback matrix is as reasonably bounded as may be expected, given the original system.

In the next section the pole assignment problem is defined in detail, and theoretical considerations are discussed. In section 3 we describe the numerical algorithms. Applications and numerical results are presented in section 4, and concluding remarks follow in section 5.

2.0 ROBUST POLE ASSIGNMENT

2.1 The Robust State Feedback Problem

We consider the time-invariant, linear, multivariable system with dynamic state equation

$$\mathcal{D}\,\underline{x} = A\underline{x} + B\underline{u}\ , \tag{1}$$

where $\underline{x}$, $\underline{u}$ are n- and m- dimensional vectors, respectively, and A, B are real, constant matrices of compatible orders. Here $\mathcal{D}$ denotes the differential operator d/dt for continuous time systems, or the delay operator for discrete time systems. Matrix B is assumed (without loss of generality) to be of full rank. The behaviour of system (1) is governed by the poles of the system, that is, by the eigenvalues of matrix A. It is often desirable to modify the poles of the system in order to obtain certain properties, such as stability. This may be achieved by using a state-feedback control

$$\underline{u} = F\underline{x} + \underline{v}\ ,$$

where F, the feedback or gain matrix, is chosen such that the modified dynamic system

$$\mathcal{D}\,\underline{x} = (A+BF)\,\underline{x} + B\underline{v}\ , \tag{2}$$

now with input $\underline{v}$, has the desired poles.

The state-feedback pole assignment problem for system (1) is formulated precisely as follows.

Problem 1. Given real matrices (A,B), of orders (nxn, nxm) respectively, and a set of n complex numbers, $\mathcal{L} = \{\lambda_1,\lambda_2,..\lambda_n\}$, closed under complex conjugation, find a real mxn matrix F such that the eigenvalues of A + BF are λ_j, j = 1,2,..n. □

Conditions for the existence of solutions to Problem 1 are well-known and the following theorem is well-established [28].

Theorem 1. A solution F to Problem 1 exists for every set $\mathcal{L}$ of self-conjugate complex numbers if and only if the pair (A,B) is completely controllable, that is, if and only if:

$$\{\underline{s}^T A = \mu \underline{s}^T \text{ and } \underline{s}^T B = \underline{0}\} \iff \underline{s}^T = \underline{0} \ . \qquad \square$$

Indeed, if (A,B) is not controllable, i.e. there exists $\underline{s}^T \neq 0$ such that $\underline{s}^T A = \mu \underline{s}^T$ and $\underline{s}^T B = 0$, then $\underline{s}^T (A+BF) = \mu \underline{s}^T$ for all F. Thus μ is an eigenvalue of $A+BF$ for all F and it cannot be modified by any feedback control. The pole μ is said to be uncontrollable and must belong to any set $\mathcal{L}$ of poles to be assigned, if a solution to Problem 1 is to exist. More specifically, the following theorem can be shown to hold (see [29]):

Theorem 2. If (A,B) is not completely controllable, then a solution F to Problem 1 exists if and only if the set $\mathcal{L} = \{\mathcal{L}_u, \mathcal{L}_c\}$, contains $\mathcal{L}_u$, the set of all uncontrollable modes of (A,B). $\square$

In the single input case $(m=1)$, the solution to Problem 1, when it exists, can be shown to be unique [16]. In the case $1 < m < n$, various solutions to Problem 1 may exist, and, to determine a specific solution, additional conditions must be supplied in order to eliminate the extra degrees of freedom. In the case $m = n$, the pair (A,B) is always completely controllable, and any given closed loop system matrix can always be achieved by feedback.

Our aim here is to develop methods for finding a feedback F, solving Problem 1, such that the closed loop system is robust, in the sense that its poles are as insensitive to perturbations as possible. We let $\underline{x}_j$ and $\underline{y}_j$, $j = 1,2,\ldots n$, be the right and left eigenvectors of the closed loop system matrix $M \equiv A+BF$, corresponding to eigenvalue $\lambda_j \in \mathcal{L}$, that is

$$M\underline{x}_j = \lambda_j \underline{x}_j \ , \quad \underline{y}_j^T M = \lambda_j \underline{y}_j^T \ . \qquad (3)$$

If M is non-defective, that is, M has n linearly independent eigenvectors, then M is diagonalizable and it can be shown [28] that the sensitivity of the eigenvalue λ_j to perturbations in the components of A, B and F depends upon the magnitude of the condition number c_j, where

$$c_j = 1/s_j = \|\underline{y}_j\|_2 \, \|\underline{x}_j\|_2 \; / \; |\underline{y}_j^T \underline{x}_j| \geq 1 \; . \tag{4}$$

In the case of multiple eigenvalues, a particular choice of eigenvectors is assumed. (For real λ_j, the sensitivity s_j is just the cosine of the angle between the right and left eigenvectors corresponding to λ_j.) More precisely, if a perturbation $O(\varepsilon)$ is made in the coefficients of the matrix M, then the corresponding first order perturbation in the eigenvalue λ_j of M is of the order of $\varepsilon n c_j$. If M is defective, then the corresponding perturbation in some eigenvalue is at least an order of magnitude worse in ε, and therefore, system matrices which are defective are necessarily less robust than those which are non-defective.

We observe that a bound on the sensitivities of the eigenvalues is given by [28]

$$\max_j c_j \leq \kappa_2(X) \equiv \|X\|_2 \, \|X^{-1}\|_2 \; , \tag{5}$$

where $\kappa_2(X)$ is the condition number of the matrix $X = [\underline{x}_1, \underline{x}_2, \ldots \underline{x}_n]$ of eigenvectors. Furthermore, the condition numbers take minimum value $c_j = 1$, for all $j = 1,2,\ldots n$, if and only if M is a normal matrix, that is $M^*M = MM^*$. In this case the eigenvectors of M may be scaled to give an orthonormal basis for $\mathbb{C}^n$, and then matrix X is perfectly conditioned with $\kappa_2(X) = 1$.

We may now formulate the robust pole assignment problem as follows:

Problem 2. Given (A,B) and $\mathcal{L}$ (as in Problem 1), find real matrix F and non-singular matrix X satisfying

$$(A+BF)X = X\Lambda \ , \tag{6}$$

where $\Lambda = \mathrm{diag}\{\lambda_1,\lambda_2,..\lambda_n\}$, such that some measure ν of the conditioning, or robustness, of the eigenproblem is optimized.

We remark that the measure ν could, for example, be chosen to be $\nu_1 = \| \underline{c} \|_\infty$, where $\underline{c}^T = [c_1,c_2,..c_n]$ is the vector of condition numbers corresponding to the selected matrix X of eigenvectors. Alternatively, we could take as a measure of robustness $\nu_2 = \kappa_2(X)$, the condition number of matrix X. The measure ν_2 then gives an upper bound on the measure ν_1, and both measures attain their (common) minimum value simultaneously. Other measures are discussed in §2.5.

The degrees of freedom available in the choice of the feedback F are reflected precisely by the degrees of freedom available in the selection of the matrix X of eigenvectors. In the case $m = 1$, if F exists, X is uniquely determined (up to scaling), and the condition numbers c_j cannot be controlled. In the case $m = n$, X may always be chosen to be orthogonal, ($X \equiv I$ suffices), and hence to be such that $c_j = 1, \forall_j$. For a general multi-input system $(1 < m < n)$ we may control the sensitivities of the assigned poles to a restricted extent by an appropriate choice of the eigenvectors comprising X.

We observe that in the robust pole placement problem (Problem 2) the choice of eigenvectors which may be assigned is restricted such that the resulting system matrix $A+BF$ is non-defective. This restriction implies certain simple conditions on the multiplicity of the poles which may be assigned. In the next section (§2.2) we discuss assignment of the entire eigenstructure of the closed loop system, and in the following section we show that minimizing the conditioning of the eigenproblem (6) leads to other desirable properties in the closed loop control system. Results on the minimal conditioning that can be achieved

for a given set of poles $\mathcal{L} = \{\lambda_1, \lambda_2, \ldots \lambda_n\}$ to be assigned are given in §2.4, and various robustness measures ν, to be used in practice, are discussed in §2.5. Such measures relate to different numerical methods for determining the feedback matrix, presented in §3.

2.2 Robust Eigenstructure Assignment

Given real matrix pair (A,B) and eigenvalue set $\mathcal{L}$, our objective is to choose eigenvectors, given by X, satisfying (6) and such that the conditioning of the eigenproblem is minimized. No restriction on the controllability of (A,B) is made, and we remark that although the uncontrollable modes of the system cannot be affected by the feedback F, the corresponding eigenvectors may be modified and the conditioning of uncontrollable modes may be improved by an appropriate choice of X.

It is reasonable now to ask under what conditions a given non-singular matrix X can be assigned to the problem. The following theorem is easily demonstrated.

Theorem 3. Given $\Lambda = \text{diag}\{\lambda_1,\lambda_2,\ldots\lambda_n\}$ and X non-singular, then there exists F, a solution to (5) if and only if

$$U_1^T(AX - X\Lambda) = 0 \tag{7}$$

where

$$B = [U_0,U_1]\begin{bmatrix} Z \\ 0 \end{bmatrix}, \tag{8}$$

with $U = [U_0,U_1]$ orthogonal and Z non-singular. Then F is given explicitly by

$$F = Z^{-1}\, U_0^T\, (X\Lambda X^{-1} - A). \tag{9}$$

Proof: The assumption that B is of full rank implies the existence of decomposition (8). From (6), F must satisfy

$$BF = X\Lambda X^{-1} - A\ , \tag{10}$$

and pre-multiplication by U^T then gives the two equations

$$\begin{aligned} ZF &= U_0^T\, (X\Lambda X^{-1} - A)\ , \\ 0 &= U_1^T\, (X\Lambda X^{-1} - A)\ , \end{aligned} \tag{11}$$

from which (7) and (9) follow directly, since X is invertible. □

We remark that (10) implies that F exists if and only if

$$R\{X\Lambda X^{-1} - A\} \subset R\{B\} \equiv R\{U_0\} , \tag{12}$$

where $R\{\cdot\}$ denotes range; that is, $R\{X\Lambda X^{-1} - A\}$ is orthogonal to $N\{B\} \equiv R\{U_1\}$, where $N\{\cdot\}$ denotes null space, and (7) holds.

We observe also that the decomposition (8) of B can be taken, for example, as the singular value (SVD) decomposition, in which case $Z = \Sigma V^T$, where $\Sigma = \text{diag}\{\sigma_1,\sigma_2,\ldots\sigma_m\}$ is a positive matrix and V is orthogonal. Alternatively the QR decomposition could be used, in which case Z is an upper triangular matrix.

An immediate consequence of Theorem 2 is the following:

Corollary 1. The eigenvector $\underline{x}_j$ of $A+BF$ corresponding to the assigned eigenvalue $\lambda_j \in \mathcal{L}$ must belong to the space

$$S_j \equiv N\{U_1^T(A - \lambda_j I)\} . \tag{13}$$

The dimension of S_j is given by

$$\dim(S_j) = m + k_{\lambda_j} , \tag{14}$$

where

$$k_{\lambda_j} = \dim(N\{[B|A - \lambda_j I]^T\}). \tag{15}$$

Proof: From (7) we have directly that

$$U_1^T(A\underline{x}_j - \lambda_j\underline{x}_j) = 0 , \qquad \forall j , \tag{16}$$

and, therefore, $\underline{x}_j \in S_j$, $j = 1,2,\ldots n$, is necessary. We have also, from (15), that

$$n - k_{\lambda_j} = \text{rank}([B|A - \lambda_j I]), \tag{17}$$

and from (7) and (8), using the definition $U = [U_0, U_1]$, we obtain

$$U^T [B|A - \lambda_j I] = \left[\begin{array}{c|c} Z & U_0^T(A - \lambda_j I) \\ \hline 0 & U_1^T(A - \lambda_j I) \end{array}\right] . \qquad (18)$$

But, matrix Z is square $(m \times m)$ and invertible, and, therefore,

$$n - m - k_{\lambda_j} = \text{rank}(U_1^T(A - \lambda_j I)) , \qquad (19)$$

from which (14) readily follows. □

The robust pole assignment problem (Problem 2) now reduces to the problem of selecting independent vectors $\underline{x}_j \in S_j$, $j = 1,2,\ldots n$, such that eigenproblem (6) is as well conditioned as possible.

From the corollary we may now deduce certain conditions which must be satisfied if the robust Problem 2 is to have a solution. In the case (A,B) is completely controllable, the dimension k_λ, defined by (15), is zero for all λ. It follows then that the multiplicity of the eigenvalues $\lambda_j \in \mathcal{L}$ to be assigned must be of multiplicity less than or equal to m, since the maximum number of independent eigenvectors which can be chosen to correspond to λ_j is equal to $\dim(S_j) = m$.

In the case (A,B) is not completely controllable, but $\mathcal{L} = \{\mathcal{L}_u, \mathcal{L}_c\}$ includes the uncontrollable modes, then a similar result holds. Now if $\lambda_j \in \mathcal{L}_u$ is uncontrollable, then $k_{\lambda_j} > 0$ and there exist at least k_{λ_j} independent (left) eigenvectors $\underline{y}_i$, $i = 1,2,\ldots k_{\lambda_j}$ of the closed loop system matrix $M \equiv A + BF$ for any choice of F. The eigenvalue λ_j must, therefore, be assigned with multiplicity at least k_{λ_j}, and for the closed loop system matrix to be non-defective, the eigenvalue λ_j can be assigned with multiplicity at most $m + k_{\lambda_j} = \dim(S_j)$.

Even when these conditions are satisfied, it is still possible that a solution to Problem 2 does not exist. Indeed, if μ is an uncontrollable mode of the pair (A,B) with $k_\mu > 0$, and there exists a vector $\underline{s} \neq 0$ such that $\underline{s}^T(A - \mu I)B = 0$, $\underline{s}^T(A - \mu I)^2 = 0$, and $\underline{s}^T(A - \mu I) \equiv \underline{s}_1^T \neq 0$, then $\underline{s}_1^T$ is a left eigenvector of $M \equiv A + BF$, corresponding to μ, for any F. If also $\underline{s}^T B = 0$, then

$$\underline{s}^T(A + BF - \mu I) = \underline{s}^T(A - \mu I) \neq 0 , \tag{20}$$

$$\underline{s}^T(A + BF - \mu I)^2 = \underline{s}^T((A - \mu I)^2 + (A - \mu I)BF + BF(A + BF - \mu I)) = 0,$$

and, therefore, $\underline{s}^T$ is always a left principal vector of M, and the system matrix $A + BF$ is defective for any feedback F. We have thus proved the following theorem.

Theorem 4. A necessary condition for the existence of a non-defective solution to the pole assignment problem (Problem 1) is that for every μ and $\underline{s}$ the following holds:

$$\underline{s}^T B = 0,\ \underline{s}^T(A - \mu I)B = 0,\ \underline{s}^T(A - \mu I)^2 = 0 \Rightarrow \underline{s}^T(A - \mu I) = 0 . \tag{21}$$

□

We note that condition (21) is satisfied trivially for all controllable modes μ of (A,B).

2.3 Properties of the Robust Closed Loop System

The objective of the robust pole placement problem is, in essence, to choose a non-defective system of eigenvectors, given by X, satisfying Theorem 3, such that X is as well conditioned as possible. We show now that minimizing the conditioning of X leads to other desirable properties of the closed loop system.

From Theorem 3 we derive bounds on the components of the feedback matrix F and the transient response $\underline{x}(t)$ (or $\underline{x}(k)$) of the closed loop system (2) in terms of the condition number $\kappa_2(X)$ and the given data of the problem. We have

Theorem 5. The gain matrix F and the transient response $\underline{x}(t)$, or $\underline{x}(k)$, of the closed-loop continuous, or discrete, time system (2), where $\underline{x}(0) = \underline{x}_0$ and $\underline{v} \equiv 0$, satisfy the inequalities

$$\| F \|_2 \leq (\| A \|_2 + \max_j\{|\lambda_j|\} \cdot \kappa_2(X))/\sigma_m\{B\} , \tag{22}$$

where $\sigma_m\{B\}$ denotes the m-th (smallest) singular value of B, and

$$\| \underline{x}(t) \|_2 \leq \kappa_2(X) \cdot \max_j\{|e^{\lambda_j t}|\} \cdot \| \underline{x}_0 \|_2 , \tag{23a}$$

or

$$\| \underline{x}(k) \|_2 \leq \kappa_2(X) . \max_j\{|\lambda_j^k|\} \| \underline{x}_0 \|_2 . \tag{23b}$$

Proof: From (9) we obtain

$$\| F \|_2 \leq \| Z^{-1} \|_2 \| U_0^T \|_2 (\| X \|_2 \| X^{-1} \|_2 \| \Lambda \|_2 + \| A \|_2) , \tag{24}$$

and from the singular value decomposition of B we have $\| Z^{-1} \|_2 = \| V \|_2 \sigma_m^{-1}\{B\}$. Then the result (22) follows from the orthogonality of V and

$U = [U_0, U_1]$. The transient response of system (2) is easily seen to satisfy

$$\underline{x}(t) = e^{(A+BF)t}\underline{x}_0 = Xe^{\Lambda t}X^{-1}\underline{x}_0 \ , \tag{25a}$$

or

$$\underline{x}(k) = (A+BF)^k\underline{x}_0 = X\Lambda^k X^{-1}\underline{x}_0 \ , \tag{25b}$$

and the inequalities (23) follow directly upon taking norms. □

This theorem demonstrates that given (A,B) and eigenvalue set $\mathcal{L}$, minimizing the conditioning $\kappa_2(X)$ of the assigned eigensystem also minimizes a bound on the feedback gains and a bound on the transient response of the closed loop system, for any given initial condition.

Also of interest is the maximum disturbance which can be made to the closed loop system such that stability is retained. We have the following.

Theorem 6 If the state feedback matrix F assigns the set $\mathcal{L}$ of stable eigenvalues λ_j, then the perturbed closed loop system matrix $A+BF+\Delta$ remains stable for all disturbances Δ which satisfy, in the continuous time case,

$$\|\Delta\|_2 < \min_{s=i\omega} \sigma_n\{sI - (A+BF)\} \equiv \delta(F) \ , \tag{26}$$

where a lower bound on $\delta(F)$ is given by

$$\delta(F) \geq \min_j \mathrm{Re}(-\lambda_j)/\kappa_2(X) \ . \tag{27}$$

In the discrete time case, the closed loop system remains stable for disturbances Δ which satisfy

$$\|\Delta\|_2 < \min_{s=e^{i\omega}} \sigma_n\{sI - (A+BF)\} \equiv \delta(F) \ ,$$

where

$$\delta(F) \geq \min_j (1 - |\lambda_j|)/\kappa_2(X) \ .$$

Proof: The first part follows directly from a standard argument. The $n \times n$ matrix $M+\Delta \equiv M(I+M^{-1}\Delta)$ is non-singular, assuming M is non-singular, provided

$$\|M^{-1}\Delta\|_2 \leq \|M^{-1}\|_2 \|\Delta\|_2 < 1 , \tag{28}$$

that is, provided $\|\Delta\|_2 < \|M^{-1}\|_2^{-1} = \sigma_n\{M\}$.

Hence, the matrix $sI - (A+BF+\Delta)$ can become singular along the imaginary axis, where $s = i\omega$, only if $\|\Delta\|_2 \geq \delta(F)$. By continuity of the eigenvalues, the matrix $A+BF+\Delta$ is, therefore, stable provided (26) holds.

From (6), it easily follows that

$$\begin{aligned}\delta(F) &= \min_{s=i\omega} \sigma_n\{sI - X\Lambda X^{-1}\} \\ &\geq \sigma_n\{X\}\sigma_n\{X^{-1}\} \min_{s=i\omega} \sigma_n\{sI - \Lambda\} \\ &\geq \min_j \mathrm{Re}(-\lambda_j)/\|X^{-1}\|_2 \|X\|_2 ,\end{aligned} \tag{29}$$

which gives the lower bound (27). For the discrete time case the result is obtained analogously. □

From Theorem 6 we can deduce a lower bound on the stability margin of the closed loop system. The result is given by

Theorem 7 The return difference $I+G(s)+\hat{\Delta}(s)G(s)$, of the disturbed closed loop system, where $G(s) = -F(sI - A)^{-1}B$, remains non-singular at $s=i\omega$ for disturbances $\hat{\Delta}(s)$ which satisfy $\|\hat{\Delta}(i\omega)\|_2 < \tilde{\delta}(F)$, where $\tilde{\delta}(F)$ is bounded below by

$$\tilde{\delta}(F) \geq \delta(F)/\|B\|_2 \|F\|_2 . \tag{30}$$

(Taking $s=e^{i\omega}$ gives the corresponding result for the discrete time system.)

Proof: It is easily shown that

$$\det(sI - (A+BF+\Delta)) = \det(sI - A)\det(I+(I+\hat{\Delta}(s))G(s)) \tag{31}$$

if $\Delta = B\hat{\Delta}(s)F$. (See [1], for example.) Hence, $I+G(s)+\hat{\Delta}(s)G(s)$ is non-singular at $s=i\omega$ provided

$$\|\Delta\|_2 \leq \|B\|_2 \|\hat{\Delta}(i\omega)\|_2 \|F\|_2 < \delta(F) \tag{32}$$

and (30) follows. □

From Theorems 5-7 it can be seen that if the conditioning $\kappa_2(X)$ of the assigned eigensystem is minimized, then a lower bound on the stability margin of the closed loop system is maximized over all feedback matrices F which assign the given (stable) eigenvalues.

Minimizing the sensitivities of the assigned poles of the closed loop system thus ensures other desirable properties of the system. For given data, the minimal conditioning that can be achieved is limited, however, and in the next section we derive a lower bound on $\kappa_2(X)$. In §2.5 we examine other measures of the conditioning of the eigenproblem (6) which are directly related to $\kappa_2(X)$.

2.4 Optimal Robustness

In this section we derive a general result on the conditioning of a matrix with columns selected from given subspaces. A lower bound on the condition number $\kappa_2(X)$ is found, where X is an $n \times n$ matrix of the form

$$X = [X_1, X_2, \ldots X_k] \ , \tag{33}$$

with $n \times r_j$ submatrices X_j which are selected from given subspaces S_j, $j = 1,2,..k$, which together span the whole space H where $H = \mathbb{R}^n$ or $H = \mathbb{C}^n$, that is, $X_j \subset S_j$ and

$$S_1 + S_2 + .. + S_k = H \ . \tag{34}$$

We let S_j be an $n \times m_j$ matrix with orthonormal columns spanning the space S_j of dimension m_j, and we write

$$X_j = S_j D_j \ , \tag{35}$$

where D_j is a $m_j \times r_j$ matrix, which is of rank r_j if the columns of X_j are linearly independent. Matrix X may be written

$$X = [S_1, S_2, \ldots S_k] \cdot \mathrm{diag}\{D_1, D_2, .. D_k\} \equiv SD \ . \tag{36}$$

Our aim now is to estimate the minimal condition number $\kappa_2(X)$ over all possible selections D_i in (36). We first require the following.

Lemma 1. If S and D are $n \times s$ and $s \times n$ matrices, respectively, of rank n, $(n \leq s)$, then

$$\| D(SD)^{-1} \|_2 \geq \| S^+ \|_2 \ . \tag{37}$$

Here $+$ denotes the Moore-Penrose inverse, given by

$$S^+ = V_s \Sigma_s^{-1} U_s^* \ , \tag{38}$$

where

$$S = U_s \begin{bmatrix} \Sigma_s, 0 \end{bmatrix} \begin{bmatrix} V_s, \hat{V}_s \end{bmatrix}^* = U_s \Sigma_s V_s^* \tag{39}$$

is the singular value decomposition of S.

Proof: If (SD) is not invertible, then the result is trivial, since S^+ is bounded. Otherwise we write the SVD decomposition of D as

$$D = [U_d, \hat{U}_d] \begin{bmatrix} \Sigma_d \\ 0 \end{bmatrix} V_d^* = U_d \Sigma_d V_d^* , \tag{40}$$

and then

$$D(SD)^{-1} = [U_d, \hat{U}_d] \begin{bmatrix} \Sigma_d \\ 0 \end{bmatrix} (\Sigma_s V_s^* U_d \Sigma_d)^{-1} U_s^* \tag{41}$$

Taking norms we have

$$\| D(SD)^{-1} \|_2 = \left\| \begin{bmatrix} \Sigma_d \\ 0 \end{bmatrix} (\Sigma_s V_s^* U_d \Sigma_d)^{-1} \right\|_2 = 1/\sigma_n\{\Sigma_s V_s^* U_d\}, \tag{42}$$

where

$$\Sigma_s V^* U_d \equiv \begin{bmatrix} \Sigma_s, 0 \end{bmatrix} \begin{bmatrix} V_s^* \\ \hat{V}_s^* \end{bmatrix} \begin{bmatrix} U_d, \hat{U}_d \end{bmatrix} \begin{bmatrix} I \\ 0 \end{bmatrix} \equiv G \begin{bmatrix} I \\ 0 \end{bmatrix} . \tag{43}$$

It can now be shown (see [25], Theorem 1) that

$$\sigma_n \left\{ G \begin{bmatrix} I \\ 0 \end{bmatrix} \right\} \leq \sigma_n\{G\} , \tag{44}$$

and, since we have here $\sigma_n\{G\} = \sigma_n\{\Sigma_s\}$, it follows readily that

$$\| D(SD)^{-1} \|_2 \geq 1/\sigma_n\{\Sigma_s\} = \| S^+ \|_2 . \qquad \square$$

Using this lemma we now establish the following theorem.

Theorem 8. If X,S,D are defined as in (33) and (36), then

$$\kappa_2(S) \equiv \| S \|_2 \| S^+ \|_2 \leq \sqrt{1+(k-1)\cos\theta_{min}} \cdot \kappa_2(X) \leq \sqrt{k}\, \kappa_2(X) , \tag{45}$$

where θ_{min} is the minimum angle between any two subspaces S_i and S_j, defined by

$$\cos\theta_{min} = \max_{i \neq j} \| S_j^* S_i \|_2 \leq 1 . \tag{46}$$

Proof: Since the columns of matrices S_j, $j = 1,2,..k$ are orthonormal by definition, we have

$$\| SD \|_2 \geq \max_j \| S_j D_j \|_2 = \max_j \| D_j \|_2 = \| D \|_2 , \tag{47}$$

and, therefore,

$$\kappa_2(SD) = \| SD \|_2 \| (SD)^{-1} \|_2 \geq \| D \|_2 \| (SD)^{-1} \|_2 . \tag{48}$$

Then by the multiplicative property of the norm and by Lemma 1

$$\kappa_2(X) = \kappa_2(SD) \geq \| D (SD)^{-1} \|_2 \geq \| S^+ \|_2 . \tag{49}$$

Now, if we let $\underline{x}$ be any vector of unit norm, partitioned conformably with S, such that $\underline{x}^T = (\underline{x}_1^T, \underline{x}_2^T, \ldots \underline{x}_k^T)$, then

$$\begin{aligned} \| S\underline{x} \|_2^2 &= \left| \sum_{i=1}^{k} \sum_{j=1}^{k} \underline{x}_i^* S_i^* S_j \underline{x}_j \right| \\ &\leq \sum_{i=1}^{k} |\underline{x}_i^* \underline{x}_i| + \sum_{i=1}^{k} \sum_{\substack{j=1 \\ j \neq i}}^{k} \| \underline{x}_i \|_2 \| \underline{x}_j \|_2 \cos\theta_{min} \\ &\leq 1 + \left\{ \left(\sum_{i=1}^{k} \sum_{j=1}^{k} \| \underline{x}_i \|_2 \| \underline{x}_j \|_2 \right) - 1 \right\} \cos\theta_{min} \\ &\leq 1 + \left\{ \left(\sum_{j=1}^{k} \| \underline{x}_j \|_2 \right)^2 - 1 \right\} \cos\theta_{min} . \end{aligned} \tag{50}$$

But, by the Cauchy-Schwarz inequality we have

$$\left(\sum_{j=1}^{k} \| \underline{x}_j \|_2 \right)^2 \leq k \left(\sum_{j=1}^{k} \| \underline{x}_j \|_2^2 \right) = k , \tag{51}$$

and it follows that for any $\underline{x}$ with $\| \underline{x} \|_2 = 1$,

$$\| S\underline{x} \|_2^2 \leq 1 + (k-1) \cos\theta_{min} \tag{52}$$

Thus

$$\| S \|_2 \leq \sqrt{1 + (k-1) \cos\theta_{min}} ,$$

and from (49) we obtain

$$\kappa_2(X)\sqrt{1 + (k-1)\cos\theta_{min}} \geq \|S^+\|_2 \|S\|_2 = \kappa_2(S) . \qquad (53)$$

The second inequality in (45) then follows directly from (46). □

Since Theorem 8 holds for all matrices X of form (33), the inequality (45) gives a lower bound on the minimal conditioning of X, and we must have

$$\min \kappa_2(X) \geq \kappa_2(S)/\sqrt{k} . \qquad (54)$$

We remark that this result is quite general and holds over all choices of $\{r_j\}_1^k$ such that

$$r_j \leq s_j , \quad \sum_{j=1}^{k} r_j = n . \qquad (55)$$

For a particular choice of $\{r_j\}_1^k$ (for example $r_j = 1, \forall_j$) the lower bound given by (54) is not necessarily realizable, however. In particular, if any submatrix of S, composed of $p < k$ of the submatrices S_j with $j \in J_p$, has rank less than $q \equiv \sum_{j\in J_p} r_j$, then a non-singular matrix X cannot be selected; if this submatrix of S has numerical rank less than q, or in other words, if the ratio of its first and q-th (non-zero) singular values, σ_1/σ_q, is large, then the solution X must of necessity, be badly conditioned, even though the condition number $\kappa_2(S)$ may be of reasonable size.

We observe also that the result (54) holds for any scaling of the matrix X, and the minimum may not be realizable if a particular scaling is imposed. For example, if the blocks X_j, $j=1,2,..k$, are constrained to have orthonormal columns, then we may take the matrix $\hat{S} = X$ with $\hat{s}_j = r_j$ and apply Theorem 6 with S replaced by $\hat{S}$. The unconstrained matrix X_0, derived from $\hat{S}$, with minimal conditioning then satisfies $\sqrt{k}\,\kappa_2(X_0) \geq \kappa_2(X)$, by (47). This suggests

that the optimal "block"-scaling cannot give much better conditioning than a scaling in which the columns of the blocks X_j are chosen to be orthonormal.

In the next section we examine other measures of robustness and demonstrate their mathematical relations to the measure $\kappa_2(X)$ of the conditioning of the eigenproblem (6).

2.5 Measures of Robustness

We now investigate relations between the condition numbers c_j (or sensitivities) of the closed loop poles, defined by (4), and various measures ν of the robustness of the eigenproblem (6). We have seen already (§2.1) that $\nu_2 = \kappa_2(X)$, the condition number of the matrix X of eigenvectors, provides an upper bound on the measure $\nu_1 = \|\underline{c}\|_\infty \equiv \max_j\{c_j\}$. We here derive other measures ν of robustness which can be bounded in terms of $\kappa_2(X)$ and which take their minimal values simultaneously when the eigenproblem (6) is perfectly conditioned, that is, when the assigned poles are as insensitive as is possible.

We assume that the right eigenvectors $\underline{x}_j$ of the closed loop system matrix are normalized such that $\|\underline{x}_j\|_2 = 1$, and write

$$Y^T \equiv [\underline{y}_1, \underline{y}_2, .. \underline{y}_n]^T = X^{-1}. \tag{56}$$

Then the condition numbers are given by

$$c_j = \|\underline{y}_j\|_2 \geq 1, \tag{57}$$

and we have

$$\|X\|_F = n^{\frac{1}{2}}, \qquad \|X^{-1}\|_F \equiv \|Y^T\|_F = \left(\sum_j c_j^2\right)^{\frac{1}{2}}, \tag{58}$$

where $\|\cdot\|_F$ is the Frobenius norm. If we define the measure ν_3 by

$$\nu_3 = n^{-\frac{1}{2}} \|X^{-1}\|_F, \tag{59}$$

then $\nu_3 = n^{-\frac{1}{2}} \|\underline{c}\|_2$ and ν_3 takes its minimal value, unity, if and only if $c_j = 1$, $\forall_j$. Furthermore, $\nu_3 = n^{-1} \|X\|_F \|X^{-1}\|_F \equiv n^{-1}\kappa_F(X)$, and by the equivalence of norms [28] we obtain

$$1 \leq \nu_3 \leq \nu_1 \leq \nu_2 \leq n\nu_3. \tag{60}$$

The measures ν_1,ν_2,ν_3 are, thus, mathematically equivalent, and the three measures take their minimal values simultaneously when the eigensystem is perfectly conditioned and X is unitary. We note that if $\nu_2 = \kappa_2(X)$ is close to unity, that is, $(\nu_2 - 1) = \varepsilon$, $\varepsilon << 1$, then the measures ν_1 and ν_3 are also close to unity, that is $(\nu_1 - 1) \leq \varepsilon$ and $(\nu_3 - 1) \leq \varepsilon$. Furthermore, if <u>any</u> of the three measures is close to unity, then the sum of the squares of the condition numbers, $\| \underline{c} \|_2^2$, is near to optimal, that is: $(\nu_1 - 1) \leq \varepsilon$, $(\nu_2 - 1) \leq \varepsilon$, or $(\nu_3 - 1) \leq \varepsilon$ implies $(n^{-\frac{1}{2}} \| \underline{c} \|_2 - 1) \leq \varepsilon$.

If a real diagonal scaling $D = \mathrm{diag}\{d_1,d_2,\ldots d_n\}$, with $d_j > 0\ \forall_j$, is applied to the matrix Y^T of left eigenvectors, we find

$$\| XD^{-1} \|_F = \| D^{-1} \|_F \equiv \left(\sum_j d_j^{-2} \right)^{\frac{1}{2}}, \qquad \| DX^{-1} \|_F \equiv \| DY^T \|_F = \left(\sum_j d_j^2 c_j^2 \right)^{\frac{1}{2}}, \tag{61}$$

and we may define a weighted measure

$$\nu_3(D) = \| DX^{-1} \|_F / \| D \|_F . \tag{62}$$

Then $\nu_3(D) = \| D \underline{c} \|_2 / \| D \|_F$, from (61), and we have also $\nu_3(D) \equiv \kappa_F(XD^{-1})/\kappa_F(D)$. The measure $\nu_3(D)$ clearly takes its minimal value, unity, if and only if $c_j = 1$, $\forall_j$, or, equivalently, X is unitary; thus, $\nu_3(D)$ attains its minimum simultaneously with the other measures. By norm equivalences, we have also

$$\{1/\kappa_2(D)\}\nu_3 \leq \nu_3(D) \leq \kappa_2(D)\nu_3, \tag{63}$$

where $\kappa_2(D) = \max_j d_j / \min_j d_j$, and $\nu_3(D)$ is equivalent to the measures ν_1,ν_2,ν_3.

The square of the measure $\nu_3(D)$ is proportional to a weighted sum of the squares of the condition numbers. By choosing appropriate weights

the relative conditioning of each eigenvalue can be controlled. Different weightings correspond to different scalings of the eigenvectors. Taking $d_j = 1$, $\forall_j$, corresponds to taking the eigenvectors to have unit length. In this case $n^{\frac{1}{2}}\nu_3(I) \equiv n^{\frac{1}{2}}\nu_3 = \| \underline{c} \|_2$, so that ν_3 is directly proportional to the ℓ_2-norm of the vector $\underline{c}$ of condition numbers. With $d_j = c_j^{-\frac{1}{2}}$, $j = 1,2,\ldots n$, we find

$$\kappa_F(D)\nu_3(D) = \kappa_F(XD^{-1}) = \sum_j c_j \equiv \| \underline{c} \|_1 , \tag{64}$$

and the ℓ_1-norm of the vector of condition numbers is obtained. This choice of weights corresponds to the optimal scaling of the eigenvectors, with respect to the Frobenius norm [24]. We note that the weights d_j here are dependent on the condition numbers c_j. More generally, we observe that with a suitable scaling, dependent on $\underline{c}$, any weighted ℓ_p-norm of $\underline{c}$ can be achieved. In particular, if $\Gamma = \mathrm{diag}\{\gamma_1, \gamma_2, \ldots \gamma_n\}$ and we choose $d_j^2 = \gamma_j^p c_j^{p-2}$, $j = 1,2,..n$, then

$$\| \Gamma \underline{c} \|_p^p \equiv \sum_j \gamma_j^p c_j^p = \sum_j d_j^2 c_j^2 \equiv \kappa_F^2(XD^{-1}) / \| D^{-1} \|_F^2 \equiv \| D \|_F^2 \nu_3^2(D). \tag{65}$$

We remark that it is not easy in practice to implement scalings which are dependent on the condition numbers, and constant weights d_j, $j = 1,2,..n$, independent of c_j, $\forall_j$, are primarily used.

The conditioning $\nu_2 = \kappa_2(X)$ of the eigenproblem (6) is optimal ($\nu_2 = 1$) if and only if the matrix X of normalized right eigenvectors $\underline{x}_j$, $j = 1,2,\ldots n$, is unitary, as we have already observed (§2.1). In essence, the aim of the robust pole placement problem is, therefore, to select eigenvectors $\underline{x}_j \in S_j$, such that $\| \underline{x}_j \|_2 = 1$ and the vectors $\underline{x}_j$ are as "orthogonal" as possible to each other. We now consider as a measure of "orthogonality" the distance between the matrix X of eigenvectors and some unitary matrix $\hat{X}$. The unitary matrix $\hat{X}$ is taken to be such that each

column $\underline{x}_j$ of X is the (normalized) orthogonal projection of the column $\hat{\underline{x}}_j$ of $\hat{X}$ into the space S_j. (The distance between $\hat{X}$ and X is then the minimal distance between $\hat{X}$ and any matrix of normalized eigenvectors belonging to the required subspaces.) Without loss of generality we may assume that $\underline{x}_j$ is scaled so that $\hat{\underline{x}}_j^* \underline{x}_j$ is <u>real</u> and <u>positive</u>, and then we have

$$\hat{\underline{x}}_j^* \underline{x}_j = \| S_j^* \hat{\underline{x}}_j \|_2 = \sqrt{1 - \phi_j^2} \equiv \cos\theta_j \leq 1, \tag{66}$$

where θ_j is the angle between $\hat{\underline{x}}_j$ and $\underline{x}_j$ (its projection into S_j), $\phi_j = \sin\theta_j$, and the columns of matrix S_j form an orthonormal basis for the space S_j. As a fourth measure of the conditioning of the eigenproblem, we may then take

$$\nu_4 = n^{-\frac{1}{2}} (\sum_j \phi_j^2)^{\frac{1}{2}} \equiv n^{-\frac{1}{2}} (\sum_j \sin^2\theta_j)^{\frac{1}{2}}. \tag{67}$$

Clearly, $0 \leq \nu_4 \leq 1$, and $\nu_4 = 0$ if and only if the set of eigenvectors $\{\underline{x}_j\}$ is orthonormal and the condition numbers $c_j = 1$, $\forall_j$.

The square of the measure ν_4 is proportional to the sum of the squares of the sines of the angles between the eigenvectors $\underline{x}_j$ and a ("closest") orthonormal set of vectors. A weighted sum may also be used as a measure. We define

$$\nu_4(D) = (\sum_j d_j^2 \phi_j^2)^{\frac{1}{2}} / (\sum_j d_j^2)^{\frac{1}{2}}, \tag{68}$$

where $D = \mathrm{diag}\{d_1, d_2, \ldots d_n\}$, as before. Here $0 \leq \nu_4(D) \leq 1$, and $\nu_4(D)$ is minimal if and only if the eigenvectors are orthonormal. We note that $\nu_4 \equiv \nu_4(I)$.

To establish a relationship between the measure $\nu_4(D)$ and the other measures of conditioning we first show that $\nu_4(D)$ is equivalent to $\| D(I - \hat{X}^* X) \|_F$. Denoting the elements of $\Delta = I - \hat{X}^* X$ by δ_{kj}, we find

that since $\| \underline{x}_j \|_2 = 1$ and $\hat{X}$ is unitary, then

$$\| \hat{X}^* \underline{x}_j \|_2 = 1 = 1 + \sum_k |\delta_{kj}|^2 - 2\delta_{jj}, \tag{69}$$

where $\delta_{jj} = 1 - \underline{\hat{x}}_j^* \underline{x}_j$ is real and positive. From (66) we have also $\delta_{jj} \leq 1$ and

$$1 + \delta_{jj}^2 - 2\delta_{jj} = 1 - \phi_j^2, \tag{70}$$

and, therefore,

$$2\delta_{jj} = \phi_j^2 + \delta_{jj}^2 \leq \phi_j^2 + \delta_{jj}. \tag{71}$$

We conclude then that $\delta_{jj} \leq \phi_j^2$, and from (69) we obtain

$$\phi_j^2 \leq 2\delta_{jj} = \sum_k |\delta_{kj}|^2 \leq 2\phi_j^2. \tag{72}$$

It follows that

$$\frac{\min d_k^2}{\max d_k^2} \sum_j d_j^2 \phi_j^2 \leq \sum_j \sum_k d_k^2 |\delta_{kj}|^2 \equiv \| D\Delta \|_F^2 \leq 2 \frac{\max d_k^2}{\min d_k^2} \sum_j d_j^2 \phi_j^2 , \tag{73}$$

and, therefore,

$$(1/\kappa_2(D)) \| D \|_F \nu_4(D) \leq \| D\Delta \|_F \leq \sqrt{2}\kappa_2(D) \| D \|_F \nu_4(D) . \tag{74}$$

We now observe that in the special case $D = I$,

$$n^{\frac{1}{2}} \nu_3 = \| X^{-1} \|_F = \| X^{-1}\hat{X} \|_F = \| (I-\Delta)^{-1} \|_F , \tag{75}$$

and if $\| \Delta \| < 1$, then

$$\| (I-\Delta)^{-1} \|_F = \| I \|_F + \sum_1^{\infty} \| \Delta \|_F^k \leq n^{\frac{1}{2}} + \| \Delta \|_F / (1 - \| \Delta \|_F). \tag{76}$$

From (74)-(76) we thus obtain

$$1 \leq \nu_3 \leq 1 + \sqrt{2}\nu_4 / (1 - \sqrt{2n}\nu_4), \tag{77}$$

provided $\nu_4 < (2n)^{-\frac{1}{2}}$. If ν_4 is small, that is, $\nu_4 < \varepsilon$, then ν_3 is close to unity, that is, $(\nu_3 - 1) < \sqrt{2}\varepsilon + O(\varepsilon^2)$, and $\| \underline{c} \|_2$ is near optimal.

In the case $D \neq I$, then the measures $\nu_3(D)$ and $\nu_4(D)$ can be compared. We obtain now

$$\| D \|_F \nu_3(D) = \| DX^{-1} \|_F = \| DX^{-1}\hat{X} \|_F = \| D(I-\Delta)^{-1} \|_F , \tag{78}$$

and then $D(I-\Delta)^{-1} = (I-D\Delta D^{-1})^{-1}D$ implies that, if $\| D^{-1} \|_F \| D\Delta \|_F < 1$,

$$\| D(I-\Delta)^{-1} \|_F \leq \| D \|_F / (1 - \| D^{-1} \|_F \| D\Delta \|_F). \tag{79}$$

From (74) and (78), it then follows that

$$1 \leq \nu_3(D) \leq 1/(1 - \sqrt{2}\kappa_2(D)\kappa_F(D)\nu_4(D)), \tag{80}$$

provided $\nu_4(D) < 1/\sqrt{2}\kappa_2(D)\kappa_F(D)$. The measure $\nu_3(D)$ thus attains its minimum simultaneously with $\nu_4(D)$, and if $\nu_4(D)$ is small, then $\nu_3(D)$ is close to unity and the weighted ℓ_2-norm of the vector of condition numbers is near optimal.

In summary, we have defined four measures of the conditioning of the eigenproblem (6):

$$\nu_1 = \| \underline{c} \|_\infty ,$$

$$\nu_2 = \kappa_2(X) ,$$

$$\nu_3 = \| X^{-1} \|_F / n^{\frac{1}{2}} \equiv \| \underline{c} \|_2 / n^{\frac{1}{2}} ,$$

$$\nu_4 = (\sum_j \sin^2\theta_j)^{\frac{1}{2}} / n^{\frac{1}{2}} ,$$

where $\underline{c} = [c_1, c_2, \ldots c_n]$ is the vector of condition numbers, $X = [\underline{x}_1, \underline{x}_2, \ldots \underline{x}_n]$ is the matrix of eigenvectors and θ_j are the angles between eigenvectors $\underline{x}_j$ and certain corresponding orthonormal vectors $\hat{\underline{x}}_j$, $j = 1,2,\ldots n$. These measures all attain their minimal values simultaneously when the eigenproblem (6) is perfectly conditioned and the assigned eigenvalues are as insensitive as possible. If

any of the measures is minimized, then an upper bound on the sum of the squares of the condition numbers is minimized.

Two weighted measures of conditioning are also defined:

$$\nu_3(D) = \| DX^{-1} \|_F / \| D \|_F \equiv (\sum_j d_j^2 c_j^2)^{\frac{1}{2}} / (\sum_j d_j^2)^{\frac{1}{2}} ,$$

$$\nu_4(D) = (\sum_j d_j^2 \sin^2\theta_j)^{\frac{1}{2}} / (\sum_j d_j^2)^{\frac{1}{2}} .$$

These measures are minimal when the eigenproblem is perfectly conditioned, and minimizing these measures minimizes an upper bound on a weighted sum of the squares of the condition numbers. A suitable set of weights is given by $d_j^{-1} = \mathrm{Re}(-\lambda_j)$ in the continuous time case, or $d_j^{-1} = (1 - |\lambda_j|)$ in the discrete time case. Then minimizing $\nu_3(D)$, or $\nu_4(D)$, corresponds to maximizing a lower bound on the stability margin of the closed loop system, as given by Theorems 6 and 7 of §2.4, with an appropriate scaling of the eigenvectors. For this choice of weights eigenvalues close to the imaginary axis, which are most likely to become unstable, are required to have much better conditioning, and hence to be less sensitive to disturbances, than those which lie far away from the axis.

In the next section we describe numerical methods for iteratively constructing a well-conditioned set of eigenvectors from the required subspaces. The procedures all aim, at each step of the iteration, to reduce the value of one of the measures discussed here.

3.0 NUMERICAL ALGORITHMS FOR ROBUST POLE ASSIGNMENT

3.1 Basic Steps

We now consider the practical implementation of the theoretical results discussed in section 2. We describe four numerical methods for obtaining solutions to the robust pole placement problem (Problem 2). Two of the methods are complementary. The procedures all consist of three basic steps:

Step A: Compute the decomposition of matrix B, given by (8), to determine U_0, U_1 and Z; construct orthonormal bases, comprised by the columns of matrices S_j and $\hat{S}_j$ for the space $\mathcal{S}_j \equiv N\{U_1^T(A-\lambda_j I)\}$ and its complement, $\hat{\mathcal{S}}_j$ for $\lambda_j \in \mathcal{L}$, $j = 1,2,..n$.

Step X: Select vectors $\underline{x}_j = S_j \underline{w}_j \in \mathcal{S}_j$ with $\|\underline{x}_j\|_2 = 1$, $j = 1,2,..n$, such that $X = [\underline{x}_1, \underline{x}_2, .. \underline{x}_n]$ is well-conditioned.

Step F: Find the matrix $M \equiv A + BF$ by solving $MX = X\Lambda$ and compute F explicitly from $F = Z^{-1} U_0^T(M-A)$.

Standard library software for obtaining QR and SVD (singular value) decompositions of matrices and for solving systems of linear equations [3] are used to accomplish these steps. The first and third steps, Step A and Step F are identical for all the methods. The key step is Step X. Here the vectors $\underline{x}_j$, $j = 1,2,..n$ are chosen by an iterative process to minimize one of the measures of conditioning described in §2.5. These measures are all equivalent in a certain sense, and when they are close to unity, the sensitivities c_j of the assigned eigenvalues λ_j are all close to minimal.

We first discuss the two basic steps common to all four methods.

3.1 Step A

The required decomposition of B is found in Step A by either the SVD or QR method, as discussed in §2.2. We note that the QR decomposition is computationally less expensive, but that the SVD decomposition gives useful information on the singular values of B.

Construction of the bases for S_j and $\hat{S}_j$ is also achieved by QR (Case 1) or SVD (Case 2) decompositions as follows.

Case 1 (QR) We determine the QR decomposition of $(U_1^T(A-\lambda_j I))^T$ partitioned as

$$(U_1^T(A-\lambda_j I))^T = [\hat{S}_j, S_j] \begin{bmatrix} R_j \\ 0 \end{bmatrix}.$$

Then $S_j, \hat{S}_j$ are the required matrices.

Case 2 (SVD) We determine the singular value decomposition of $U_1^T(A-\lambda_j I)$ in the partitioned form

$$U_1^T(A-\lambda_j I) = T_j[\Gamma_j, 0][\hat{S}_j, S_j]^T,$$

where Γ_j is the diagonal matrix of singular values. Then the columns of $S_j, \hat{S}_j$ give the required orthonormal bases.

We note that the decompositions can be carried out most efficiently if the matrix $[B|A]$ is first reduced to staircase form (see [26]). This requires less than $n^2(3n+m)$ operations. The number of operations needed to find each subspace is then $m(n-m)(2n-m)$, or a total of $O(n^3m)$ operations.

3.2 Step F.

The matrix $\tilde{M} = X\Lambda X^{-1}$ is constructed in Step F by solving the equation $X^T M^T = (X\Lambda)^T$ for M^T using a direct L-U decomposition (or Gaussian elimination) method. This process is stable for a well-conditioned matrix X. The computation of F is then achieved by straightforward matrix multiplication in the case Z is given by the SVD process, or by using back substitution to solve the equations $ZF = U_0^T(M-A)$ in the case Z is given by the QR process.

The computation of M requires $O(n^3)$ operations and the computation of F needs $O(nm^2)$ operations. We remark that the total amount of work required in Steps A and F is comparable to the number of operations needed for one iteration in Step X, and is not a significant factor in the total operation count.

3.3 Step X

We describe four methods for accomplishing the main step, Step X, in the basic algorithm. The methods are each iterative and each aims to minimize a different measure of the conditioning of matrix X, although two of the methods use complementary measures. We discuss the techniques only for the case where the eigenvectors are required to be real. Detailed descriptions of the methods, together with modifications for the complex case and other special cases, are given elsewhere [7] [12].

3.3.1 Method 0: The objective here is to choose vectors $\underline{x}_j \in S_j$, $j = 1,2,..n$ such that each vector is as orthogonal as possible to the space spanned by the remaining vectors; that is, such that the angle between vector $\underline{x}_j \in S_j$ and the space $X_j = < \underline{x}_i, i \neq j >$ is maximized $\forall_j$. Equivalently, we choose $\underline{x}_j \in S_j$ to minimize the angle between $\underline{x}_j$ and the normalized vector $\tilde{\underline{y}}_j$ orthogonal to the space X_j, $\forall_j$.

The solution is found by an iteration in which each vector $\underline{x}_j$ is replaced by a new vector with maximum angle to the current space X_j for each $j = 1,2,..n$, in turn. The new vector is obtained by the QR method. The decomposition

$$X_j \equiv [\underline{x}_1, \underline{x}_2, \ldots \underline{x}_{j-1}, \underline{x}_{j+1}, \ldots \underline{x}_n] = [\tilde{Q}_j, \tilde{\underline{y}}_j] \begin{bmatrix} \tilde{R}_j \\ \underline{0}^T \end{bmatrix}$$

is formed, in order to find $\tilde{\underline{y}}_j$ orthogonal to X_j, and then the projection of $\tilde{\underline{y}}_j$ into S_j, given by

$$\underline{x}_j \equiv S_j \underline{w}_j = S_j S_j^T \tilde{\underline{y}}_j / \| S_j^T \tilde{\underline{y}}_j \|_2 ,$$

is the vector in S_j which has minimum angle to $\tilde{\underline{y}}_j$. The iteration is continued until the reduction in $\nu_2 = \kappa_2(X)$, after a full sweep of the process ($j = 1,2,..n$), is less than some positive tolerance. Any set of

independent vectors $\underline{x}_j \in S_j$ can be taken to give the initial matrix X.

In effect, at each step of the iteration a rank-one up-date to the matrix X is made such as to minimize the sensitivity of the eigenvalue λ_j. The vector $\underline{\tilde{y}}_j^T$ is just the normalized left eigenvector corresponding to λ_j, and therefore the condition $c_j = 1/|\underline{\tilde{y}}_j^T \underline{x}_j|$ is minimized by the choice of $\underline{x}_j$. The conditioning of the remaining eigenvalues λ_i, $i \neq j$ is disturbed, however, when the new vector replaces the old vector $\underline{x}_j$, and the overall conditioning is not necessarily improved at each step. The process does not necessarily converge to a fixed point, therefore. The method is simple to implement, however, and gives good solutions in practice. At each step j of the iteration, the QR decomposition of X_j is obtained by a rank-one update of the decomposition of X_{j-1}, which requires $O(n^2)$ operations. The computation of $\underline{x}_j$ as the projection of $\underline{\tilde{y}}_j$ into S_j requires $O(nm)$ operations, and the operation count for one full sweep of the procedure is thus $O(n^3) + O(n^2m)$.

3.3.2 Method 1 As in Method 0, the solution is here found by an iteration in which a rank-one update is made to matrix X at each step. The objective of the update is now to select a new vector $\underline{x}_j \in S_j$, for each $j = 1,2,..n$, such as to minimize the measure of conditioning $\nu_3(D) = \|DX^{-1}\|_F^2 / \|D\|_F$ discussed in §2.5, over all $\underline{x}_j \in S_j$, at each step. Thus, at each step, a non-linearly constrained least square problem must be solved. This can be accomplished explicitly by QR decompositions. The measure $\nu_3(D)$ represents a weighted sum of the squares of all the condition numbers c_j, and hence, the overall conditioning of the solution X is improved at each step of the iteration and the process converges. Any set of independent vectors $\underline{x}_j \in S_j$ may be used to start the procedure. The iteration is stopped when the reduction in the measure $\nu_3(D)$, after a full sweep $(j = 1,2,..n)$ is

less than a given tolerance.

The technique for determining the update at step j is described here for the case $D = I$. (A complete discussion of the process is given elsewhere [12].) The problem is to find $\underline{w}_j$ with $\| \underline{w}_j \| = 1$ to minimize $\| X^{-1} \|_F$ where $\underline{x}_j = S_j \underline{w}_j$, and $X_j = [\underline{x}_1, \underline{x}_2, \ldots \underline{x}_{j-1}, \underline{x}_{j+1}, \ldots \underline{x}_n]$ is assumed known. We may write

$$\| X^{-1} \|_F = \| [S_j \underline{w}_j, X_j]^{-1} \|_F = \| [\underline{y}_j, Y_j]^T \|_F \equiv \| Y^T \|_F .$$

By QR decomposition we obtain

$$X_j = [Q_j, \underline{q}_j] \begin{bmatrix} R_j \\ \underline{0}^T \end{bmatrix},$$

and then $Y^T X = I$ implies

$$\| Y^T \|_F = \left\| \begin{bmatrix} \underline{0}^T & \rho_j \\ R_j^{-1} & -\rho_j R_j^{-1} Q_j^T S_j \underline{w}_j \end{bmatrix} \right\|_F$$

where $\rho_j = 1/(\underline{q}_j^T S_j \underline{w}_j)$. Using $\| \underline{w}_j \|_2 = 1$, we find, therefore, that to minimize $\| Y^T \|_F$ we must minimize

$$\rho_j^2 + \rho_j^2 \| R_j^{-1} Q_j^T S_j \underline{w}_j \|_2 \equiv \left\| \begin{bmatrix} R_j^{-1} Q_j^T S_j \\ I_m \end{bmatrix} \rho_j \underline{w}_j \right\|_2 .$$

Here ρ_j is a normalizing factor, dependent upon $\underline{w}_j$, which may be eliminated by making a further orthogonal decomposition. We find unitary matrix $\tilde{P}_j$ such that

$$\underline{q}_j^T S_j = \sigma_j \underline{e}_m^T \tilde{P}_j^T \equiv \sigma_j \underline{e}_m^T [P_j, \underline{p}_j]^T = \sigma_j \underline{p}_j^T ,$$

and then $\rho_j^{-1} = \sigma_j \underline{p}_j^T \underline{w}_j$. Writing $\tilde{\underline{w}}_j = \rho_j P_j^T \underline{w}_j$ we obtain

$$\rho_j \underline{w}_j = \rho \tilde{P}_j \tilde{P}_j^T \underline{w}_j = (P_j P_j^T \underline{w}_j + \underline{p}_j \underline{p}_j^T \underline{w}_j)/(\sigma_j P_j^T \underline{w}_j)$$

$$= \sigma_j^{-1}(P_j \tilde{\underline{w}}_j + \underline{p}_j).$$

The problem thus reduces to the linear least-square problem

$$\min \left\| \begin{bmatrix} R_j^{-1} Q_j^T S_j \\ I_m \end{bmatrix} (P_j \tilde{\underline{w}}_j + \underline{p}_j) \right\|_2 ,$$

which is solved for $\tilde{\underline{w}}_j$ by a standard technique using a further QR decomposition. The required update is then given by

$$\underline{x}_j = S_j \underline{w}_j \equiv (\rho_j \sigma_j)^{-1} S_j (P_j \tilde{\underline{w}}_j + \underline{p}_j) ,$$

where the normalizing constant ρ_j is determined by

$$\rho_j^2 = \rho_j^2 \underline{w}_j^T \underline{w}_j = \sigma_j^{-2} (\tilde{\underline{w}}_j^T \tilde{\underline{w}}_j + 1) .$$

In practice, this procedure gives solutions very similar to those of Method 0, although it is rather more complicated to implement. The computation of the QR decompositions of X_j (by a rank-one update of the decomposition of X_{j-1}) and of $\underline{q}_j^T S_j$ requires $O(n^2)$ and $O(2m)$ operations, respectively, and the solution of the least-square problem for $\tilde{\underline{w}}_j$ uses $O(nm^2)$ operations. The principal expense at each step of the iteration is in the computation of $R_j^{-1} Q_j^T S_j$, (obtained by back-substitutions,) which requires $O(n^2 m)$ operations. A full sweep of the process, therefore, requires a total of $O(n^3 m) + O(n^2 m^2)$ operations. Method 1 is thus rather more expensive per sweep than Method 0, but it is guaranteed to converge.

3.3.3 Methods 2/3 The objective of Method 2 is to determine an orthonormal set of vectors $\tilde{\underline{x}}_j$, $j = 1,2,..n$, such that some measure of the distance between the vectors $\tilde{\underline{x}}_j$ and the subspaces S_j is minimized; then the

required eigenvectors $\underline{x}_j$, $j = 1,2,..n$ are taken as the normalized projections of $\tilde{\underline{x}}_j$ into S_j. The resulting $\underline{x}_j$ are approximately orthogonal to each other and the conditioning of X is expected to be reasonably close to unity. The complementary objective of Method 3 is to select orthonormal vectors $\tilde{\underline{x}}_j$ such as to <u>maximize</u> the distance between the vectors and the complementary spaces $\hat{S}_j$, and then project $\tilde{\underline{x}}_j$ into S_j.

The measure of distance to be minimized in Method 2 is the weighted sum of the squares of the sines of the angles between the vectors and the subspaces, given by

$$\nu_4(D) = \left(\sum_j d_j^2 \phi_j^2\right)^{\frac{1}{2}} / \left(\sum_j d_j^2\right)^{\frac{1}{2}} ,$$

where

$$\phi_j = \sin\theta_j = \| \hat{S}_j^* \, \tilde{\underline{x}}_j \|_2$$

is the sine of the angle θ_j between $\tilde{\underline{x}}_j$ and its projection $\underline{x}_j$ into S_j. As shown in §2.5, minimizing $\nu_4(D)$ corresponds to minimizing $\nu_3(D)$, and if $\nu_4(D)$ is sufficiently small, then $\| D \, \underline{c} \|_2 = \left(\sum_j d_j^2 \, c_j^2\right)^{\frac{1}{2}}$ is also small and the solution $X = [\underline{x}_1, \underline{x}_2, .. \underline{x}_n]$ is well-conditioned (in a weighted sense).

The complementary measure to be maximized in Method 3 is given by

$$\hat{\nu}_4(D) = \left(\sum_j d_j^2 \cos^2\theta_j\right)^{\frac{1}{2}} / \left(\sum_j d_j^2\right)^{\frac{1}{2}} = 1 - \nu_4(D),$$

where

$$\cos\theta_j = (1-\phi_j^2)^{\frac{1}{2}} = \| S_j^* \, \tilde{\underline{x}}_j \|_2 .$$

The solutions obtained by Methods 2 and 3 are identical, but the complementary measure is computationally more efficient to use when $m < n - m$.

The vectors $\tilde{\underline{x}}_j$ are determined iteratively by applying plane rotations to the matrix $\tilde{X} = [\tilde{\underline{x}}_1, \tilde{\underline{x}}_2, .. \tilde{\underline{x}}_n]$ such that each rotation reduces (or increases) the measure $\nu_4(D)$ (or $\hat{\nu}_4(D)$) by an optimal quantity. Initially any orthogonal matrix, say $\tilde{X} = I$, may be taken. At each step of the iteration two indices $1 \leq j < k \leq n$ are selected and the two vectors $\tilde{\underline{x}}_j$, $\tilde{\underline{x}}_k$ of the current matrix $\tilde{X}$ are updated by a rotation in the plane which maintains their orthogonality and minimizes

$$d_j\phi_j^2 + d_k\phi_k^2 \equiv d_j^2 \| \hat{S}_j^* \tilde{\underline{x}}_j \|_2^2 + d_k^2 \| \hat{S}_k^* \tilde{\underline{x}}_k \|_2^2$$

(or equivalently, maximizes

$$d_j^2(1-\phi_j^2) + d_k^2(1-\phi_k^2) \equiv d_j^2 \| S_j^* \tilde{\underline{x}}_j \|_2^2 + d_k^2 \| S_k^* \tilde{\underline{x}}_k \|_2^2).$$

The required rotation is easy to compute explicitly. (See [7] for details.) The rotations are applied in a natural order in sweeps through the matrix, each full sweep comprising $\frac{1}{2}n(n-1)$ rotations. (In practice it is generally more economical to perform only rotations which result in significant improvements at first.) The sweeps are repeated until the improvement in the measure is less than a specified tolerance. The projections of the resulting vectors $\tilde{\underline{x}}_j$ into subspaces S_j, for $j = 1,2,..n$, are then determined explicitly by

$$\underline{x}_j = S_j S_j^* \tilde{\underline{x}}_j / \| S_j^* \tilde{\underline{x}}_j \|_2 .$$

The procedure generates the same iterates for Methods 2 and 3 and is convergent. The computation of each rotation requires $O(n-m)$ (or $O(m)$) operations and the update of matrix $\tilde{X}$ takes $O(n)$ operations. The operation count for a full sweep of the process is, therefore, $O(n^3) + O(n^2m)$, making it of the same order of efficiency as Method 0. In practice, however, Methods 2/3 require fewer sweeps than either Methods 0 or 1 to obtain a solution, and the procedure 2/3 is, in general the least expensive of the techniques.

In cases where a well-conditioned solution is obtainable, the three methods produce similar results and Methods 2/3 are preferred. In cases where the pole assignment problem, itself, is poorly posed, that is, where $\kappa_2(S)/\sqrt{n} >> 1$, then Methods 2/3 do not perform well, since the conditioning $\| D \underline{c} \|_2$ is only bounded by the measure $\nu_4(D)$ when it is sufficiently small. Method 1 is then the more reliable technique to use. We note that in such cases Methods 2/3 can be used to produce a reasonable initial solution for use with Method 1.

We remark that Klein and Moore [14] have proposed a pole placement algorithm which uses the plane rotation technique of Methods 2/3 to make rank-two updates to matrix X with the objective of minimizing the measure $\| D X^{-1} \|_F$, used in Method 1. In their algorithm the optimal rotation which minimizes the measure cannot be obtained explicitly, however, and an "inner" iteration process must be performed to obtain each update. Similar solutions to those constructed by Method 1 appear to be obtained, and comparisons for a particular example are presented in §4.2.

3.4 Implementation

The three steps, Step A, Step X and Step F of the numerical methods described in §3.1-3.3 have been implemented using the system MATLAB [19]. This system uses standard library routines from the software packages LINPACK [3] and EISPACK [23]. For experimental purposes we have developed a set of executive files for use with MATLAB to carry out the various steps of the pole assignment procedures. These files, together with a number of test examples, have been incorporated in a small package, which is available on request from the authors. In the next section of the paper some of the test examples are presented and the numerical results obtained by the pole placement procedures for these examples are reported.

4.0 APPLICATIONS

The four procedures described in Section 3 have been applied to a number of examples collected from the literature. In some cases, the given control system is unstable, and a feedback matrix which stabilizes the system is to be assigned. In other cases, the system is already stable, and the objective of the pole assignment is to move some of the eigenvalues into new positions in the left half-plane, and also to improve the conditioning of the system. To illustrate the behaviour of the methods, we give here the results obtained for two test problems. Other results are given in [7] [11].

4.1 Example 1 CHEMICAL REACTOR (Munro, [17]).

$n = 4 \qquad m = 2$

$$A = \begin{bmatrix} 1.380 & -0.2077 & 6.715 & -5.676 \\ -0.5814 & -4.290 & 0 & 0.6750 \\ 1.067 & 4.273 & -6.654 & 5.893 \\ 0.0480 & 4.273 & 1.343 & -2.104 \end{bmatrix}$$

$$B^T = \begin{bmatrix} 0 & 5.679 & 1.136 & 1.136 \\ 0 & 0 & -3.146 & 0 \end{bmatrix}$$

$EIG(A) = 1.991, \quad 6.351_{10}-2, \quad -5.057\dot{}, \quad -8.666\dot{}.$

This system is unstable and a feedback matrix is required to stabilize the system. We therefore move the two positive real modes into the left-half plane, keeping the original stable modes. We assign the set $\mathcal{L} = \{-0.2, \ -0.5, \ -5.0566..., \ -8.6659...\}$. The condition of S is $\kappa_2(S) = 3.761$ and a feedback system with good conditioning is expected.

In Table 1(a) various measures of the conditioning of the solutions obtained after two sweeps of the procedures are shown. (The weights are taken here as $D = I$.) The magnitude of the gain matrix F, given by $\|F\|_2$, is also given in the Table. In Table 1(b) the same results are shown for the converged solutions, together with the number of iterations required for convergence to a tolerance of 10^{-5}. The computed feedback matrices are given in Appendix 1.

In all cases, well-conditioned solutions, close to the optimal attainable, are determined after only two sweeps of each procedure. The converged results, obtained by Methods 1 and 2/3, have slightly better overall conditioning, measured by $\kappa_2(X)$ or $\|\underline{c}\|_2$, (although with Method 2/3, the maximum condition number $\max_j c_j \equiv \|\underline{c}\|_\infty$ is increased slightly, as is the magnitude $\|F\|_2$ of the gains). Method 1 is very slow to converge in comparison with Method 2/3.

Method 0 is not convergent in this case, and the best result is actually obtained after only one sweep of the procedure.

TABLE 1 Conditioning - Example 1

	(a) Solutions after two sweeps				(b) Solutions at convergence				
Method	$\|\underline{c}\|_\infty$	$\kappa_2(X)$	$\|\underline{c}\|_2$	$\|F\|_2$	$\|\underline{c}\|_\infty$	$\kappa_2(X)$	$\|\underline{c}\|_2$	$\|F\|_2$	No. of sweeps
0	1.82	3.43	3.28	1.47	*	*	*	*	*
1	1.79	3.38	3.27	1.44	1.76	3.32	3.23	1.40	106
2/3	2.36	4.56	3.71	1.16	2.37	4.54	3.68	1.17	6

To demonstrate the robustness of the solutions obtained, the computed feedback matrix F is rounded to three significant figures and the eigenvalues of the resulting closed loop system matrix are calculated. Rounding the gain matrix F corresponds to introducing maximum absolute errors of $\pm\frac{1}{2}10^{-3} \max_{ij} \{|B||F|\}_{ij}$ into the system matrix. For robust solutions such perturbations should only cause errors of the same order of magnitude ($\times$ n) in the poles of the feedback system. For this example, the absolute and percentage errors in each assigned eigenvalue due to the perturbation of the closed loop system matrix are given in Table 2. The absolute errors are of the expected order of magnitude. The eigenvalues of the perturbed system are all within less than 3% of the assigned values and most of the errors are considerably smaller. It may be observed that, for this example, although the overall conditioning obtained with Method 2/3 is somewhat worse than that obtained by Methods 0 and 1, the effects of the perturbation in F is rather less, particularly in the case of the smallest eigenvalue.

TABLE 2 Perturbation Errors - Example 1

	Method 0*		Method 1		Method 2/3	
λ_j	Abs.	%	Abs.	%	Abs.	%
-0.2	0.0051	2.5	0.0012	0.60	0.0005	0.25
-0.5	0.0032	0.64	0.0018	0.35	0.0012	0.25
-5.0566...	0.0039	0.08	0.0004	0.01	0.0015	0.03
-8.6659...	0.0067	0.08	0.0003	0.003	0.0022	0.03

* Result after 2 sweeps.

The conditioning of individual eigenvalues can be controlled by the choice of the weights used in Methods 1 and 2/3. To illustrate the effect of using the weighted measures $\nu_3(D)$ and $\nu_4(D)$, we give in Table 3 the condition numbers c_j for each assigned eigenvalue λ_j obtained after five sweeps of each procedure with various choices of $D = \text{diag}\{d_j\}$. The overall conditioning of the solutions obtained with the weighted measures is worse than in the case $D = I$, as expected, but the conditioning of the smaller eigenvalues is improved.

TABLE 3 Weighted Conditioning - Example 1

	Condition no. c_j Method 1			Condition no. c_j Method 2/3		
Weights d_j	1	$1/\lvert\lambda_j\rvert$	$1/\lvert\lambda_j\rvert^2$	1	$1/\lvert\lambda_j\rvert$	$1/\lvert\lambda_j\rvert^2$
λ_j = -0.2	1.48	1.28	1.33	2.37	1.58	1.55
-0.5	1.76	1.91	2.20	1.08	1.65	1.70
-5.0566..	1.45	1.69	5.00	2.36	1.46	1.45
-8.6659..	1.81	1.96	4.29	1.07	1.79	1.79

4.2 Example 2 DISTILLATION COLUMN (Klein and Moore [14]).

$n = 5 \qquad m = 2$

$$A = \begin{bmatrix} -0.1094 & 0.0628 & 0 & 0 & 0 \\ 1.306 & -2.132 & 0.9807 & 0 & 0 \\ 0 & 1.595 & -3.149 & 1.547 & 0 \\ 0 & 0.0355 & 2.632 & -4.257 & 1.855 \\ 0 & 0.00227 & 0 & 0.1636 & -0.1625 \end{bmatrix}$$

$$B^T = \begin{bmatrix} 0 & 0.0638 & 0.0838 & 0.1004 & 0.0063 \\ 0 & 0 & -0.1396 & -0.206 & -0.0128 \end{bmatrix}$$

EIG(A) = -0.07732, -0.01423, -0.8953, -2.841, -5.982.

We assign the eigenvalue set $\mathcal{L} = \{-0.2, -0.5, -1.0, -1.0 \pm 1.0i\}$, which includes a complex conjugate pair. The condition number $\kappa_2(S) = 57.78$ is large for this choice of poles, and we cannot expect to obtain a very well-conditioned solution to the feedback problem. In Table 4 the various measures of the conditioning of the solutions obtained after five sweeps of Methods 1 and 2/3 are shown, together with the magnitude $\|F\|_2$ of the gain matrix. The computed feedback matrices are given in Appendix 1. The solution obtained by Method 1 is considerably better conditioned than that determined by Method 2/3, although neither is very robust, as expected. We note that the gains are large here, which is anticipated, since $\|A\|_2 = 6.246$, and $\sigma_m\{B\} = 0.057$ and the ratio $\|A\|_2 / \sigma_m\{B\}$ indicates the order of magnitude of $\|F\|_2$.

TABLE 4 Conditioning - Example 2

Method	$\|\underline{c}\|_\infty$	$\kappa_2(X)$	$\|\underline{c}\|_2$	$\|F\|_2$
1	15.3	39.4	22.4	311.5
2/3	30.0	66.1	44.1	283.1

In Table 5 the errors in the assigned poles due to rounding the computed feedback matrices derived by Methods 1 and 2/3 are shown. Corresponding results for solutions to this test problem constructed by two other algorithms, given in [6] and [14], are also shown. The percentage errors due to perturbations in the closed loop systems obtained by Methods 1 and 2/3 are all reasonably small and are comparable with the solution derived in [14]. These three procedures all give solutions which are much less sensitive to general perturbations in the system coefficients than the solution obtained in [6].

TABLE 5 Percentage Errors - Example 2

λ_j	Method 1	Method 2/3	Method [14]	Method [6]
-0.2	1.8%	1.5%	2.5%	73%
-0.5	0.1	0.2	1.2	85
-1.0	0.2	5.0	0.3	40
-1±1i	2.4	1.9	3.0	130

4.3 Summary of Results

For problems where well-conditioned solutions may be expected, that is, where the conditioning $\kappa_2(S)$ is reasonably close to unity, the four methods described here all perform well and lead to robust solutions to the pole assignment problem. The measures of robustness used closely reflect the actual eigenvalue sensitivities, and the condition $\kappa_2(X)$ of the assigned eigenvectors, X, determined by all the methods, are near to the optimal bound $\kappa_2(S)/\sqrt{n}$. The components of the gain matrices F are also as reasonably small as may be expected, given A, B and the choice of the assigned eigenvalue set $\mathcal{L}$. We note that the upper bound given by (22) (Theorem 5) considerably over-estimates the computed magnitude $\|F\|_2$ of the gain matrix, which is found generally to be of the same order of magnitude as $\|A\|_2/\sigma_m\{B\}$. As expected for these problems, small perturbations in the gain matrices lead to proportionately small errors in the assigned poles.

In most cases, for problems where $\kappa_2(S)$ is small, good results are obtained by all three procedures after only two sweeps of the iteration in Step X. Further iteration gives some improvement in robustness, but often leads to small increases in the gains (measured by $\|F\|_2$). For the same initial vectors, Methods 0 and 1 give very similar solutions; Method 0 is rather less expensive per sweep than Method 1, but cannot be guaranteed to converge. Method 2/3 is comparatively less expensive than Method 1 and generally converges more quickly. The solutions constructed by Methods 2/3 are frequently similar to those produced by Methods 0 and 1; the results can, however, be quite different, whilst being equally robust, as shown here in Example 1.

For problems where $\kappa_2(S)$ is large, that is, where well-conditioned solutions cannot be achieved, the methods perform more erratically. In these cases Methods 0 and 1 generally give more robust solutions than Methods 2/3,

which is expected, since minimizing $\nu_4(D)$ does not now necessarily minimize an upper bound on the conditioning $\|\underline{c}\|_2$ of the closed loop system. The methods all produce acceptable solutions, however, within the limitations of the problem. To obtain more robust solutions in cases where the conditioning of matrix S is poor, it is advisable to modify the set $\mathcal{L}$ of eigenvalues to be assigned so as to reduce $\kappa_2(S)$, and hence improve the feasible conditioning of the eigenvector selection problem.

5. CONCLUSIONS

The problem of pole assignment by state feedback for multivariable control systems is essentially under-determined. We demonstrate here that the extra degrees of freedom in the problem may be used to determine a robust, or well-conditioned, solution such as to minimize the sensitivities of the closed loop poles to perturbations in the system and gain matrices. For such robust solutions it is shown that bounds on the (mean square) magnitude of the closed loop transient response and on the norm of the feedback gain matrix are also minimized, and that a lower bound on the stability margin is maximized. A measure of the optimal conditioning that may be expected for a particular system with a given set of closed loop poles is described and used to assess the suitability of the given poles for assignment.

Four novel numerical methods are derived for constructing robust, well-conditioned solutions to the state feedback pole placement problem. The methods are applied to practical test examples, and numerical results are presented and discussed. The tests indicate that the methods are stable and efficient. In cases where well-conditioned solutions may be expected, near optimal results are obtained. Introducing perturbations in the computed gain matrices leads only to correspondingly small errors in the assigned poles. The methods are based on different principles and exhibit different behaviour, however, and certain of the procedures may be regarded as more reliable than others.

Generalizations of these methods for degenerate systems and for the output feedback problem are currently being developed. Certain necessary theoretical results have already been derived and numerical techniques are now available [9] [10] [2]. Extensions to techniques for modifying the locations of the assigned closed loop poles to improve further their insensitivity are also being examined.

6. REFERENCES

[1] Barnett, S., Introduction to mathematical control theory. Oxford University Press, Oxford (1975).

[2] Chu, E. K-w. and Nichols, N. K., "Robust pole assignment by output feedback," Proc. 4th IMA Conf. on Control Theory, Sept., 1984 (to appear).

[3] Dongarra, J. J., Moler, C. B., Bunch, J. R. and Stewart, G. W., LINPACK User's Guide. SIAM, Philadelphia (1979).

[4] Fahmy M. M. and O'Reilly, J., "On eigenstructure assignment in linear multivariable systems", IEEE Trans. Auto Control AC-27. 690-693 (1982).

[5] Flamm, D. S., "A new proof of Rosenbrock's theorem on pole assignment," IEEE Trans. Auto. Control AC-25, 1128-1133 (1980).

[6] Gourishankar, V. and Ramar, K. "Pole assignment with minimum eigenvalue sensitivity to plant variations," Int. J. Control., 23, 493-504 (1976).

[7] Kautsky, J. and Nichols, N. K., Robust eigenstructure assignment in state feedback control, Univ. of Reading, Dept. of Math., Num. Anal. Rpt. NA/2/83 (1983).

[8] _____, "Matrix factorization and some applications in linear systems theory," IEE Colloquium on Reliable Numerical Procedures in Control System Design, IEE Digest No. 1983/81, pp.5/1-5/3 (1983).

[9] _____, "Pole assignment in degenerate linear systems - necessary conditions for regular solutions," (to appear).

[10] Kautsky, J., Nichols, N. K. and Chu, E., "Robust pole assignment for descriptor systems," presented at the IMA/SERC Meeting on Control Theory at Warwick University, July, 1983. (To appear as Univ. of Reading, Dept. of Math., NA Rpt.)

[11] Kautsky, J., Nichols, N. K., Van Dooren, P. and Fletcher, L., "Numerical Methods for robust eigenstructure assignment in control system design," Numerical Treatment of Inverse Problems for Differential and Integral Equations, (ed. P. Deuflhard, E. Hairer) Birkhauser/Boston pp. 171-178, (1983).

[12] Kautsky, J., Nichols, N. K. and Van Dooren, P., "Numerical Methods for selecting a well-conditioned set of vectors from given subspaces." (to appear).

[13] Klein, G. and Moore, B. C., "Eigenvalue-generalized eigenvector assignment with state feedback," IEEE Trans. Auto. Control AC-22, 140-141 (1977).

[14] _____, "Feedback design of linear systems by eigenvector selection," Dept. of Mech. Eng., Columbia Univ., pre-print (1982).

[15] Maki, M. C. and Van de Vegte, J., "Optimization of multiple-input systems with assigned poles," IEEE Trans. Auto. Control AC-19, 130-133 (1974).

[16] Mayne, D. Q. and Murdock, P., "Model control of linear time invariant systems," Int. J. Control, 11, 223-227 (1970).

[17] Minimis, G. S. and Paige, C. C., "An algorithm for pole assignment of time invariant linear systems," Int. J. Control, 35, 341-354 (1982).

[18] _____, "An algorithm for pole assignment of time invariant multi-input linear systems," Proc. IEEE Conf. Dec. & Contr., pp. 62-67 (1982).

[19] Moler, C. B. MATLAB User's Guide. University of New Mexico Dept. of Computer Science (1981).

[20] Moore, B. C., "On the flexibility offered by state feedback in multivariable systems beyond closed loop eigenvalue assignment," IEEE Trans. Auto. Control. AC-21, 689-692 (1976).

[21] Munro, N., "Pole assignment," Proc. IEE 126, 549-555 (1979).

[22] Porter, B. and D'Azzo, J. J., "Closed-loop eigenstructure assignment by state feedback in multivariable linear systems," Int. J. Control, 27, 487-492 (1978).

[23] Smith, B. T., Boyle, J. M., Dongarra, J. J., Garbow, B. S., Ikebe, Y., Klema, V. C., and Moler, C. B., Matrix Eigensystem Routines - EISPACK Guide. Lecture Notes in Computer Science, Vol. 6, 2nd edition, Springer-Verlag (1976).

[24] Smith, R. A., "The condition numbers of the matrix eigenvalue problem," Numer. Math. 10, 232-240 (1967).

[25] Thompson, R. C., "Principal submatrices IX : interlacing inequalities for singular values of submatrices," Lin. Alg. & Appl., 5, 1-12 (1972).

[26] Van Dooren, P., "The generalized eigenstructure problem in linear system theory," IEEE Trans. Auto. Contr., AC-26, 111-129 (1981).

[27] Varga, A., "A Schur method for pole assignment," IEEE Trans. Auto. Control AC-26, 517-519 (1981).

[28] Wilkinson, J. H., The algebraic eigenvalue problem. Oxford University Press, Oxford (1965).

[29] Wonham, W. M., Linear multivariable control: a geometric approach. (2nd ed.) Springer, New York, (1979).

APPENDIX 1 Computed Feedback Matrices

Example 1

Method 0*

$$F = \begin{bmatrix} 0.23416 & -0.11423 & 0.31574 & -0.26872 \\ 1.1673 & -0.28830 & 0.68632 & -0.24241 \end{bmatrix}$$

Method 1

$$F = \begin{bmatrix} 0.14454 & -0.051421 & 0.13265 & -0.12868 \\ 1.1101 & -0.033345 & 0.78416 & -0.23384 \end{bmatrix}$$

Method 2/3

$$F = \begin{bmatrix} 0.10277 & -0.63333 & -0.11872 & 0.14632 \\ 0.83615 & 0.52704 & -0.25775 & 0.54269 \end{bmatrix}$$

*Result after two sweeps.

Example 2

Method 1

$$F = \begin{bmatrix} -47.690 & 102.01 & -213.70 & 179.86 & -42.552 \\ -22.596 & 30.633 & -48.077 & 33.799 & 2.2776 \end{bmatrix}$$

Method 2/3

$$F = \begin{bmatrix} -159.68 & 69.844 & -165.24 & 125.23 & -45.748 \\ -99.348 & 7.9892 & -14.158 & -5.9382 & -1.2542 \end{bmatrix}$$