ILL-CONDITIONED EIGENSYSTEMS AND THE COMPUTATION OF THE JORDAN CANONICAL FORM

by

G. H. Golub
J. H. Wilkinson

STAN-CS-75-478
FEBRUARY 1975

COMPUTER SCIENCE DEPARTMENT
School of Humanities and Sciences
STANFORD UNIVERSITY

ILL-CONDITIONED EIGENSYSTEMS AND THE COMPUTATION OF THE JORDAN CANONICAL FORM

BY G H GOLUB AND J H WILKINSON

ABSTRACT

The solution of the complete eigenvalue problem for a non-normal matrix A presents severe practical difficulties when A is defective or close to a defective matrix. Moreover in the presence of rounding errors one cannot even determine whether or not a matrix is defective. Several of the more stable methods for computing the Jordan canonical form are discussed together with the alternative approach of computing well-defined bases (usually orthogonal) of the relevant invariant subspaces.

* The work of G H Golub was supported in part by National Science Foundation, GJ35135X and Atomic Energy Commission, AT(04-3)-326PA #30.

ILL-CONDITIONED EIGENSYSTEMS AND THE COMPUTATION OF THE JORDAN CANONICAL FORM

BY GH GOLUBANDJ HWILKINSON

1 INTRODUCTION

From the standpoint of classical algebra the algebraic eigenvalue problem has been completely solved. The problem is the subject of classical similarity theory and the fundamental result is embodied in the Jordan canonical form (J.c.f.). Most mathematicians encounter similarity theory in an abstract setting but since we are concerned here with practical algorithms we first review the basic result purely in matrix terms.

The J.c.f. is described with reference to matrices known as elementary Jordan blocks. A Jordan block of order r associated with an eigenvalue λ_i will be denoted by $J_r(\lambda_i)$ and its general form is adequately illustrated by the definition

$$J_4(\lambda_i) = \begin{bmatrix} \lambda_i & 0 & 0 & 0 \\ 1 & \lambda_i & 0 & 0 \\ 0 & 1 & \lambda_i & 0 \\ 0 & 0 & 1 & \lambda_i \end{bmatrix} . \qquad (1.1)$$

The basic theorem is that given any nx n matrix with complex elements there exists a *non-singular* matrix X such that

$$X^{-1}AX = J \ , \quad AX = XJ \ , \qquad (1.2)$$

where 3, the J.c.f. of A, is block diagonal, each diagonal matrix being an elementary Jordan block. Apart from the ordering of the blocks along the diagonal of J (which can he arbitrary) the 3.c.f. is unique, although X is far from unique. It will be convenient to order the block in some stahdard way. Unless reference is made to the contrary we assume that the $|\lambda_i|$ are in order of non-increasing magnitude and that the blocks associated with a specific λ_i are ordered to be of non-decreasing size.

Thus if the matrix h of order 12 has only 2 distinct eigenvalues λ_1 and λ_2 with $|\lambda_1| \geq |\lambda_2|$ and λ_1 is associated with 2 blocks of order 2 and one of order 3 while λ_2 is associated with one block of order 2 and one of order 3 its J.c.f. will be presented in the form

$$\left[\begin{array}{ccccc|c} J_2(\lambda_1) & & & & & \\ & J_2(\lambda_1) & & & & \\ & & J_3(\lambda_1) & & & \\ & & & J_2(\lambda_2) & & \\ & & & & J_3(\lambda_2) & \end{array}\right] . \qquad (1.3)$$

Here λ_1 is an eigenvalue of multiplicity 2+2+3=7 and λ_2 of multiplicity 2+3=5. The example illustrates that there may be more than one block or a given dimension associated with a specific A_i.

Let us consider the significance of the existence of a block $J_r(\lambda_i)$ in J, where $J_r(\lambda_i)$ starts in rows and columns s and ends in rows and columns t and

$$r = t - s + 1 . \qquad (1.4)$$

Equating columns s to t on both sides of equation (1.2)

$$\begin{array}{ll} Ax_s = \lambda_i x_s , & (A-\lambda_i I)x_s = 0 \\ Ax_{s+1} = \lambda_i x_{s+1} + x_s , & (A-\lambda_i I)x_{s+1} = x_s \\ Ax_{s+2} = \lambda_i x_{s+2} + x_{s+1} , & (A-\lambda_i I)x_{s+2} = x_{s+1} \\ \cdots\cdots\cdots\cdots & \cdots\cdots\cdots\cdots \\ Ax_t = \lambda_i x_t + x_{t-1} , & (A-\lambda_i I)x_t = x_{t-1} , \end{array} \qquad (1.5)$$

where, here and later, we shall denote the ith column of a matrix X (say) by x_i. The first of these relations implies that x_s is an eigenvector corresponding to λ_i. The remaining equations imply that

$$(A-\lambda_i I)^2 x_{s+1} = 0, \ (A-\lambda_i I)^3 x_{s+2} = 0, \ \ldots \ (A-\lambda_i I)^{t-s+1} x_t \equiv (A-\lambda_i I)^r x_t = 0 . \qquad (1.6)$$

Notice that in general the x_{s+i} satisfy the relations

$$(A-\lambda_i I)^{p-1} x_{s+p-1} = x_s \neq 0 \text{ and } (A-\lambda_i I)^p x_{s+p-1} = 0 . \qquad (1.7)$$

We shall refer to any vector x such that $(A-\lambda I)^{p-1}x \neq 0$, $(A-\lambda I)^p x = 0$ as a vector of grade p and for uniformity an eigenvector becomes a vector of grade 1. It is evident, for example, that

$$(A-\lambda I)^2 (a_2 x_{s+2} + a_1 x_{s+1} + a_0 x_s) = a_2 x_s, \ (A-\lambda I)^3 (a_2 x_{s+2} + a_1 x_{s+1} + a_0 x_s) = 0 \qquad (1.8)$$

so that $a_2 x_{s+2} + a_1 x_{s+1} + a_0 x_s$ is a vector of grade 3 for all a_i provided $a_2 \neq 0$. The vectors x_{s+i} arising in the Jordan canonical reduction are special in that they satisfy the chain relations (1.5). We shall refer to the vectors of grades 1, 2, 3, ... associated with a Jordan block as principal vectors of grades 1, 2, 3

Clearly $\det(\lambda I - J_r(\lambda_i)) = (\lambda - \lambda_i)^r$ and we may associate such a polynomial with each of the blocks in the J.c.f. These polynomials are called the elementary divisors of A. An enumeration of the elementary divisors gives a unique specification of the J.c.f. Corresponding to a Jordan block of dimension unity the elementary divisor is $(\lambda - \lambda_i)$, ie it is linear. If all the Jordan blocks in the J.c.f. are of dimension unity then the J.c.f. is strictly diagonal, the matrix has n independent eigenvectors given by the columns of X and all the elementary divisors are linear. These four properties are fully equivalent to each other. Notice that if there are n distinct λ_i then all the blocks are necessarily of dimension unity. Departure from strictdiagonal form can occur only if there is at least one multiple eigenvalue, though even in this case the J.c.f. can be diagonal.

A matrix is said to be defective if the J.c.f. in not strictly diagonal. In this case at least one elementary divisor is non-linear and the number of independent eigenvectors is less than n; the remaining columns of X are principal vectors of the appropriate grades.

A matrix is said to bc derogatory if there is at least one λ_i which is associated with more than one diagonal block in the J.c.f. If such a λ_i is *associated* with k different blocks then there are precisely k independent eigenvectors associated with λ_i.

It. should be emphasized that a matrix may be defective without being derogatory and vice versa or it can be both defective and derogatory. if the λ_i are distinct it cannot be either. If A is normal (including Hermitian, skew Hermitian or unitary) then its J.c.f. is always strictly diagonal and the X producing the J.c.f. may be *chosen to* be unitary. A normal matrix witii a multiple eigenvalue is therefore derogatory but not defective.

We do not report on numerical experiments in this paper although many of the algorithms described have been implemented with success. It is the aim of this paper to emphasize the problems associated with computing invariant subspaces and to stimulate research in this area. We have not attempted to be encyclopaedic (despite the length of the paper) but state those principles which we feel are of importance in this area.

a LINEAR DIFFERENTIAL EQUATIONS AND THE J.c.f.

The practical significance of the J.c.f. of a matrix A is that it provides the general solution of the associated system of linear differential equations with constant coefficients defined by

$$\frac{du}{dt} = A u \ , \tag{2.1}$$

where u is a vector of order n. Under the linear transformation u = Xv the equation becomes

$$X \frac{dv}{dt} = AXv \text{ or } \frac{dv}{dt} = X^{-1} A X v = Jv, \tag{2.2}$$

Hence the J.c.f. gives a simplified version of the original system. If J is strictly

diagonal (ie A is not defective) the transformed system is

$$\frac{dv_i}{dt} = \lambda_i v_i \qquad (2.3)$$

and in terms of the variables v_i the equations are completely decoupled. The general solution is

$$v_i = v_i^{(o)} e^{\lambda_i t} \qquad u = \sum v_i^{(o)} x_i e^{\lambda_i t} \qquad (2.4)$$

and is therefore directly expressible in terms of the n independent eigenvectors x_i and n independent constants $v_i^{(o)}$, the initial values of the v_i. Notice that the analysis is not affected by any multiplicities in the λ_i provided J is strictly diagonal. An eigenvalue λ_i of multiplicity r is then associated with r independent eigenvectors and r arbitrary $v_j^{(o)}$. When A is defective the linear transformation does not give a complete decoupling of the equations, but there is a decoupling of those equations involving the v_i associated with each specific block from those associated with all other v_j. The general solution is most readily exposed in terms of the concept of the "exponential" of a matrix. We define exp (B) by the relation

$$\exp(B) = I + \frac{1}{1!} B + \frac{1}{2!} B^2 + \ldots + \frac{1}{r!} B^r + \ldots \qquad (7.5)$$

the matrix series being convergent for all B. The solution of (2.1) such that $u = u^{(o)}$ when $t = 0$ is given by

$$u = \exp(At) u^{(o)}. \qquad (2.6)$$

From the series expansion it will readily be verified that

$$\exp(XBX^{-1}t) = X \exp(Bt) X^{-1} \qquad (2.7)$$

and hence the solution of (2.1) is

$$u = X \exp(Jt) X^{-1} u^{(o)}$$

or $$v = \exp(Jt) v^{(o)} \text{ where } v = X^{-1} u . \qquad (2.8)$$

If $J_r(\lambda_i)$ is a typical block in J then exp(Jt) has the same block structure, with $\exp(J_r(\lambda_i)t)$ in place of each $J_r(\lambda_i)$ and the form of $\exp(J_r(\lambda_i)t)$ is fully illustrated by the relation

$$\exp(J_4(\lambda_i)t) = \exp(\lambda_i t) \begin{bmatrix} 1 & t/1! & t^2/2! & t^3/3! \\ & 1 & t/1! & t^2/2! \\ & & 1 & t/1! \\ & & & 1 \end{bmatrix} . \qquad (2.9)$$

Hence on transforming back from the v coordinates to the u coordinates the solution corresponding to the initial value problem is again given in terms of the vectors x_i but corresponding to a Jordan block $J_r(\lambda_i)$, terms involving $\exp(\lambda_i t)\, t^s/s!$ $(s = 0, \ldots, r-1)$ arise.

This discussion gives the impression that the theoretical significance of the J.c.f. is fully matched by its practical importance since it is precisely because of its relationship to the solution of systems of linear differential equations that the algebraic eigenvalue problem occupies such a prominent position in practical applied mathematics. The principal objective of the remainder of this paper is to show the basic limitations of the J.c.f. from the point of view of practical computation and indeed to cast doubt on the advisability of trying to determine it.

Before proceeding it is useful to consider the degree of arbitrariness in the matrix X involved in the reduction to J.c.f. If the λ_i are distinct, J is diagonal and the x_i are the unique eigenvectors. The only degree of arbitrariness is in the scaling of the x_i. We have

$$D^{-1}X^{-1}AXD = D^{-1}JD = J \qquad (2.10)$$

where D is a non-singular diagonal matrix.

Turningnow to the case when A has a single block of dimension r we see that there is already a wide freedom of choice in X. Suppose for illustration that there is a block of order4associated with λ_i, then from equations (1.5) we see, writing $B \equiv A-\lambda_i I$, that

$$\begin{aligned} B(ax_{s+3} + bx_{s+2} + cx_{s+1} + dx_s) &= ax_{s+2} + bx_{s+1} + cx_s \\ B(ax_{s+2} + bx_{s+1} + cx_s) &= ax_{s+1} + bx_s \\ B(ax_{s+1} + bx_s) &= ax_s \\ B(ax_s) &= 0 , \end{aligned} \qquad (2.11)$$

where the a, b, c, d arc arbitrary but $a \neq 0$. Hence the chain of *x-actors* $x_{s+3}, x_{s+2}, x_{s+1}, x_s$ may be replaced by the chain of vectors given in (2.11) and on this account X may he replaced by XP where

$$P = \begin{bmatrix} I & & \\ & \begin{matrix} a & b & c & d \\ & a & b & c \\ & & a & b \\ & & & a \end{matrix} & \\ & & I \end{bmatrix} . \qquad (2.12)$$

The derogatory case, ie when there is more than one block associated with a given λ_i may be illustrated by the case when there are blocks of orders 2 and 3 starting in pssitions s and t respectively. From the tno chains

$$\begin{aligned} Bxs &= 0 & Bx_t &= 0 \\ Bx_{s+1} &= x_s & Bx_{t+1} &= x_t \\ Bx_{st2} &= x_{stl} \end{aligned} \qquad (2.13)$$

the two generalised chains defined by

$$\begin{aligned} B(ax_{s+2} + bx_{s+1} + cx_s + dx_{t+1} + ex_t) &= ax_{s+1} + bx_s + dx_t \\ B(ax_{s+1} + bx_s + dx_t) &= ax_s \\ B(ax_s) &= 0 \\ \text{and} & \\ B(fx_{st1} + gx_s + hx_{t+1} + ix_t) &= fx_s + hx_t \\ B(fx_s + hx_t) &= 0 \end{aligned} \qquad (2.14)$$

may be derived, where the c, b, ... i are arbitrary except that $a \neq 0$, $h \neq 0$, and X may be varied correspondingly.

3 SENSITIVITY OF THE EIGENVALUES OF A DEFECTIVE MATRIX

Blocks of dimension greater than unity in the J.c.f. can emerge, if at all, only as the result of the *presence* of multiple eigenvalues. In the classical theory there is a clear cut distinction between equal and unequal eigenvalues. In practice the situation is very different since a matrix may not be representable exactly in the computer and in any case rounding errors are, in general, involved in computing transformations. Let us consider the effect of small perturbations on the eigenvalues of en elementary Jordan block $J_r(\lambda_i)$. If the zero element in position (r, 1) is replaced by ε the characteristic equation

$$(\lambda-\lambda_i)^r = \varepsilon \qquad (3.1)$$

and the multiple eigenvalue λ_i is replaced by r distinct eigenvalue $\lambda_i + \varepsilon^{1/r}(\cos\frac{2s\pi}{r} + i\sin\frac{2s\pi}{r})$ (s = 0, ... r-1). Suppose λ_i is of order unity, $r = 10$ and $\varepsilon = 10^{-10}$. Then the separation of the perturbed roots is of order 10^{-1} and they cannot in any reasonable sense be regarded as "close".

In practice we have todiagnose multiplicities and the degree of defectiveness from computed eigenvalues. When these are determined by a very stable algorithm

we cannot rely on any of them being recognisably 'close', even when the given A really does have some multiple eigenvalues. When A has an elementary divisor of high degree this danger appears to be particularly severe.

However, *even* this remarksomewhat oversimplifies the situation. One tends to be seduced by the simplicity of the J.c.f. and as a result to attach too much significance **to every detail of it.** When attempting to construct "difficult" matrices for practical experiments it is common to take a nsn-diagonal J.c.f., subject it to some exact similarity transformation and then to regard the resulting matrix as wholly typical of a defective matrix.

Fut this is to attach too much significance to the unity elements in the Jordan blocks. If $D = \mathrm{diag}(d_i)$ is any non-singular diagonal matrix then from (1.2) we have

$$D^{-1}X^{-1}AXD = D^{-1}JD \ , \qquad (3.2)$$

Hence if J has a unity element in position $(p, p+1)$ the matrix $D^{-1}JD$ has $d_p^{-1} d_{p+1}$ in this position; by a suitable choice of the d_i the unity element? may be given arbitrary values. The choice of the unity elements in the J.c.f.is purely for notational convenience. However, in classical mathematics we can make a sharp distinction between zero and non-zero elements, a luxury we are denied in practical computation. We refer to a matrix as being in quasi-J.c.f. if the only difference from strict J.c.f. is that some of the super-diagonals have values other than unity.

It is possible for a matrix A to be highly defective without its eigenvalues being unduly sensitive. Suppose, for example, that A is such that there is an orthogonal matrix X for which

$$X^{-1}AX = \tilde{J} \ , \qquad (3.3)$$

where $\tilde{J}$ is of quasi-J.c.f. in which non-zero super-diagonal elements are all 10^{-10}.

Perturbations of order 10^{-10} in J (nhirh correspond Co perturbatiens of order 10^{-10} in A since X is orthogonal) produce perturbations nf order 10^{-10} at most in the eigenvalues. If $\|A_2\|$ is of the order of unity, then from the point of view of 10 digit decimal computation the eigenvalues of A are not at all sensitive. One cannot even rely on defectiveness being characterised by sensitivity of the corresponding eigenvalues. Nevertheless it is true that $\frac{\partial \lambda_i}{\partial \varepsilon} = O(\varepsilon^{\frac{1}{r}-1})$ for some perturbations when J has 2 block of order r *and* hence $\frac{\partial \lambda_i}{\partial \varepsilon} \to \infty$ as $\varepsilon \to 0$. This means that if we are prepared to extend the precision of computation indefinitely we shall ultimately gain only one figure of accuracy for r extra figures of precision.

At this stage one might ask what is the 'natural' quasi-J.c.f. for computational purposes. A reasonable definition is that it is that $\tilde{J}$ for which the corresponding $\|X\|_2 \|X^{-1}\|_2 = \kappa(X)$ is a minimum. If this $\tilde{J}$ has super-diagonal elements which are all small relative to $\|\tilde{J}\|_2$ the matrix A will not have sensitive eigenvalues.

As a final result relating small eigenvalues and small singular values we note the following theorem.

Let A be an $n \times n$ matrix with $\lambda_n = \varepsilon$ and $|\lambda_n| \leq |\lambda_j|$ and such that there are p Jordan blocks of dimensions $k_1, k_2, \ldots, k_p$ with $k_1 \leq k_2 \leq \ldots \leq k_p$ associated nith λ_n. Then if $A = XJX^{-1}$

$$\sigma_{n-j+1}(A) \leq \|X\|_2 \|X^{-1}\|_2 |\varepsilon|^{k_{p-j+1}} + O(|\varepsilon|^{k_{p-j+1}+2})$$

$$(j = 1,2,\ldots,p) \qquad (3.4)$$

Proof

$$\sigma_{n-j+1}(A) = \sigma_{n-j+1}(XJX^{-1}j \leq \sigma_1(X)\,\sigma_{n-j+1}(JX^{-1})$$

$$\leq \sigma_1(X)\,\sigma_1(X^{-1})\,\sigma_{n-j+1}(J) . \qquad (3.5)$$

Since the singular values of J arc given by $\left[\lambda_i(JJ^T)\right]^{\frac{1}{2}}$ it is obvious that they are the singular values of the elementary Jordan blocks. Consider the $k \times k$ block

$$K = \begin{bmatrix} \varepsilon & 1 & & & \\ & \varepsilon & 1 & & \\ & & & & \\ & & & 1 & \\ & & & \varepsilon & \end{bmatrix} . \qquad (3.6)$$

From the form of KK^T, k-l of the singular value of close to unity and since their product is ε^k the reasining singular value is $O(\varepsilon^k)$. In fact

$$\sigma_k(K) = \min_{x \neq 0} \frac{\|Kx\|_2}{\|x\|_2} \qquad (3.7)$$

and taking $\tilde{x}^T = (1, -\varepsilon, \varepsilon^2, \ldots, (-1)^{k-1}\varepsilon^{k-1})$ we have

$$\sigma_k(K) = |\varepsilon|^k + O(|\varepsilon|^{k+2}) . \qquad (3.8)$$

The result is thus establi-bed. Note that although we have shown that the singular values are small, we have not and cannot show that the elements of the corresponding singular vectors are correspondingly small.

4 ILL-CONDITIONED EIGENVALUES

Since in practice it will usually be impossible to determine whether a matrixhas exactly equal eigenvalues it is necessary to consider the problem of the sensitivity of a simple eigenvalue with respect to perturbations in A. If J is the J.c.f. we have

$$AX = XJ \,, \quad ZA = JZ \,, \quad Z = X^{-1} . \qquad (4.1)$$

When A, is 3 simple eigenvalue, x_1 is the corresponding right-hand eigenvector and

$$Ax_1 = \lambda_1 x_1 . \qquad (4.2)$$

If z_1^T is the first row of Z then

$$z_1^T A = z_1^T \lambda_1 . \qquad (4.3)$$

It is customary to define the left-hand eigenvector y_1 of A corresponding to λ_1 as the vector satisfying

$$y_1^H A = y_1^H \lambda_1 \tag{4.4}$$

and hence if we write $Y = Z^H$ the first column of Y gives this eigenvector and

$$Y^H X = I \tag{4.5}$$

Consider now the corresponding eigenvalue $\lambda_1(\varepsilon)$ and right-hand eigenvector $x_1(\varepsilon)$ of $A+\varepsilon B$ where $\|B\|_2 = 1$. For sufficiently small ε it is easy to show that $X_1(\varepsilon)$ and $x_1(\varepsilon)$ may be expanded as convergent power series

$$\lambda_1(\varepsilon) = \lambda_1 + p_1\varepsilon + p_2\varepsilon^2 + \ldots, \quad x_1(\varepsilon) = x_1 + v_1\varepsilon + v_2\varepsilon^2 + \ldots \tag{4.6}$$

where the v_i lie in the space spanned by $x_2, \ldots x_n$. (Note that in general these x_i will include principal vectors which are not eigenvectors). Equating coefficients of ε in the relation

$$(A+\varepsilon B)(x_1 + v_1\varepsilon + \ldots) = (\lambda_1 + p_1\varepsilon + \ldots)(x_1 + v_1\varepsilon + \ldots) \tag{4.7}$$

gives

$$Bx_1 + Av_1 = \lambda_1 v_1 + p_1 x_1. \tag{4.8}$$

Now both v_1 and Av_1 lie in the space spanned by $x_2, \ldots x_n$ and from (4.5) $y_1^H x_i = 0$ $(i = 2, \ldots n)$. Hence premultiplying (4.6) by y_1^H

$$p_1 = y_1^H Bx_1 / y_1^H x_1 . \tag{4.9}$$

As derived above $y_1^H x_1 = 1$ but clearly in (4.9) x_1 and y_1 may be arbitrarily scaled

and it is convenient computationally to have $\|x_1\|_2 = \|y_1\|_2 = 1$. In this case $y_1^H x_1 = s_1$ (in the notation of [25]) where s_1 is the cosine of the angle between x_1 and y_1. From (4.9)

$$\left|\frac{\partial \lambda_1}{\partial \varepsilon}\right|_{\varepsilon=0} = |p_1| \leq \|y_1\|_2 \|B\|_2 \|x_1\|_2 / s_1 = 1/|s_1| . \tag{4.10}$$

The derivative it finite for any 'direction' of B. This is in contrast to the case where λ_i is associated with a defective matrix when $\left|\frac{\partial \lambda_i}{\partial \varepsilon}\right|_{\varepsilon=0} = \infty$. This latter result is in agreement with (4.10) since the left-hand and right-hand eigenvectors are orthogonal corresponding to a 'defective' λ_i. The bound in (4.10) is attained where $B = y_1 x_1^H$ since then

$$y_1^H B x_1 = y_1^H y_1 x_1^H x_1 = 1 . \tag{4.11}$$

Further, taking $B = e^{i\theta} y_1 x_1^H$ we can make $\left(\frac{\partial \lambda_1}{\partial \varepsilon}\right)_{\varepsilon=0}$ have any required phase.

There is one very unsatisfactory feature of the above analysis. The quantity s_i is not invariant with respect to diagonal similarity transformation. Consider the matrix

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \tag{4.12}$$

with

$$\lambda_1 = 3, \ \lambda_2 = 1, \ y_1^H = [1, 1]/2^{\frac{1}{2}}, \ x_1^H = [1, 1]/2^{\frac{1}{2}}, \ s_1 = 1. \tag{4.13}$$

The eigenvalue λ_1 is therefore very well-conditioned, as indeed are all eigenvalues of all normal matrices. However we have

$$D^{-1}AD = \begin{bmatrix} 2 & \alpha \\ \alpha^{-1} & 2 \end{bmatrix} \quad \text{where } D = \begin{bmatrix} 1 & \\ & \alpha \end{bmatrix} \tag{4.14}$$

and now

$$y_1^H = [1, \alpha]/(1+\alpha^2)^{\frac{1}{2}} \ , \ x_1^H = [\alpha\ 1]/(1+\alpha^2)^{\frac{1}{2}} \ , \ s_1 = 2\alpha/(1+\alpha^2). \qquad (4.15)$$

Hence we may make s_1 arbitrarily small by taking a sufficiently large or sufficiently small. It is clear that a small s_i induced in this way is a very artificial phenomenon. In this example when s_1 is small $\|D^{-1}AD\|_2 >> \|A\|_2$. In practice the relevant values of s_1 are those for $D^{-1}AD$ where D has been chosen so that $\|D^{-1}AD\|_2$ is a minimum. Reducing this norm to a true minimum is not vital and in practice the process of balancing described by Parlett and Reinsch in [12] is usually adequate.

High sensitivity of an eigenvalue λ_i has now been encountered in two different contexts, first when λ_i is associated with defectiveness and secondly when a value of s_i is small. We now show that when an s_i is small, A is necessarily relatively close to a matrix with a multiple eigenvalue. Let

$$Ax_1 = \lambda_1 x_1 \ , \ y_1^H A = y_1^H \lambda_1 \ , \ s_1 = y_1^H x \ , \ \text{with } \|x_1\|_2 = \|y_1\|_2 = 1 \qquad (4.16)$$

and suppose P is a unitary matrix such that $Px_1 = e_1$ where $e_1^T = (1, 0, \ldots, 0)$. Then

$$PAP^H Px_1 = \lambda_1 Px_1 \ , \ (PAP^H)e_1 = \lambda_1 e_1 \qquad (4.17)$$

and $B = PAP^H$ must be of the form

$$B = \left[\begin{array}{c|c} \lambda_1 & b_1^H \\ \hline 0 & B_1 \end{array}\right]. \qquad (4.18)$$

Further

$$s_1 = y_1^H x_1 = (y_1^H P^H)(Px_1) = (Py_1)^H e_1 \qquad (4.19)$$

and writing $Py_1 = p_1$ we have

$$p_1^H B = p_1^H PAP^H = y_1^H A P^H = \lambda_1 (y_1^H P^H) = \lambda_1 p_1^H \qquad (4.20)$$

while

$$s_1 = p_1^H e_1 = \bar{p}_{11}. \qquad (4.21)'$$

Hence if we write $p_1^H = (\bar{p}_{11} \mid v^H)$ where v is of order n-1

$$\bar{p}_{11} b_1^H + v^H B_1 = \lambda_1 v^H$$

$$v^H(B_1 - \lambda_1 I + \bar{p}_{11} \frac{v b_1^H}{v^H v}) = 0 \qquad (4.22)$$

ie the matrix $B_1 + \bar{p}_{11} \frac{v b_1^H}{v^H v}$ has A, as an eigenvalue and v as a left-hand eigenvector. NOW

$$\left\| \bar{p}_{11} \left(\frac{v b_1^H}{v^H v} \right) \right\| \leq \left| \bar{p}_{11} \right| \frac{\|v\| \|b_1\|}{\|v\|^2} = \frac{|s_1| \|b_1\|}{(1-s_1^2)^{\frac{1}{2}}} \leq \frac{s_1 \|B\|_2}{(1-s_1^2)^{\frac{1}{2}}} = \frac{s_1 \|A\|_2}{(1-s_1^2)^{\frac{1}{2}}} \qquad (4.23)$$

and when s_1 is small a small relative perturbation in B converts λ_1 into an eigenvalue of multiplicity at least two. Since the l_2 norm is invariant with respect to unitary transformations the same remark is true of A. By a similar argument Kahan in an unpublished paper has shown that the denominator $(1-s^2)^{\frac{1}{2}}$ may be replaced by 1 in the final bound. However the above argument shows that the relevant bound is $|s_1| \|b_1\|_2 / (1-s_1^2)^{\frac{1}{2}}$ and in replacing $\|b_1\|_2$ by $\|B\|_2$ and hence by $\|A\|_2$ the result is weakened. When A is normal, B is also normal and $b_1 = 0$. Hence if $|s_1| < 1$ for a normal matrix λ_1 must already be a multiple eigenvalue. This is otherwise obvious since if λ_1 is a simple eigenvalue of a normal matrix $y_1 x$ and $s_1 = 1$. The bound we have given is, in general, a considerable improvement on the bound given by Ruhe [16].

5 ALMOST LINEARLY DEPENDENT EIGENVECTORS

The perturbation analysis described above can be used to give the first order perturbation of x_1 resolved in the directions $x_2, \ldots, x_n$. In the case when A is non-defective this lends to

$$X_i(\varepsilon) = x_i + \varepsilon \left\{ \sum_{2}^{n} \left(\frac{y_j^H B x_i}{s_j(\lambda_j - \lambda_i)} \right) x_j \right\} + O(\varepsilon^2) \qquad 5.1$$

and the coefficient of x_j is bounded by $1/|s_j(\lambda_j-\lambda_i)|$. Hence we obtain a large perturbation in the direction of x_i if s_j or $\lambda_j-\lambda_i$ is small. However this analysis is rather unsatisfactory. When A has an ill-conditioned eigenvalue problem the set of x_i will be almost linearly dependent as we show below. The fact that some of the x_j have large coefficients need not necessarily mean that the perturbation as a whole is large.

The left-hand eigenvector y_1 is orthogonal to $x_2, \ldots, x_n$ and hence x_1 may be expanded in terms of $y_1, x_2, \ldots, x_n$. In fact

$$x_1 = s_1 y_1 + \sum_{2}^{n} \alpha_i x_i \qquad (5.2)$$

since $y_1^H x_1 = s_1$ and $y_1^H x_i = 0$ $(i=2, \ldots, n)$. Equation (5.2) may be expressed in the form

$$\sum_{i=1}^{n} \beta_i x_i = s_1 y_1 / (1+\Sigma \alpha_i^2)^{\frac{1}{2}}, \qquad (5.3)$$

where

$$\beta_1 = 1/(1+\Sigma\alpha_i^2)^{\frac{1}{2}}, \quad \beta_i = -\alpha_i/(1+\Sigma\alpha_i^2)^{\frac{1}{2}}, \quad \|\beta\|_2 = 1 \quad . \qquad (5.4)$$

Hence we have a unit vector $\hat{\rho}$ so that

$$\|X\beta\|_2 = |s_1|/(1+\Sigma\alpha_i^2)^{\frac{1}{2}} < |s_1| \qquad (5.5)$$

and when s_1 is small the vectors x_i are "almost linearly dependent". (Note that in general the x_i $(i=2, \ldots, n)$ will include principal vectors which are not eigenvectors). Anticipating section 7 equation (5.5) implies that $\sigma_n(X) < |s_1|$.

Conversely if a set of the x_i are almost linearly dependent then at least one of the associated s_i is small and A has an ill-conditioned eigenvalue. Suppose for example

$$\sum_{1}^{p} \alpha_i x_i = u \text{ where } \|u\|_2 = \varepsilon, \quad \sum_{1}^{p} \alpha_i^2 = 1 \; . \qquad (5.6)$$

Then if the vectors y_i are the normalized columns of $(X^{-1})^H$ we have

$$\alpha_i y_i^H x_i = y_i^H u, \quad s_i = y_i^H u/\iota_i, \quad |s_i| \leq \varepsilon/|\alpha_i| \; , \qquad (5.7)$$

Since at least one α_i is such that $|\alpha_i| > p^{-\frac{1}{2}}$ this means that at least one s_i is small. In fact it is obvious that at least two of the s_i must be small since otherwise just one of the eigenvalues would be sensitive and the remainder insensitive; as the trace is obviously not sensitive this is impossible.

This result emphasizes one more unsatisfactory feature of ill-conditioned eigensystems. Suppose we have managed (in spite of the practical difficulties) to obtain correctly rounded versions of a set of ill-conditioned eigenvectors $x_1, \ldots, x_p$. We may now wish to determine an accurate orthogonal basis for this sub-space of dimension p. However since the vectors $x_1, \ldots, x_p$ are almost linearly dependent, when we perform the Schmidt orthogonalisation process on these x_i the orthogonal basis is bound to be poorly determined. In fact information about the last of the orthogonal vectors will be completely vitiated by the rounding errors which will usually be inherent in the representation of the x_i in the computer.

This casts doubt on the advisability of attempting to determine the x_i themselves and suggest that it might be better to determine directly an orthogonal basis for the sub-space corresponding to such vectors.

6 ORTHOGONAL BASES FOR INVARIANT SUBSPACES

The eigenvectors of A corresponding to λ are the solutions of the equation $(A-\lambda I)x = 0$. If $A-\lambda I$ is of nullity n_1 (rank n-n_1) then there will be n_1 independent

eigenvectors. These vectors span a subspace P_1, the null-space of A-XI. Let $x_1^{(1)}, x_2^{(1)}, \ldots x_{n_1}^{(1)}$ be an orthogonal basis of this subspace P_1.

Turning now to the solutions of the equation $(A-\lambda I)^2 x$ clearly they include any vector in P_1 since if $(A-\lambda I)x$ then certainly $(A-\lambda I)^2 x = 0$. The nullity of $(A-\lambda I)^2$ may therefore be denoted by n_1+n_2 where $n_2 \geq 0$. If the null-space is denoted by P_2 then $P_2 \supset P_1$ *and* the basis $x_1^{(1)}, x_2^{(1)}, \ldots, x_{n_1}^{(1)}$ may be extended to an orthogonal basis of P_2 by the addition of further orthogonal vectors $x_1^{(2)}, x_2^{(2)}, \ldots) x_{n_2}^{(2)}$. These additional vectors satisfy the relations

$$u_i = (A-\lambda I)x_i^{(2)} \neq 0 \ , \ (A-\lambda I)^2 x_i^{(2)} = 0 \qquad (i = 1, \ldots n_2) \qquad (6.1)$$

and hence they are vectors of grade 2.

We now shon that $n_2 \leq n_1$. For the vectors u_i are non-null and satisfy the relation $(A-\lambda I)u_i = 0$. Hence they lie in P_1 and if $n_2 > n_1$

$$\sum \alpha_i u_i = 0 \text{ ie } (A-\lambda I)\sum \alpha_i x_i^{(2)} = 0 \qquad (6.2)$$

which means that $\sum \alpha_i x_i^{(2)} \subset P_1$. But $\sum \alpha_i x_i^{(2)}$ is orthogonal to the $x_i^{(1)}$ by the construction and hence we have a contradiction.

Continuing in this way by considering the nullities of $(A-\lambda I)^3, (A-\lambda I)^4, \ldots$ we obtain numbers $n_3, n_4, \ldots$ s.t. $n_{i+1} \leq n_i$ and orthogonal bases of subspaces P_i such that $P_{i+1} \supset P_i$. The subspece Pi is-of dimension $m_i = n_1 + \ldots + n_i$. In general the orthogonal vectors $x_i^{(s)}$ are such that $(A-\lambda I)^{s-1} x_i^{(s)} \neq 0$ but $(A-\lambda I)^s x_i^{(s)} = 0$.

The sequence comes to an end with $(A-\lambda I)^k$ where $(A-\lambda I)^{k+1}$ is of the same nullity as $(A-\lambda I)^k$.

Comparing these spaces with those spanned by the chains of vectors associated with λ in the J.c.f. P_1 is the space spanned by the principal vectors of grade 1, P_2

that spanned by principal vectors of grades 1 and 2 etc. Notice though, that the space spanned by $x_1^{(2)}, \ldots x_{n_2}^{(2)}$ is not in general the same as that spanned by the principal vectors of grade 2 in the Jordan chains.

n_1 is equal to the number of blocks associated with λ in the J.c.f. and in general n_s is the number of those blocks which are of dimension not less than s.

The derivation of these orthogonal bases is in some ways more satisfactory than that of the Jordan chains themselves and though the chains may be derived from the orthogonal bases there will in *general* be a loss of digital information in this process.

7 THE SINGULAR VALUES

In the previous section it was shown that in the solution of the complete eigenvalue problem we ere concerned with the determination of the nullities or ranks of sequences of matrices. Rank determination is a notoriously dangerous numerical problem end in practice the only reliable way of doing it is via the singular value decomposition (S.V.D.). Accordingly we now give a brief review of the S.V.D. and the properties of singular values.

For our purposes the singular values of a complex $m \times n$ matrix A may be defined to be the non-negative square roots of the eigenvaluea of the matrix $A^H A$. Clearly $A^H A$ is an $n \times n$ non-negative definite Hermitian matrix and its eigcnvalues may be denoted by σ_i^2 (i=1, ... n); the σ_i are the singular values of A. Although apparently a more sophisticated concept than the eigenvalues, the determination of the singular values is more satisfactory from the computational point of *view*. The σ_i are defined in terms of the eigenvalues of a Hermitian matrix and these are always insensitive to small perturbations in elements of that matrix. We shall assume that the σ_i are ordered so that $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n$. The σ_i^2 may be defined via the mini-max properties $\frac{x^H A^H A x}{x^H x}$ ieof

$$\frac{\| Ax \|_2^2}{\| x \|_2^2}$$

$$\sigma_1(X) = \max \frac{\|Ax\|_2}{\|x\|_2} = \|A\|_2 , \quad \sigma_n(X) = \min \frac{\|Ax\|_2}{\|x\|_2} \qquad (7.1)$$

and

$$\sigma_r(A) - \|B\|_2 = \sigma_r(A) - \sigma_1(B) \leq \sigma_r(A+B) \leq \sigma_r(A) + \sigma_1(B) \leq \sigma_r(A) + \|B\|_2 . \qquad (7.2)$$

From the last of these relations the well-conditioned nature of the σ_r is well exposed.

Although we have defined the σ_i via A^HA they should not be determined in this way. In practice they are computed via the S.V.D. which is defined as follows.

Any m×n complex matrix A may be factorized in the form

$$A = U\Sigma V^H , \qquad (7.3)$$

where U and V are m× m and n×n unitary matrices respectively and Σ is anmxn matrix with $\Sigma_{ii} = \sigma_i$ and $\Sigma_{ij} = 0$ othernise. Golub and Reinsch [4] have described an extremely efficient and stable method for determining the S.V.D. and hence the σ_i. The computed U and V^H are almost orthogonal to the working accuracy and the computed σ_i correspond to those of some (A+E) where $\|E\|_2 / \|A\|_2$ is a modest multiple of tte machine precision. Since the σ_r are insensitive to E, this is very satisfactory.

Clearly from (7.3)

$$A^HA = V\Sigma^2V^H , \quad A^HAV = V\Sigma^2 \qquad (7.4)$$

so that the columns of V are orthogonal eigenvectors of A^HA. Similarly

$$AA^H = U\Sigma^2U^H , \quad AA^HU = U\Sigma^2 \qquad (7.5)$$

and the columns of U we orthogonal eigenvectors of AA^H.

Turning now to the case when A is n×n we have from the definitions of the eigenvalues and of the singular values

$$\Pi \lambda_i = \det(A) , \quad \Pi(\sigma_i^2) = \det(A^HA) = |\det(A)|^2 \qquad (7.6)$$

and hence

$$\Pi|\lambda_i| = \Pi\sigma_i . \qquad (7.7)$$

We have the fundamental result that $\lambda_n = 0$ iff $\sigma_n = 0$ and both imply that A is singular. The three properties are fully equivalent.

From this it is intuitively obvious that if A is "nearly" singular, λ_n and σ_n are "small" with appropriate determination of the terms "nearly" singular and "small". As a measure of the proximity of A to singularity we shall take $\|E\|_2 / \|A\|_2 = \varepsilon$ where E is the matrix of minimum norm such that A+E is singular. Since A+E is singular there exists a y such that

$$(A+E)y = 0 . \qquad (7.8)$$

Hence

$$\sigma_n = \min \frac{\|Ax\|}{\|x\|} \leq \frac{\|Ay\|}{\|y\|} = \frac{\|-Ey\|}{\|y\|} \leq \|E\|_2 = \varepsilon\|A\|_2 . \qquad (7.9)$$

On the other hand since $\min \|Ax\| / \|x\|$ is attained for some unit *vector, y (say)*

$$\sigma_n = \|Ay\| , \quad Ay = \sigma_n z \text{ with } \|z\|_2 = 1 . \qquad (7.10)$$

Hence $(A - \sigma_n zy^H)y = 0$ and $A - \sigma_n zy^H$ must be singular. But $\|\sigma_n zy^H\| = \sigma_n$ and $\varepsilon = \min \|E\|_2 / \|A\|_2 \leq \sigma_n / \|A\|_2$; hence $\sigma_n / \|A\|_2 = \varepsilon$.

Turning now to λ_n we have

$Ay = \lambda_n y$ for some y with $\|y\|_2 = 1$

and

$$\sigma_n = \min \frac{\|Ax\|}{\|x\|} \leq \frac{\|Ay\|}{\|y\|} = |\lambda_n| \quad . \tag{7.11}$$

On the other hand from (7.7)

$$|\lambda_n|^n \leq \sigma_n \sigma_1^{n-1} \tag{7.12}$$

$$|\lambda_n/\sigma_1|^n \leq \sigma_n/\sigma_1 = \sigma_n/\|A\|_2 = \epsilon \tag{7.13}$$

giving $$\|\lambda_n\| \leq \sigma_1 \epsilon^{\frac{1}{n}} \quad . \tag{7.16}$$

This last relation is disappointing but unfortunately it is a best possible result as is illustrated by the matrices K_n typified by

$$K_4 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \epsilon & 0 & 0 & 0 \end{bmatrix} \quad . \tag{7.15}$$

In general $|x_i| = \epsilon^{\frac{1}{n}}$ $(i = 1, \ldots, n)$, but $\sigma_1 = \ldots = \sigma_{n-1} = 1$ and $\sigma_n = \epsilon$. All extreme examples are of this kind since we have equality in equation (7.12) *only* if $|\lambda_i| = |\lambda_n|$ (all n) and $\sigma_1 = \sigma_2 = \ldots = \sigma_{n-1}$. In practice then we may well have a matrix which is singular to working accuracy and therefore has a negligible singular value but which hes no eigenvalues which can be regarded as in any sense small.

The practical consequences of this theoren are very serious. The most stable algorithms for computing eigenvalues can guarantee only that each computed eigenvalue λ_i' is exact for some $A+E_i$ where $\|E_i\|_2 / \|A\|_2$ is a modest multiply of the machine precision and it is difficult to conceive how such algorithms can be improved upon except, of course, by working to higher accuracy at least *in* some significant part

of tne computation. This means that $(A+E_i-\lambda_i'I)$ is exactly singular and hence that $A-\lambda_i'I$ is within $\|E_i\|_2$ of a singular matrix. Hence $A-\lambda_i'I$ has a singular value bounded by $\|E_i\|_2$ but the bound for the smallest eigenvalue of $A-\lambda_i'I$ involves $\|E_i\|_2^{\frac{1}{n}}$. All that we can guarantee a priori is that each computed λ_i' will have an error which involves the i'actor $\|E_i\|_2^{\frac{1}{n}}$ and this may be far from small.

For a normal matrix $|\lambda_r| = \sigma_r$ end hence this weakness disappears. If λ_i is an eigenvalue of A then $A+E_i$ has an eigenvalue λ_i' such that

$$|\lambda_i - \lambda_i'| \leq \|E_i\| \quad . \tag{7.16}$$

Unfortunately the realisation that this result is true has tended to lead to an overconfidence when dealing with real symmetric and Hermitian matrices, which are the commonest examples of normal matrices.

8 FACTORIZATIONS OF ALMOST-SINGULAR MATRICES

If B is an exactly singular matrix and B=XY ic a factorisation of B then oither X or Y (or both) is exactly singular. Most of the common factorizations used in practice ensure that one of the factors is certainly not singular and hence with exactly singular B and exact factorisation the other factor must be singular.

A factorisation which is frequently used is B=QR where Q is unitary and R is upper triangular. Clearly Q is non-singular and hence if B is singular, R must also be singular and therefore have a zero eigenvalue and a zero singular value. But the eigenvalue of R are its diagonal elements and hence at least *one* r_{ii} must be zero, indeed r_{nn} unless B is "special".

Now consider the case when B is almost singular and let us assume for simplicity that B is factorisod exactly. We have $\sigma_i(R) = \sigma_i(B)$ since the σ_i are invnriant with respect to unitary transformations. Hence R must still have a negligible singular value. However we can no longer guarantee that any r_{ii} is pathologically

small since the r_{ii} are merely the eigenvalues the bound for which involves $(\sigma_n(E))^{\frac{1}{n}}$.

This result is important in practice because many algorithms for solving the complete eigenproblem of a matrix first compute the eigenvalues and then attempt to determine the eigenvectors from them. If λ is an eigenvalue given by a stable algorithm $(A+E-\lambda I)$ sill be erectly singular with $\|E\|/\|A\|$ small and hence $B = A-XI$ will be almost singular. The situation appears particularly favourable when A is normal since the computed λ will then have an error which is small relative to $\|A\|_2$ ie to $|\lambda_1|$. Unfortunately although B is normal, the same is not true of R and hence we still cannot guarantee that R will have any pathologically small r_{ii}. Now the weak bound for λ_n is attained only when B is extremely pathological and hence one might expect that failure of R to have a small diagonal element would be rare. Unfortunately this is far from true. Attempts were made to construct an algorithm based on this factorisation in the case when A is a symmetric tridiagonal matrix. Por such matrices a particularly satisfactory algorithm is known for the determination of the λ's. Nsvcrtheless it was found in practice that when the QR factorisation of $A-\lambda I$ was performed for each of the n computed λ in turn almost invariably some of the R were such that they had no small r_{ii}, and all algorithms based on a search for a negligible r_{ii} failed disastrously.

The LL^T factorisation of a poeitive definite matrix A is known to be extremely stable and it might be thought that when such an A was near to singularity this would be bound to reveal itself in the corresponding L. That this is not true is illustrated by the matrices $A=L_nL_n^T$ where L_n is of the form illustrated by

$$L_4 = \begin{bmatrix} 1 & & & \\ -1 & 1 & & \\ -1 & -1 & 1 & \\ -1 & -1 & -1 & 1 \end{bmatrix}. \qquad (8.1)$$

It is easy to shorn that $\sigma_n(A_n) = \lambda_n(A_n) = O(4^{-n})$ and hence for quite modest

valves of n the matrix A_n is almost singular. Yet there is no obvious indication of this in the factor L_n since all of its diagonal elements are unity.

Finally we consider the factorisation given by Gaussian elimination with complete pivoting. This too would appear to be quite favourable and yet it can fail quite catastrophically. Indeed if A_n is of the form illustrated by

$$A_4 = \begin{vmatrix} 1 & -1 & -1 & -1 \\ & 1 & -1 & -1 \\ & & 1 & -1 \\ & & & \end{vmatrix} \qquad (8.2)$$

then it can be shown that $\sigma_n(A_n) = O(2^{-n})$ and hence A_n is almost singular for quite modest n. Yet the factorisation given by Gaussian elimination with complete pivoting is

$$A_n = I \times A_n, \qquad (8.3)$$

ie A_n is itself the upper triangular factor, and its diagonal elements are all unity.

These examples illustrate the fact that the determination of singularity, much less the *rank*, by means of simple factorisations is not a practical proposition. On the other hand the S.V.D. is extremely reliable and since the computed σ_i correspond to A+E where $\|E\|_2/\|A\|_2$ is of the *order* of the machine precision it provides an excellent means of determining the numerical rank.

9 VECTORS BY MATRIX POWERING

In the next three sections we discuss some of the algorithms which have been designed to find bases for the successive null spaces of *powers* of $(A-XI)$ corresponding to an eigenvalue λ.

For simplicity of notation we shall work throughout with $B=A-\lambda I$. We shall not for the moment discuss numerical stability but knowing that most simple factorisations are numerically unreliable for finding the rank of a matrix we shall use only the S.V.D. for this purpose. Let the S.V.D. of B be denoted by

$$B_1 \equiv B = U_1 \Sigma_1 V_1^H , \tag{9.1}$$

where U_1 and V_1 are $n \times n$ unitary matrices. Since λ is an eigenvalue, B is a singular matrix. If it is of nullity n_1 then B_1 will have n_1 zero singular value and we may write

$$BV_1 = U_1 \Sigma_1 = \left[\underbrace{A_2}_{n-n_1} \middle| \underbrace{0}_{n_1} \right] . \tag{9.2}$$

For consistency with later stages we write $W_1 \equiv V_1$ and the last n_1 columns of W_1 clearly give an orthogonal basis for the principal vectors of grade 1 while the matrix A_2 has orthogonal columns.

Proceeding to the null space of B^2 we have

$$B^2 V_1 \equiv B^2 W_1 = \left[BA_2 \middle| \underbrace{0}_{n_1} \right] = \left[B_2 \middle| \underbrace{0}_{n_1} \right] \tag{9.3}$$

the zero columns obviously persisting. We now compute the S.V.D. of B_2

$$B_2 = U_2 \Sigma_2 V_2^H , \tag{9.4}$$

where U_2 is an $n \times n$ unitary matrix and V_2 an $(n-n_1) \times (n-n_1)$ unitary matrix. Writing

$$\left[\underbrace{\quad}_{n-n_1} \middle| \underbrace{\quad}_{n_1} \right] \qquad W_2 = W_1 V_2 , \tag{9.5}$$

we have

$$B^2 W_2 = \left[\underbrace{U_2 \Sigma_2}_{n-n_1} \middle| \underbrace{0}_{n_1} \right] . \tag{9.6}$$

Since the nullity of B^2 is n_1+n_2, B_2 will have n_2 singular values and we have

$$B^2 W_2 = \left[A_3 \middle| \underbrace{0}_{n_2} \middle| \underbrace{0}_{n_1} \right] . \tag{9.7}$$

Writing $\sum_1^s n_i = m_s$ the matrix A_3 has $n-m_2$ orthogonal columns. The last m_2 columns of W_2 give an orthogonal basis for vectors of grade 2 and grade 1. The last n_1 of these columns are those of W_1 having been unaltered by this second step.

The general step is then as follows

$$B^s W_s = \left[\underbrace{A_{s+1}}_{n-m_s} \middle| \underbrace{0}_{n_s} \middle| \cdots \middle| \underbrace{0}_{n_1} \right] , \tag{9.8}$$

$$B^{s+1} W_s = \left[B_{s+1} \middle| 0 \middle| \cdots \middle| 0 \right] \text{ where } B_{s+1} = BA_{s+1} , \tag{9.9}$$

$$B_{s+1} = U_{s+1} \Sigma_{s+1} V_{s+1}^H , \tag{9.10}$$

where U_{s+1} is an $n \times n$ unitary matrix and V_{s+1} an $(n-m_s) \times (n-m_s)$ unitary matrix. B_{s+1} has n_{s+1} zero singular values and writing

$$\tilde{V}_{s+1} = \begin{bmatrix} V_{s+1} & \\ & I \end{bmatrix} , \qquad W_{s+1} = W_s \tilde{V}_{s+1} , \tag{9.11}$$

$$B_{s+1} W_{s+1} = \left[U_{s+1} \Sigma_{s+1} \middle| 0 \middle| \cdots \middle| 0 \right] , \tag{9.12}$$

$$= \left[A_{s+2} \middle| \underbrace{0}_{n_{s+1}} \middle| \cdots \middle| \underbrace{0}_{n_1} \right] . \tag{9.13}$$

The process terminates when A_{S+5} is of full rank.

The main weakness of this algorithm is the difficulty of recognising which of the elements of σ_i may be treated as zero. This is well illustrated when A and therefore B is normal. If such a matrix were inserted into this algorithm then at the first step the singular value would be $|\lambda_1|, |\lambda_2| \ldots |\lambda_n|$ of which n_1 would be treated as zero. For a normal matrix the process should terminate here since all vectors are of grade 1. However if one continue:;, the singular values in the second step would be $|\lambda_1|^2, |\lambda_2|^2, \ldots |\lambda_{n-n}|^2$ and some of these might wellbe regarded as negligible. The algorithm can be modified to limit this shortening but even then It compares unfavourably in most respects with the algorithm of the next section.

10 VECTORS BY ORTHOGONAL DEFLATION

Again it is convenient to work with B and we assume that it has an eigenvalue of multiplicity k. We write $B^{(1)} = B$ and denote the S.V.D. of $B^{(1)}$ by

$$B^{(1)} = U^{(1)}\Sigma^{(1)}(V^{(1)})^H , \tag{10.1}$$

where there will be n_1 zero singular values. Hence

$$B^{(2)} = (V^{(1)})^H B^{(1)} V^{(1)} = (V^{(1)H} U^{(1)}\Sigma^{(1)} = W^{(1)}\Sigma^{(1)} \tag{10.2}$$

and we may write

$$B^{(2)} = \left[\begin{array}{c|c} B_{11}^{(2)} & 0 \\ \hline B_{21}^{(2)} & 0 \end{array}\right] \begin{matrix} \}n-n_1 \\ \}n_1 \end{matrix} . \tag{10.3}$$

From the orthogonality of $W^{(1)}$ the first n-n, column: of $B^{(2)}$ are orthogonal and therefore independent. Relation (10.1) shows that the last n_1 columns of $V(1)$ give n_1 orthogonal eigenvectors (ie vectors of grade 1) of $B(1)$ corresponding to $\lambda=0$.

If $n_1 = k$ then we have dealt with all the eigenvalues. Otherwise $B_{11}^{(2)}$ will have $k-n_1$ zero eigenvalues and we can proceed to the consideration of vectors of grade 2. Let z be an arbitrary non-null vector partitioned conformally with $B^{(2)}$ so that $z^T = [x^T | y^T]$. Then

$$B^{(2)}z = \left[\begin{array}{c} B_{11}^{(2)} \\ \hline B_{21}^{(2)} \end{array}\right] x \tag{10.4}$$

and when $x = 0$ and $y \neq 0$, z is a vector of grade 1. If $x \neq 0$ then it follows from the independence of the first $n-n_1$ columns of $B^{(2)}$ that $B^{(2)}z \neq 0$. However we have

$$(B^{(2)})^2 z = \left[\begin{array}{c} B_{11}^{(2)} \\ \hline B_{21}^{(2)} \end{array}\right] B_{11}^{(2)} x \tag{10.5}$$

and from the same linear independence z is a vector of grade 2 iff $B_{11}^{(2)}x = 0$.

Hence we may proceed as follows. Let the S.V.D. of $B_{11}^{(2)}$ be given by

$$B_{11}^{(2)} = U^{(2)}\Sigma^{(2)}(V^{(2)})^H , \tag{10.6}$$

where $\Sigma^{(2)}$ has n_2 zero diagonal elements if $B_{11}^{(2)}$ is of nullity n_2. Hence

$$(V^{(2)})^H B_{11}^{(2)} V^{(2)} = (V^{(2)})^H U^{(2)}\Sigma^{(2)} = W^{(2)}\Sigma^{(2)} \tag{10.7}$$

and we may write

$$(V^{(2)})^H B_{11}^{(2)} V^{(2)} = \left[\begin{array}{c|c} B_{11}^{(3)} & 0 \\ \hline B_{21}^{(3)} & 0 \end{array}\right] \begin{matrix} \}n-m_2 \\ \}n_2 \end{matrix} . \tag{10.8}$$

Again the first $n-m_2$ columns of $(V^{(2)})^H B_{11}^{(2)} V^{(2)}$ are orthogonal and hence independent. Introducing the unitary matrix

$$\tilde{V}^{(2)} = \left[\begin{array}{c|c} V^{(2)} & 0 \\ \hline 0 & I \end{array}\right] \qquad (10.9)$$

$$B^{(3)} = (\tilde{V}^{(2)})^H B^{(2)} \tilde{V}^{(2)} = \left|\begin{array}{ccc} B_{11}^{(3)} & 0 & C \\ B_{21}^{(3)} & 0 & 0 \\ B_{31}^{(3)} & B_{32}^{(3)} & 0 \end{array}\right| \cdot \qquad (10.10)$$

$\underbrace{}_{n-m_2}$ $\underbrace{}_{n_2}$ $\underbrace{}_{n_1}$

it is obvious that $n_2 \leq n_1$ otherwise $B^{(3)}$ and hence $B^{(1)}$ would have been of nullity greater than n_1.

Again if $m_2 = k$ the process is complete. Otherwise $B_{11}^{(3)}$ has some zero eigenvalues and we proceed via its S.V.D., this next stage being typical. If

$$B^{(3)} = U^{(3)} \Sigma^{(3)} (V^{(3)})^H \qquad (10.11)$$

then

$$(V^{(3)})^H B_{11}^{(3)} V^{(3)} = (V^{(3)})^H U^{(3)} \Sigma^{(3)} = W^{(3)} \Sigma^{(3)} \qquad (10.12)$$

and again introducing

$$\tilde{V}^{(3)} = \left[\begin{array}{c|c} V^{(3)} & 0 \\ \hline 0 & I \end{array}\right] \qquad (10.13)$$

$\underbrace{}_{I}$

we are led to

$$B^{(4)} = (\tilde{V}^{(3)})^H B^{(3)} \tilde{V}^{(3)} = \left|\begin{array}{cccc} B_{11}^{(4)} & 0 & 0 & 0 \\ B_{21}^{(4)} & 0 & 0 & 0 \\ B_{31}^{(4)} & B_{32}^{(4)} & 0 & 0 \\ B_{41}^{(4)} & B_{42}^{(4)} & B_{43}^{(4)} & 0 \end{array}\right| , \qquad (10.14)$$

$\underbrace{}_{n-m_3}$ $\underbrace{}_{n_3}$ $\underbrace{}_{n_2}$ $\underbrace{}_{n_1}$

where n_3 is the nullity of $B_{11}^{(3)}$. By an argument similar to that used above the non-null columns of $B^{(4)}$ and of leading principal submatrices of orders $n-m_1$, $n-m_2$ are linearly independent. The process clearly terminates when $m_s = k$ at which stage $B_{11}^{(s+1)}$ is no longer singular. Since

$$B^{(s+1)} = V^H B^{(1)} V \equiv V^H B V , \qquad (10.15)$$

where $V = V^{(1)} \tilde{V}^{(2)} \tilde{V}^{(s)} \ldots \tilde{V}^{(s)}$ the principal vectors of $B^{(1)}$ may be found via those of $B^{(s+1)}$. For simplicity of notation we expose the case when $s=3$ which is wholly typical. We may write

$$B^{(4)} = \left[\begin{array}{c|c} B_{11}^{(4)} & 0 \\ \hline P & C \end{array}\right] \begin{matrix} n-m_3 \\ m_3 \end{matrix} \qquad (10.16)$$

and it is evident that

$$\left[B^{(4)}\right]^t = \left[\begin{array}{cc} (B_{11}^{(4)})^t & 0 \\ P_t & C^t \end{array}\right] . \qquad (10.17)$$

Hence

$$\left[B^{(4)}\right]^t \left[\begin{array}{c} x \\ y \end{array}\right] = \left[\begin{array}{c} (B_{11}^{(4)})^t x \\ P_t x + C^t y \end{array}\right] \qquad (10.18)$$

and since $B_{11}^{(4)}$ is non-singular $(B_{11}^{(4)})^t x$ is not null unless $x=0$. All vectors in the relevant invariant subspace have their first $n-n_3$ components equal to zero and since

$$\left[B^{(4)}\right]^t \left[\begin{array}{c} 0 \\ y \end{array}\right] = \left[\begin{array}{c} 0 \\ C^t y \end{array}\right] \qquad (10.19)$$

it is evident that we may concentrate on the matrix C given explicitly by

$$C = \begin{bmatrix} 0 & 0 & 0 \\ B_{32}^{(4)} & 0 & 3 \\ B_{42}^{(4)} & B_{43}^{(4)} & 0 \end{bmatrix} . \qquad (10.20)$$

A discussion of vectors of grade 3 will be fully illustrative. Let us take any vector x of order m_3 end partition conformally into $I = [x_1^T | x_2^T | x_3^T]$. If $x_1 \neq 0$ we have

$$Cx = \begin{bmatrix} 0 \\ B_{32}^{(4)} \\ B_{42}^{(4)} \end{bmatrix} x_1 + \begin{bmatrix} 0 \\ 0 \\ B_{43}^{(4)} \end{bmatrix} x_2 \qquad (10.21)$$

$$C^2x = \begin{bmatrix} 0 \\ 0 \\ B_{43}^{(4)} \end{bmatrix} B_{32}^{(4)} x_1 = \begin{bmatrix} 0 \\ 0 \\ B_{43}^{(4)} \end{bmatrix} z \text{ (say)} . \qquad (10.22)$$

But since we know the columns of $B_{32}^{(4)}$ are independent $z \neq 0$ and side also the columns of $B_{43}^{(4)}$ are independent $C^2x \neq 0$. on the other hand $C^3x = 0$ for eny x. The last n_1 columns of the identity matrix therefore give n_1 orthogonal vectors of grade 1, the next n_2 columns of it give vectors of grade 2 and the next n_3 columns give vectors of grade 3.

Interpreting this result in terms of B for the general case the last n_1 columns of V give orthogonal vcciors of grade1, the next n_2 give orthogonal vectors of grade 2 etc.

When the process terminate $B_{11}^{(s+1)}$ is non singular and its eigenvalues are the remaining eigenvalues cf B, ie $B_{11}^{(s+1)} + \lambda I$ gives the remaining eigenvalues of A. We can no;: turn to the next eigenvalue of A and repeat this process starting from $B^{(s+1)} + AI$. In this way a canonical form is ultimately attained which may be illustrated in the case when A has only three distinct eigervalues $\lambda_1, \lambda_2, \lambda_3$ by

$$V^H A V = \left[\begin{array}{cc|cc|ccc|c} \lambda_1 I & & & & & & & n_2^{(1)} \\ Y_{21} & \lambda_1 I & & & & & & n_1^{(1)} \\ \hline X_{31} & X_{32} & \lambda_2 I & & & & & n_3^{(2)} \\ X_{41} & X_{42} & Y_{43} & \lambda_2 I & & & & n_2^{(2)} \\ X_{51} & X_{52} & Y_{53} & Y_{54} & \lambda_2 I & & & n_1^{(2)} \\ \hline X_{61} & X_{62} & X_{63} & X_{64} & X_{65} & \lambda_3 I & & n_2^{(3)} \\ X_{71} & X_{72} & X_{73} & X_{74} & X_{75} & Y_{76} & \lambda_3 I & n_1^{(3)} \end{array}\right] . \qquad (10.23)$$

In the example given here there were tco stages with λ_3, three stages nith λ_2 and two stages with A_1 and the integers $n_j^{(i)}$ are the mullities exposed in the successive stages of the *process*. The matrix V being the product of unitary matrices is itself unitary. Note that we havedenoted the submatrices in the diagonal blocks by $Y_{i,j}$ and outside these blocks by X_{ij}. From the definition of the algorithm we have $n_j^{(i)} \geq n_{j+1}^{(i)}$ and the columns of $Y_{i+1,i}$ are linearly independent. be already know that $n_1^{(3)}$, $n_2^{(3)}$ give the number of vectors of grades 1 and 2 respectively associated with λ_5 and the corresponding columns of V provide the vectors themselves. The remaining columns of V cannot, of course, give vectors corresponding to λ_2 and λ_1 since, in general, the latter will not be orthogonal to those of λ_3. We have not yet established that $n_1^{(2)}$, $n_2^{(2)}$, $n_3^{(2)}$ give the number of vectors of Grades 1, 2, 3 associated with λ_2, and that $n_1^{(1)}$ and $n_2^{(1)}$ the vectors of grades 1, 2 associated with λ_1 and this we now do.

We cannot proceed further with the reduction without departing from unitary similarities. However if we now admit general sinilsrities the submatrices denoted by the X_{ij} may be annihilated. To annihilate X_{42} for example se premultiply by Z_{42}^{-1} and postmultiply by Z_{42} where Z_{42} is equal to the identity matrix with a bloc!: $X_{42}/(\lambda_1-\lambda_2)$ in the same position as is occupied by X_{42}. The X_{ij} ere eliminated in this way in the order X_{52}, X_{31}, X_{42}, X_{41}, X_{52}, X_{51}, X_{65}, X_{64}, ...

It is obvious that the Y_{ij} are unaffected by this. The final result is that we have

$$C = M^{-1}AM = \left|\begin{array}{c|c|c} \lambda_1 I & 0 & 0 \\ Y_{21}\ \lambda_1 I & & \\ \hline & \lambda_2 I & \\ 0 & Y_{43}\ \lambda_2 I & 0 \\ & Y_{53}\ Y_{54}\ \lambda_2 I & \\ \hline 0 & 0 & \lambda_3 I \\ & & Y_{76}\ \lambda_3 I \end{array}\right| , \qquad (10.24)$$

where M is no longer unitary (though its last $n_1^{(3)} + n_2^{(3)}$ columns are still orthogonal). From the properties of the $Y_{i+1,i}$ described above it is now evident that the $n_j^{(i)}$ have the significance described above and indeed that all the columns of M give independent (but not orthogonal) vectors of the relevant grades corresponding to $\lambda_1, \lambda_2, \lambda_3$. Notice that we have now *proved* that the canonical form (10.23) which is achieved purely by unitary transformations gives a full specification of the J.c.f. There is no need actually to proceed to the form (10.24) in order to find the J.c.f. However, from the form C and the rank properties of the $Y_{i+1,i}$ we may proceed to a demonstration of the J.c.f. itself. It is easy to show that by further similarities (10.24) may be reduced to

$$\left|\begin{array}{c|c|c} \lambda_1 I & & \\ K_{21}\ \lambda_1 I & & \\ \hline & \lambda_2 I & \\ & K_{43}\ \lambda_2 I & \\ & K_{54}\ \lambda_2 I & \\ \hline & & \lambda_3 I \\ & & K_{76}\ \lambda_3 I \end{array}\right| , \qquad (10.25)$$

where $K_{i+1,i}$ is a matrix of the form

$$\left|\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right| \qquad (10.26)$$

having the same dimension as $Y_{i+1,i}$. Apart from the ordering of the rows and columns this is the J.c.f.

It should be emphasized that we are not recommending proceeding beyond the form (10.23) and indeed if one requires an orthogonal basis associated with each of the λ_i one should return to the original matrix with each eigenvalue in turn.

The outstanding weakness of the algorithms of this section and the previous one is that the volume of work may be excessive. To find the vectors for a matrix of order n corresponding to an eigenvalue λ_1 of multiplicity k having just one block $J_r(\lambda_1)$ in the J.c.f. one must perform a S.V.D. on matrices of orders n, n-1, ... n-r in succession (the last one merely to reveal that there are no more eigenvalues equal to λ_1!).

Both algorithms were suggested by Kublanovskaya[10] but not in terms of the S.V.D. 2nd have also been described by Ruhe[14] though in different terms from those used here.

11 ECONOMICAL ALGORITHM FOR DETERMINATION OF VECTORS

An alternative algorithm suggested by Golub and Wilkinson is considerably more economical in general (though not necessarily superior in other respects). Again corresponding to an eigenvalue λ one works with $B = A-\lambda I$. We first give the basic motivation. Suppose we have already determined independent vectors u_1, u_2, u_3 of grade 1, vectors v_1, v_2 of grade 2 and vectors w_1, w_2 of grade 3 (not necessarily orthogonal).

If x is any vector of grade 5 then Bx is of grade 2 and hence lies in the subspace spanned by the u_i, v_i, w_i. In fact x must satisfy a relation

$$Bx = \left[u_1 \mid u_2 \mid u_3 \mid v_1 \mid v_2 \mid w_1 \mid w_2\right]\alpha \quad , \tag{11.1}$$

where α is a vector of order 7. However the totality of independent solutions of (11.1) includes v_1, v_2, w_1, w_2 which will have been obtained by previously solving

$$Bx = \left[u_1 \mid u_2 \mid u_3 \mid v_1 \mid v_2\right]\beta \text{ and } Bx = \left[u_1 \mid u_2 \mid u_3\right] . \tag{11.2}$$

We need a procedure which will reject these previous solutions. Indeed the solutions needed at the current stage are solutions of (11.1) which are independent of v_1, v_2, w_1, w_2. To this end we observe that instead of solving (11.1) we may equally well solve

$$Bx = \left(\left[u_1 \mid u_2 \mid u_3 \mid v_1 \mid v_2 \mid w_1 \mid w_2\right] Z\right)\alpha , \tag{11.3}$$

where Z is any non-singular 7×7 matrix, preferably unitary if one does not wish to sacrifice numerical information. Now B is a singular matrix and a convenient method for solving (11.1) is via the S.V.D. of B

$$B = U\Sigma V^H , \tag{11.4}$$

where Σ has n_1 zero elements, assumed to be in the last n_1 diagonal positions. Hence we wish to solve

$$\Sigma(V^H x) = \Sigma y = U^H \left(\left[u_1 \mid u_2 \mid u_3 \mid v_1 \mid v_2 \mid w_1 \mid w_2\right] Z\right)\alpha \tag{11.5}$$

$$= \left(\left[U^H u_1 \mid U^H u_2 \mid U^H u_3 \mid U^H v_1 \mid U^H v_2 \mid U^H w_1 \mid U^H w_2\right] Z\right)\alpha . \tag{11.6}$$

Solutions are obtained via those values of α for which the right-hand side p has zero components in the last n_1 positions. In our example $n_1 = 3$ and the equations become

$$\left|\begin{array}{l} \sigma_{11} y = p_1 \\ \sigma_2 y_2 = p_2 \\ \sigma_3 y_3 = p_3 \\ \sigma_4 y_4 = p_4 \\ 0\, y_5 = 0 \\ 0\, y_6 = 0 \\ 0\, y_7 = 0 \end{array}\right| \tag{11.7}$$

Components y_5, y_6, y_7 are therefore arbitrary and in our algorithm are taken to be zero since they merely result in including multiples of u_1, u_2, u_3 in the vector α derived from y.

We have still to discuss how we avoid duplicating the vectors we have previously produced. Suppose at the stage when we are determining the vectors v_1 and v_2 we have computed a Z (which might be called Z_1) such that

$$\left[U^H u_1 \mid U^H u_2 \mid U^H u_3\right] Z_1 = \begin{bmatrix} X & X & X \\ X & X & X \\ X & X & X \\ X & X & X \\ X & 0 & 0 \\ X & 0 & 0 \\ X & 0 & 0 \end{bmatrix} = \left[p^{(1)} p^{(2)} p^{(3)}\right] . \tag{11.8}$$

Then v_1 and v_2 are obtained by solving $\Sigma y = p$ where p takes in turn each of the vectors $p^{(2)}$ and $p^{(3)}$, giving independent solutions. Now when we have to solve the set

$$y = (\left[U^H u_1 \mid U^H u_2 \mid U^H u_3 \mid U^H v_1 \mid U^H v_2\right] Z)\alpha \qquad (11.9)$$

if we take $X = \begin{bmatrix} Z_1 & \\ & I \end{bmatrix}$ then (11.9) becomes

$$y = (\left[p^{(1)} \mid p^{(2)} \mid p^{(3)} \mid U^H v_1 \mid U^H v_2\right] \left[W_2\right])\alpha \quad . \qquad (11.10)$$

Columns two and three have already been dealt with and gave us v_1 and v_2. The new solutions are obtained by solving,

$$y = (\left[p^{(1)} \mid U^H v_1 \mid U^H v_2\right] Z_2)\alpha \; . \qquad (11.11)$$

Notice in this stage we are left with only three columns (ie n_1) just as in the previous stage. Again we determine a Z_2 which will make the trailing columns of the matrix in parenthesis on the right of (11.10) of the requisite form ie with zero in positions 5, 6, 7 (the last n_1 positions). The number of vectors of this form willdecide how many vectors of grade 3 we obtain. The algorithm is now complete and cc observe that at each stage we are dealing with a system of the form.

$$y = (RZ)\alpha \; , \qquad (11.12)$$

I here R is always an $n \times n_1$ matrix 2nd we wish to determine Z so that RZ is of the requisite form, ie its trailing columns have zeros in the last n_1 positions. This is conveniently done via an S.V.D. composition. We wri to

$$R = \begin{matrix} \left[\begin{array}{c} R_1 \\ \hline R_2 \end{array}\right] \\ n_1 \end{matrix} \begin{matrix} \} n_1' \end{matrix} \qquad (11.13)$$

where R_2 is an $n_1 \times n_1$ matrix. If $R_2 = U_2 \Sigma_2 V_2^H$ where Σ_2 has at the sth stage n_s zero diagonals. Then taking $Z = V_2$

$$RZ = \begin{bmatrix} R_1 V_2 \\ U_2 \Sigma_2 \end{bmatrix} \qquad (11.14)$$

and the lest n_s columns are of the required form having regard to the n_s zero elements in Σ_2.

The general algorithm may now be described by its typical stage at which we determine vectors of grade s+1. We assume that by this time we **have**

$$\begin{aligned} &n_1 \text{ vectors of grade 1 , } && u_1^{(1)}, u_2^{(1)} \ldots u_{n_1}^{(1)} \\ &n_2 \text{ vectors of grade 2 , } && u_1^{(2)}, u_2^{(2)} \ldots, u_{n_2}^{(2)} \\ &\cdots\cdots\cdots \\ &n_s \text{ vectors of grade s , } && u_1^{(s)}, u_2^{(s)}, \ldots u_{n_s}^{(s)} \end{aligned} \qquad (11.15)$$

We then assemble an $n \times n_1$ matrix $R^{(s+1)}$ the first $n-n_s$ columns of which will be denoted by $P^{(s)}$ the origin of which will become obvious during the description of this next stage. The remaining n_s columns are $U^H u_1^{(s)}, \ldots U^H u_{n_s}^{(s)}$. This matrix $R^{(s+1)}$ is partitioned in the form

$$R^{(s+1)} = \begin{matrix} \left[\begin{array}{c} R_1^{(s+1)} \\ \hline R_2^{(s+1)} \end{array}\right] \\ n_1 \end{matrix} \begin{matrix} \} n_1 \\ \end{matrix} \; . \qquad (11.16)$$

If the S.V.D. of $R_2^{(s+1)}$ is $U^{(s+1)} \Sigma^{(s+1)} (V^{(s+1)})^H$ where the number of zero elements in $\Sigma^{(s+1)}$ is denoted by n_{s+1}, then

$$R^{(s+1)} V^{(s+1)} = \left[\begin{array}{c} R_1^{(s+1)} V^{(s+1)} \\ \hline U^{(s+1)} \Sigma^{(s+1)} \end{array}\right] = \left[\begin{array}{c|c} \underbrace{P_{11}^{(s+1)}}_{n_1 - n_{stl}} & \underbrace{P_{12}^{(s+1)}}_{n_{stl}} \\ \hline P_{21}^{(s+1)} & 0 \end{array}\right] \qquad (11.17)$$

and the vectors $y_1^{(s+1)}, \ldots, y_{n_{s+1}}^{(s+1)}$ are obtained by solving $\sum y = p$ where p takes each of the last n_{s+1} columns of the matrix on the right in turn. The $u_i^{(s+1)}$ are then obtained by multiplying these vectors by V. The $n \times (n_1 - n_{s+1})$ matrix in the first $n-n_{s+1}$ columns of the matrix on the right of (11.17) is the matrix $P^{(s+1)}$ required in the next stage. The process terminates when $n_{s+1} = 0$.

In the first stage we solve

$$\sum y = 0 \qquad (11.18)$$

and obtain the solutions $y = e_{n-n_1+1}, e_{n-n_1+2}, \ldots, e_n$, the last n_1 columns of the identity matrix and hence $u_1^{(1)}, \ldots, u_{n_1}^{(1)}$ are the last n_1 columns of V. The vectors of grade 1 are therefore orthogonal but this is not true of any of the subsequent sets of vectors though by taking $y_{n-n_1+1}, \ldots, y_n$ to be zero in each of the subsequent solutions of $\sum y = p$ one ensures that all vectors of grades higher than one are orthogonal to those of grade 1.

Observe that the successive S.V.D.'s are all performed on a matrix of order $n_1 \times n_1$. In the case when $n_1 = 1$ and there is only one block in the J.c.f. associated with the current eigenvalue this will be a 1×1 matrix at each stage and the process come to an end when the last element of $U^H u_1^{(s)}$ is non-zero!

12 COMMENTS ON ALGORITHMS FOR PRINCIPAL VECTORS

So far we have concentrated mainly on the formal aspects of the algorithms though in using the S.V.D. we are tacitly recognising numerical difficulties. The first problem is how to select our λ when forming $B = A - \lambda I$. In practice the eigenvalues of A should have been found using some stable algorithm such as the QR algorithm. Although the computed λ_i may be arbitrarily bad each should be exact for some matrix $A+E_i$ where E_i is such that $\| E_i \|_2 / \| A \|_2$ is merely a modest multiple of the computer precision. Hence $B = A - \lambda_i I$ should have at least one negligible singular value relative to $\| A \|_2$ however "poor" λ_i may be in an absolute sense. However if A really is defective the computed λ_i are probably not the best values to use. If for example A has 3 well defined J.c.f. (ie in the optimum quasi-J.c.f. the superdiagonal elements are of the order of magnitude of $\|A\|_2$) and there is just one block $J_r(\lambda_1)$ associated with λ_1 one will expect the computed λ_i to include a set of r values $\bar{\lambda}_i, \ldots, \bar{\lambda}_r$ which, though not particularly close to λ_1, will be such that their sum is very close to $r\lambda_1$. If one could recognise such a block one should use the mean of those values $\bar{\lambda}$ and then work with $B = A - \bar{\lambda} I$. However in practice the situation will be much more obscure than this and it is a difficult problem to decide which values of λ to use.

Whichever of the algorithms we use we shall need at each stage when an S.V.D. is performed, a satisfactory criterion for deciding which singular values may be regarded as 'zero'. The situation is most satisfactory in connexion with the deflation technique. At each stage the matrix on which the S.V.D. is performed has been determined by a unitary similarity on (A-AI) and it is reasonable to use some tolerance $\varepsilon \| A \|_2$ throughout when ε is 'small' but appreciably larger than the machine precision.

In the powering algorithm the rth matrix is of degree r in the elements of A and the decision is much less satisfactory. A modification of the procedure has been developed which ameliorates this difficulty but matrix powering would seem to have nothing to recommend it in comparison with the deflation algorithm.

The Golub-Wilkinson algorithm is far superior from the point of view of economy of computation; while the first S.V.D. is done on A-XI the others are all performed on a submatrix of a set of n_1 vectors. If the vectors $v_j^{(i)}$ are normalised at each stage a negligible singular value would be one which is small compared with unity. If in the matrix Σ obtained from $B = A - \lambda I$ itself, the smallest singular value to be regarded as non-zero is quite close to the tolerance then in determining all subsequent solution of equations of the form $\sum y = p$ the element y_{n-n_i} is obtained by dividing by this almost negligible σ_{n-n_i}. The vectors obtained with this process are not orthogonal as they are with the other two and there does appear to

be a larger that they may be almost linearly dependent with a consequent loss of digital information.

None of the three processes gives principal vectors satisfying the chain reaction typical of the columns of the X producing the J.c.f. Modified vectors satisfying the chain reaction can be determined from the computed vectors but the volume of work is substantial and care is needed to avoid losing digital information. Some such loss is inevitably involved in going from the orthogonal sets given by tire powering and deflation algorithms since the vectors in the chains may be arbitrarily near to linear dependence. Indeed one might well ask whether one should move from the orthogonal sets to sets satisfying thechain relations. The answer must depend on what the vectors are needed for and here numerical analysts would welcome discussion with applied mathematicians since this is clearly a subjective matter. Further experimentation is necessary before the algorithm can be fully assessed.

13 POORLY DEFINED J.c.f.

As mentioned previously there is a natural tendency to construct "difficult" examples for testing purposes by taking a J.c.f. and subjecting it to some simple similarity transformation. Suchexamples severely underestimate the difficulties associated rith ill-conditioned matrices. The point is well illustrated by considering the Frank matrices F_n defined typically by

$$F_5 = \begin{bmatrix} 5 & 4 & 3 & 2 & 1 \\ 4 & & 3 & 2 & 1 \\ & & & 2 & 1 \\ & & 2 & 2 & 1 \\ & & & & 1 \end{bmatrix} . \qquad (13.1)$$

Even for quite modest values of n some of the eigenvalues and eigenvectors are very ill-conditioned and yet one has a simple method of determining them by observing that for example

$$\begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix} (F_5 - \lambda I) = \begin{bmatrix} 1-? & & & & \\ 4 & 1-h & & & \\ & 3 & 1-\lambda & \lambda & \\ & & 2 & 1-\lambda & \lambda \\ & & & 1 & 1-\lambda \end{bmatrix} = G . \qquad (13.2)$$

This result is quite general and enables us to determine the eigenvalues of F_5 for example fromthose of the quasi symmetric tridiagonalmatrix

$$T_5 = \begin{vmatrix} 0 & 1 & & & \\ 4 & 0 & 1 & & \\ & 3 & 0 & 1 & \\ & & 2 & 0 & 1 \\ & & & 1 & 0 \end{vmatrix} \qquad (13.3)$$

The determination of these latter eigenvalues is a cell-conditioned problem for all values of n. We are able to remove the ill-condition in this way because the transformation can be performed exactly, io sithout rounding error. The eigenvalues of F_n a-e very sensitive to perturbation in elements in the top right-hand corner and by transforming to T_n and then working explicitly with a tri-diagonal matrix one ensures that no rounding errors are effectively made in these elements! From this transformation it is easy to show that eigenvalues of Fn are such that $\lambda_r = 1/\lambda_{n-r+1}$. It is the smaller eigenvalues which we ill-conditioned.

To illustrate the nature of the ill-conditioning we concentrate for the moment on F_{12} and discuss the problem from the point of view of computation on KDF9 which has a 39 digit binary mantissa, ie rather less than 12 decimal digits of accuracy.

By row transformations we see that $\det(F_n) = 1$ and If $\tilde{F}_n$ is the matrix resulting from a perturbation c in position (c,n) we have $\det(\tilde{F}_n) = 1 \pm (n-1)!c$. Since the determinant is the productof the eigenvalues it is evident that changes of $\pm 1/(n-1)!$ in this element alter the product of the eigenvalues from the true value,

1, to 0 and 2 respectively. When n = 100, for example, this represents a change of approximately 10^{-158}. To obtain the eigenvalues correct to 10 decimals even with an extremely stable general purpose algorithm would require computation with a mantissa of about 170 decimal digits. Yet the eigenvalues may be determined via T_{100} working to ten decimals only.

For n=12 the situation is not yet too serious with 12 digit decimal computation since $11! = 4 \times 10^7$. One can expect to obtain some correct digits even in the most ill-conditioned eigenvalues. The quantities s_i for the four smallest eigenvalues are

$$s_{12} = 5 \times 10^{-8}, \; s_{11} = 3 \times 10^{-8}, \; s_{10} = 4 \times 10^{-8}, \; s_9 = 15 \times 10^{-8} \qquad (13.4)$$

the corresponding eigenvalue being

$$\lambda_{12} = 0.0310\ldots, \; \lambda_{11} = 0.0495\ldots, \; \lambda_{10} = 0.0812\ldots, \; \lambda_9 = 0.1456\ldots \qquad (13.5)$$

where we have given only the order of magnitude of the s_i. In fact the errors in the eigenvalues as computed on KDF9 using the very stable QR algorithm were 4×10^{-6}, 7×10^{-6}, 5×10^{-6} and 10^{-7} respectively and from the sensitivity considerations discussed in section 4 these results are seen to be extremely creditable.

From the discussion in that section we also know that there is certainly a matrix having a double eigenvalue λ_{11} at a distance within $\|F_{12}\| s_{11}$, but in fact F_{12} is much nearer to a defective matrix than this. Indeed it is new to quite a number of different defective matrices. Let us consider first the possibility of inducing defectiveness by a perturbation ε in the (1, 12) element only. The modified characteristic equation is

$$\prod (\lambda_i - \lambda) - \varepsilon \, 11! = 0 \qquad (13.6)$$

If we draw the graph $y = \prod (\lambda_i - \lambda)$ then the modified eigenvalues are at the values of A for which $\prod (\lambda_i - \lambda) = 11!\varepsilon$. The situation is as illustrated in Fig 1.

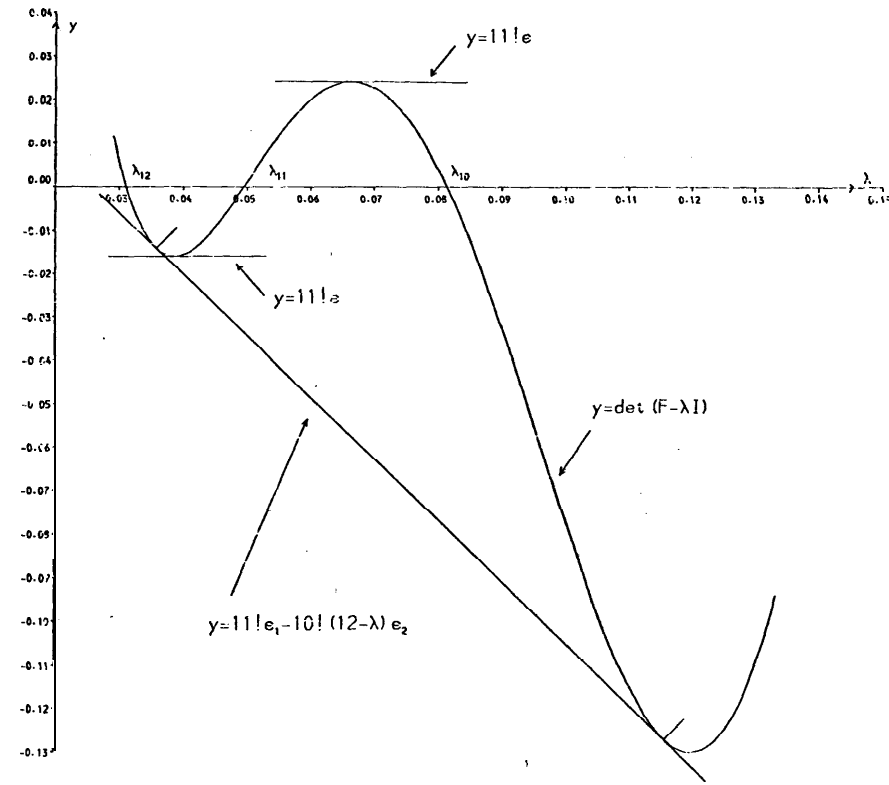

Figure 1

Taking ε to be negative we obtain a double root when the line $y = 11!\varepsilon$ is tangential to the curve, which first occurs at a point between λ_{11} and λ_{12}. If we take ε positive a double root is obtained when the line is tangential to the curve at a point between λ_{10} and λ_{11}. It is surprisingly easy to compute these points quite accurately provided $\prod(\lambda_i-\lambda)$ is computed from G_{12} not from F_{12}! The value of ε is quite a lot smaller than $\|F_{12}\| s_{11}$ and on reflexion this is not surprising. In establishing that result we attempted to induce a double eigenvalue at the value λ_i itself for which the s_i is small. It is to be expected that a smaller perturbation is needed to produce a double eigenvalue at some point "between" that λ_i and some other λ_j. As we have seen there are always perturbations εB for which $\frac{\delta\lambda_i}{\delta\varepsilon} = \pm\frac{1}{s_i}$. At least one of the other λ_j must be changing fast "to keep the trace correct" and we would expect to be able to make $\lambda_i(\varepsilon)$ and some $\lambda_j(\varepsilon)$ move towards each other. As they get nearer we would expect s_i to get even smaller and one feels intuitively that a perturbation nearer the order of magnitude $\left|\frac{1}{2}(\lambda_j-\lambda_j)s_i\right|$ is likely to give a double eigenvalue at a value of roughly $\frac{1}{2}(\lambda_i+\lambda_j)$. The quantity $\left|\frac{1}{2}(\lambda_j-\lambda_j)s_i\right|$ is likely to be much smaller than $\|A\| s_i$ since the relevant λ_j is likely to be at least "fairly close" to λ_i. This certainly proves to be true for F_{12}. In fact a value $\varepsilon = -10^{-10}(3.95\ \ldots)$ gives a double root between λ_{11} and λ_{12} at $\lambda = 0.038\ \ldots$

If perturbations ε_1 and ε_2 are made in F_{12} **(1, 12)** and F_{12}(2, 12) respectively then the characteristic equation becomes

$$\prod(\lambda_i-\lambda) - 11!\,\varepsilon_1 + 10!\,\varepsilon_2 \quad (12\text{-}A) \quad = 0 \qquad (13.7)$$

and the eigenvalues are at the intersection of the straight line $y = 11!\varepsilon_1 - 10!\varepsilon_2(12-\lambda)$ with the curve $y = \prod(\lambda_i-\lambda)$. By appropriate choices of ε_1 and ε_2 this line can be made tangential to the curve at two points, one between λ_{12} and λ_{11} and one between λ_{10} and λ_9. The values are in fact $\varepsilon_1 = -10^{-7}(6.24\ \ldots)$ and $\varepsilon_2 = -10^{-7}(3.9\ldots)$ and it gives coincident eigenvalues at 0.036 . . . and 0.116.

Notice that if one attempts to solve these perturbed matrices by the QR algorithm on KDF9 the separation of the "paired" eigenvalues may, at first sight, seem disappointing. Two points should be emphasized. First, since the KDF9 has a mantissa with less than 12 decimal digits the perturbations ε_i cannot be inserted with any great precision since they occur via entries $1+\varepsilon_1$. Hence even if the ε_i are determined accurately they cannot be included in $1+\varepsilon_1$ without incurring an error of between 10^{-11} and 10^{-12}. Further in solving the perturbed matrix A+E on KDF9 the effect of rounding errors will imply that a computed λ_i is an eigenvalue of $A+E+E_i$ where $\|E_i\|_2 / \|A\|_2$ is likely to be a modest multiple of 2^{-39} (ie $10^{-11.7}$). Since we are now extremely close to a defective matrix s_i will be quite a lot smaller than the corresponding value for A itself. In fact with $\varepsilon = -10^{-10}(3.9..)$ the two close computed values of λ were 0.03758... and **0.03963** . . . the mean of these being 0.03863.. ; this is very close to the minimum point of $\prod(\lambda-\lambda_i)$ between λ_{12} and λ_{11}. Again working with the perturbed version of G_{12} it is possible not only to insert the perturbations accurately (since they now arise as $\varepsilon_1-\varepsilon_2$ and ε_2 and not as $1+\varepsilon_1$ and $1+\varepsilon_2$) but also to compute the eigenvalues of the perturbed matrix accurately. Altogether the Frank matrices provide good material for investigating ill conditioned eigenvalues and eigenvectors. It is clear that by the time n = 20, F_n is very near to a large number of defective matrices having different sets of multiple eigenvalues and even elementary divisors of different degrees. It is natural to ask what information one should really extract and shy.

Continuing with F_{12} and KDF9 (2nd we make no excuse for being so specific, the "difficulty" involved in dealing with 2 matrix is intimately associated with the precision of computation one is prepared to use; on 2 40 decimal digit computer F_{12} could reasonably be regarded as well-conditioned!) the dilemma is particularly acute. The computed λ_9, λ_{10}, λ_{11}, λ_{12} all have <u>some</u> accuracy and it is debatable whether there is anything to be gained by pretending that they are equal or equal in pairs etc. On the other hand if one treats them as distinct and computes the corresponding eigenvectors, not only will these eigenvectors inevitably be

accurately determined but they will also be almost linearly dependent. Indeed if we use the QR algorithm they will be exact for some A+E with $\|E_2\|/\|A\|_2$ of the order of 2^{-39}. The s_i for this matrix will be quite close to those of A itself and the smallest of these is roughly 3×10^{-8}. How much information do we have at best about the space of dimension four spanned by the corresponding eigenvectors? From section 5 we see that these vectors are linearly dependent to within less than 3×10^{-8}. Certainly the fourth orthogonal direction is extremely poorly determined. Indeed all four vectors are "fairly" parallel and in performing the Schmidt orthogonalisation process there will be a loss of figures at each stage.

Would it not be better to group these four eigenvalues together and to attempt to determine directly a set of four orthogonal vectors spanning the corresponding invariant subspace? One can certainly determine the subspace in this way much more accurately. Whether it is better or not depends on what one really wants. If accuracy is an over-riding consideration the 'best' thing to do is to group all 12 eigenvalues together and then any 12 orthogonal vectors specify the subspace exactly, e_1, e_2, ..., e_{12} being an obvious choice! Here we have the ultimate absurdity of perfect accuracy in a set of vectors but no information.

A sensible compromise would seem to be the following. On a t digit computer we might aim to determine the smallest groupings of the eigenvalues for which one can claim that all the computed orthogonal bases "have t′ correct digits". Obviously one must have $t' < t$ and if one insists on t′ being too close to t one runs the risk of being forced into large groups with a consequent loss of information. There is no need to get into abstruse discussions about the meaning to be attached to the angle between a computed set of s orthogonal vectors and an exact set of s vectors defining the subspace. Since we are unlikely to require less than 3 decimal digits (say) we would merely be arguing about the relative merits of θ, $\sin\theta$, $\tan\theta$, $2\sin\frac{\theta}{2}$ etc when $\theta < 10^{-3}$ and clearly such matters are of no importance. The following is a perfectly adequate measure of the angle between the orthonormal set $u_1, u_2, \ldots, u_s$ and the orthonormal set $v_1, v_2, \ldots, v_s$. We may write

$$u_i = a_{i1}v_1 + \ldots + a_{is}v_s + r_i \qquad (i = 1, \ldots, s) \tag{13.8}$$

and the r_i might reasonably be called the residual vectors. If the two bases spanned the same subspace then $r_i = 0$. $\mathrm{Max}\,\|r_i\|$ may therefore be regarded as a measure of the errors in the u_i relative to the v_i. In fact $\|r_i\|$ is the sine of the angle between u_i and the space spanned by the v_i.

14 CALCULATION OF ORTHOGONAL BASES OF INVARIANT SUBSPACES

In classical similarity theory, unitary similarities play quite an important role, since when $X^H = X$

$$B = X^{-1}AX = X^HAX \tag{14.1}$$

and hence matrices which are unitarily similar ere also conjunctive. The fundamental result with respect to unitary similarities is that for any complex matrix A there exists a unitary matrix X such that

$$X^HAX = T\,, \tag{14.2}$$

where T is upper triangular with the eigenvalues of A on its diagonal. This is known as the Schur canonical form. The ordering of' the λ_i on the diagonal may be chosen arbitrarily.

Unitary transformations are of great significance for numerical analysts because a wide range of algorithms based on them are numerically stable. When A is real it may in general have some complex eigenvalues though these of course occur in conjugate pairs. It is convenient to remain in the real field whenever possible and there is a single modification of Schur's result which states that when A is

real there is an orthogonal X (ie a real unitary X) such that

$$X^T AX = T , \tag{14.3}$$

where T is now almost triangular except that corresponding to each complex conjugate pair of eigenvalues, T has a 2×2 block on the diagonal having as its two eigenvalues this complex pair. This is usually known as the Wintner-Murnaghan canonical form [29].

It is precisely this form which is produced by the double Francis QR algorithm, perhaps the most widely used general-purpose algorithm for finding the eigensystem of a non-normal real matrix. This algorithm works directly with a real-upper Hessenberg matrix but a general real matrix may be reduced to this form by a (real) orthogonal similarity. (For detailed discussions see [28]). The combined reduction from general form to real almost trigular T is extremely stable and it has been proved [25] that the computed matrix is such that

$$T = X^T(A+E)X , \tag{14.4}$$

where X is exactly orthogonal and $\| E\|/\| A\|$ is a modest multiple of the machine precision. Further the computed $\bar{X}$ is very close to the exactly orthogonal X for which (14.4) is true and hence, in particular, has columns which are orthogonal almost to working accuracy. Since the computed T is exactly orthogonally similar to A+E and the s_i are Invariant nith respect to orthogonal transformations, the s_1 of the matrix T give information that really is relevant. The left-hand and right-hand eigenvectors of T may be readily computed; the right-hand eigenvectors are required in any case and the additional work needed to compute the left-hand eigenvectors of T is a negligible percentage of that fo- the complete reduction. Ignoring the 2×2 blocks for the moment the left-hand and right-hand vectors for the eigenvalue in posit-ion r on the diagonal are determined by a triangular back

substitution with matrices of order n−r, and r respectively. The vectors are of the forms

$$(0, \ldots 0, y_r, y_{r+1}, \ldots y_n) \text{ and } (x_1, x_2 \ldots, x_r, 0, \ldots, 0) \tag{14.5}$$

and if these are normalised vectors the corresponding $s = x_r y_r$. The complication caused by the 2×2 blocks is not substantial and is discussed in detail in [28] pp 372, 374. The computed s_i are invaluable in any case since they give the sensitivities of the eigenvalues of T, ie of (A+E).

Now let us consider the eigenvalues in the first s positions along the diagonal of T. We may write

$$X^{-1}(A+E)X = \begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix} \begin{matrix} \}\ s \\ \}\ n-s \end{matrix} \tag{14.6}$$

(with column blocks of widths s and $n-s$)

and hence

$$(A+E)\, X_s = X_s T_{11} \tag{14.7}$$

where X_s consists of the first s columns of the orthogonal matrix X. Notice that this is true even if there arc 2×2 blocks included in T_{11} provided the first of a pair of conjugate eigenvalue eigenvalues is not in position s. These s orthogonal vectors therefore provide an orthogonal basis for the invariant subspace of A+E corresponding to this groupof s eigenvalues and, as we have remarked, even the computed columns of x are accurately orthogonal. They do, of course, provide information only about the subspaces of A+E rather than of A itself but any loss of accuracy due to this perturbation is inherent in the problem and cannot be avoided without working to a higher precision (or exactly!) at least in come significant part of the computation. Although the individual eigenvectors corresponding to those s eigenvalues may be almost linear dependent the columns of

X_r being orthogonal, cannot have this shortcoming.

There is no comparatively simple method for computing a set of orthogonal vectors giving the invariant subspace corresponding to a set of λ_i which are not in the leading position. However given any collection of λ_i it is possible to transform T into an upper triangular $\tilde{T}$ having these λ_i in the leading positions by means of an orthogonal similarity. Hence we have an orthogonal Y such that

$$Y^T(T+F)Y = \tilde{T}, \tag{14.8}$$

where F is the result of rounding errors, and since the process is stable $\|F\|/\|T\|$ and hence $\|F\|/\|A\|$ is of the order of the machine precision. Hence finally

$$Y^T X^T(A+E+G)YX = \tilde{T}, \tag{14.9}$$

where $G = XFX^T$ and $\|G\|_2 = \|F\|_2$, and the first s columns of YX give an orthogonal basis of the subspace corresponding to the selected s eigenvalues.

The transformation from T to $\tilde{T}$ was first described by Rube [14]. It is achieved by a sequence of orthogonal similarities each of which is a plane rotation and is based on the following observation. If

$$T = \begin{bmatrix} p & q \\ 0 & r \end{bmatrix} \tag{14.10}$$

then there is a plane rotation R such that $R^T T R = \begin{bmatrix} r & q \\ 0 & p \end{bmatrix}$. If this is true then clearly

$$R^T \begin{bmatrix} 0 & q \\ 0 & r-p \end{bmatrix} R = \begin{bmatrix} r-p & q \\ 0 & 0 \end{bmatrix} \tag{14.11}$$

or

$$\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 0 & q \\ 0 & r-p \end{bmatrix} = \begin{bmatrix} r-p & q \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}. \tag{14.12}$$

For this to be true $(r-p)\cos\theta - q\sin\theta = 0$ Giving

$$\cos\theta = q/a, \quad \sin\theta = (r-p)/a, \quad a = \sqrt{(r-p)^2 + q^2} \tag{14.13}$$

and a simple verification shows that with this choice of R the relation is true. Ruhe gave the analogous result in the complex case; in this q becomes $\bar{q}$ in the transformed matrix. Using this algorithm any eigenvalue may be brought into any required position along the diagonal by a sequence of plane rotations. When T is real but has 2×2 blocks corresponding to complex eigenvalues an analogous result is true in which complex pairs are always kept together in the form of a real 2×2 block. One needs only two additional algorithms which serve to interchange the position of a single real diagonal element and a real 2×2 block and to interchange the positions of two 2×2 blocks. (NB the 2×2 blocks need not remain invariant; only their eigenvalues). The relevant algorithms have been coded on KDF9 and are numerically stable.

There remains the problem of the grouping and there does not get appear to be a perfectly satisfactory method of deciding on this. It cannot be decided purely on the basis of the separation since even multiple eigenvalues corresponding to elementary divisors of moderate degree will not in general lead to "close" eigenvalues in the computed set. Further even when the exact λ_i and λ_j are by no means pathologically close, they may be so sensitive that small perturbations in A may make them so. A good working test is that a perturbation E may make them coincident if

$$\frac{\|E\|_2}{|s_i|} + \frac{\|E\|_2}{|s_j|} \geq |\lambda_i - \lambda_j| \tag{14.14}$$

though since $\|E\|/s_i$ is merely a first order perturbation, a smaller $\|E\|$ than

this may well be adequate.

However we are not merely concerned with whether the computed λ_i and λ_j could belong to a multiple root. If this is our criterion then the groups will be much smaller than is advisable. A reasonably satisfactory rule is that if our aim is to have t' decimal digits correct in the subspace on a t digit decimal computer then λ_j should be coupled with λ_i when

$$|\lambda_i-\lambda_j| \max(|s_i|, |s_j|) \leqslant 10^{(t'-t)} \|A\|_F, \qquad (14.15)$$

where $\|A\|_F$ rather than $\|A\|_2$ is used since a practical criterion is required. This criterion has been applied on KDF9 and found to be quite sound. We may illustrate this in action by means of a simple example. Consider the matrix

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ \varepsilon & 0 & 1 \end{bmatrix} \qquad (14.16)$$

The eigenvalues are $1+\varepsilon^{\frac{1}{3}}$, $1+\omega\varepsilon^{\frac{1}{3}}$, $1+\omega^2\varepsilon^{\frac{1}{3}}$ where ω is a complex cube root of unity. The separation is $\varepsilon^{\frac{1}{3}}\sqrt{3}$ and hence when ε is of the order of machine precision the eigenvalues will not appear unduly close. But the left-hand eigenvector corresponding to ε^3 is $[\varepsilon^{\frac{2}{3}}, \varepsilon^{\frac{1}{3}}, 1]$ and the right-hand eigenvector is $[1, \varepsilon^{\frac{1}{3}}, \varepsilon^{\frac{2}{3}}]$ and hence the corresponding $s_1 = 3\varepsilon^{\frac{2}{3}}/(1+\varepsilon^{\frac{1}{3}}+\varepsilon^{\frac{2}{3}})$ with similar results for the other eigenvalues. Hence $(\lambda_1-\lambda_2)s_1 \doteq 3\sqrt{3}\,\varepsilon$ and this product fully exposes the danger. These two eigenvalues would be grouped together even if one were making the tolerance very lax.

One difficulty encountered in experimentation with algorithms for finding invariant subspaces is that of obtaining a correct orthogonal basis against which to test computed subspaces except in the case of rather artificially constructed matrices. In practice we have found it useful to work with A itself and with $\tilde{A}$

such that $\tilde{a}_{ij} = a_{n+1-i,\, n+1-j}$ i.e. A is the reflection of A in its centre point. Eigenvectors etc of $\tilde{A}$ are merely those of A with components in the reverse order. If A and $\tilde{A}$ are solved by the same algorithm then one can compare orthogonal bases obtained with the two matrices. At least one computed subspace has an error which is of the order of magnitude of the angle between the two computed subspaces. Where it has been possible to determine a correct basis by independent means, the error in each of the computed subspaces has proved to be of the same order of magnitude as the angle between them. One might expect this to be true generally unless there is some special reason for errors to be correlated in some way.

For matrices with well defined J.c.f.'s the orthogonal bases determined by an algorithm based on the above have been correct almost to working accuracy. Even if one takes t' almost equal to t only the eigenvalues associated with multiple roots have been grouped together.

The results obtained with the Frank matrices are interesting. For $n=16$ the 9 smallest computed eigenvalue and their true values are given in Table 1. six of the computed values are complex and with imaginary parts which are quite comparable with the real parts. Only with λ_{10} do we begin to have any significant accuracy and λ_9 has four correct figures. The largest eigenvalues were given very accurately.

Orthogonal bases were computed for subspaces of dimensions 2, 4, 6, 7, 8, 9 obtained by grouping the corresponding number of smallest eigenvalues together. (Notice xc did not compute spaces of dimension 3, 5 since conjugate pairs were always kept together in order to be able to work in the real field). The angles between the computed bases and the true subspace are given in Table 2. The subspaces of order 2 and 4 are scarcely of any significant accuracy but that of order 6 is correct to about 3 decimals and that of order 7 to almost six decimals. Notice that this accuracy in the subspace is attained although some of the λ_i are very poor. (It

TABLE 1

Computed Eigenvalues	True Eigenvalues
λ_{16} = -.02710+i (.04506)	λ_{16} = .02176
λ_{15} = -.02710-i (.04506)	λ_{15} = .03133
λ_{14} = .06121+i (.09907)	λ_{14} = .04517
λ_{13} = .06121-i (.09907) ,	λ_{13} = .06712
λ_{12} = .1882+i (.06248)	λ_{12} = .1051
λ_{11} = .1882-i (.06248)	λ_{11} = .1775
λ_{13} = .3342	λ_{10} = .3307
λ_{9} = .6809	λ_{9} = .6809
λ_{8} = 1.469	λ_{8} = 1.469

TABLE 2

Dimension	Angle between computed & true subspace
2	3.05×10^{-2}
4	1.73×10^{-2}
6	6.23×10^{-4}
7	1.74×10^{-6}
8	1.73×10^{-8}
9	2.67×10^{-10}

should be emphasized though that every computed λ_i is an eigenvalue of some $A+E$ with $\|E\|/\|A\|$ of the order of 2^{-39}).

We may look at these results from an alternative point of view. If the matrix F_{16} is regarded as having relative errors of order 10^{-12} in its elements then the invariant subspace corresponding to its two smallest elements is scarcely determined at all, while that corresponding to its smallest 7 eigenvalues for example is determined to about six decimals.

15 INVERSE ITERATION AND ILL-CONDITIONED EIGENSYSTEMS

Inverse iteration is one of the main tools used in practice for the calculation of eigenvectors from computed eigenvalues. The motivation for inverse iteration, due to Wielandt[23], springs from the observation thst if A is a matrix with a complete set of eigenvectors x_i then an arbitrary vector may be expressed in the form

$$y = \sum_{1}^{n} a_i x_i \tag{15.1}$$

and hence

$$(A-kI)^{-1}y = \sum_{1}^{n} a_i x_i/(\lambda_i-k) \ . \tag{15.2}$$

If $|\lambda_j-k| << |\lambda_i-k|$ $(i \neq j)$ the components of x_j will be very much larger than the coefficients of the remaining x_i unless the vector y happens to be very deficient in x_j. If in particular k is a very accurate approximation to λ_j the right-hand side of (15.2) may be written in the form

$$\left(\frac{1}{\lambda_j-k}\right)\left[a_j x_j + \sum_{i\neq j} a_i(\lambda_j-k)x_i/(a_i-k)\right] \tag{15.3}$$

and the normalized form of this vector will be x_j to very high accuracy. However for non-normal matrices a computed λ may not be particularly near to any eigenvalue and it appears that one can no longer expect such a spectacular performance in one iteration.

Varah was the first to point out this is not so. The simplest way to see this is to forget about the expansion of y and concentrate directly on the solution of $(A-\lambda I)z=y$ where $\|y\|_2 = 1$. We may write

$$(A-\lambda I)z/\|z\|_2 = y/\|z\|_2 , \quad w = z/\|z\|_2 , \quad y/\|z\|_2 = r \qquad (15.4)$$

giving

$$(A-\lambda I)w = r \ , \ \|w\|_2 = 1 \ , \ \|r\| = 1/\|z\|_2 = \varepsilon \text{ (say)} . \qquad (15.5)$$

The first of equations (15.5) may be expressed in the form

$$(A-rw^H)w = \lambda w \qquad (15.6)$$

and hence λ and w are an exact eigenvalue of eigenvector of the matrix $A-rw^H$. Since $\|rw^H\|_2 = \|r\|_2 = \varepsilon$ it is evident that if $\|z\|_2$ is 'large' A and w are satisfactory since they arc exact for a neighbouring matrix.

Now if we start with a value of λ which is an exact eigenvalue of A+E then however poor λ may otherwise be

$$(A+E-\lambda I)q = G \text{ for some } \|q\|_2 = 1 . \qquad (15.7)$$

Hence $(A-\lambda I)q = -Eq$ and if one takes $y = -Eq/\|Eq\|_2$ the solution of $(A-\lambda I)z = y$ is $z = q/\|Eq\|_2$ and $\|z\| \geq 1/\|E\|_2$. With this choice of y then we obtain a very large z in one iteration and the corresponding $w = z/\|z\|_2$ is a satisfactory eigenvector corresponding to λ. Obviously if we take as initial y an arbitrary unit vector the probability of it being very deficient in the vector $-Eq/\|Eq\|_2$ is very small and hence inverse iteration will "work" in one iteration with almost any starting vector.

However Varah also produced an argument which suggested that when λ is related to an ill-conditioned eigenvalue there are severe disadvantages in performing more

than one step of inverse iteration ond a satisfactory analysis of the phenomenon was subsequently given by Wilkinson [27]. It is instructive to analyse this phenomenon in terms of the S.V.D. docomposition. We showed in section 5 that if λ_i is an ill-conditioned eigenvalue the associated s_i is small and the matrix X of eigenvectors has a small singular value $\sigma_n < |s_i|$. If the S.V.D. of X is

$$X = U\Sigma V^H \ , \ XV = U\Sigma \qquad (15.8)$$

then

$$u_n = Xv_n/\sigma_n = \frac{1}{\sigma_n}\left[a_1x_1 + a_2x_2 + \ldots + a_nx_n\right], \text{ where } a_i = v_{in} \qquad (15.9)$$

and hence the unit vector u_n expanded in terms of the x_i has very large coefficients. If we take an abitrary vector y it *can* be expressed in the form

$$y = \beta_1u_1 + \ldots + \beta_nu_n \qquad (15.10)$$

where the β_i are distributed in a natural way. When transformed to its expansion in term of the x_i we have

$$y = \left[\frac{a_1\beta_n}{\sigma_n} + \ldots\right]x_1 + \left[\frac{a_2\beta_n}{\sigma_n} + \ldots\right]x_2 + \ldots + \left[\frac{a_n\beta_n}{\sigma_n} + \ldots\right]x_n \qquad (15.11)$$

and in general all the coefficients of the x_i will be very large but will be in ratios which are independent of the β_i provided β_n is not small. From (15.11)

$$z = (A-\lambda I)^{-1}y = \left[\frac{a_1\beta_n}{\sigma_n} + \ldots\right]\frac{x_1}{\lambda_1-\lambda} + \left[\frac{a_2\beta_n}{\sigma_n} + \ldots\right]\frac{x_2}{\lambda_2-\lambda} + \ldots + \left[\frac{a_n\beta_n}{\sigma_n} + \ldots\right]\frac{x_n}{(\lambda_n-\lambda)} \qquad (15.12)$$

and z will in general be a large vector for two reasons. First because σ_n is small and secondly because usually one of the $(\lambda_i-\lambda)$ will be moderately small (though not usually pathologically so). Now when z is normalised prior to doing the second iteration the coefficients of the x_i in this normalised z will no

longer be special in the way that they were in the first "arbitrary" y.

In fact the normalised vector will be essentially

$$\frac{\lambda_i-\lambda}{\lambda_1-\lambda}\left[\alpha_1+\ldots\right]x_1+\frac{\lambda_i-\lambda}{\lambda_2-\lambda}\left[\alpha_2+\ldots\right]x_2+\ldots+\left[\alpha_i+\ldots\right]x_i+\ldots+\frac{\lambda_i-\lambda}{\lambda_n-\lambda}\left[\alpha_n+\ldots\right]x_n \qquad (15.13)$$

and the coefficients of the x_i will be of order unity. In the first vector these coefficients were all large but cancelled out to give a vector of normal size. Consequently in the second step of inverse iteration the growth in size will come only from the comparative smallness of a $\lambda_i-\lambda$ and will not be reinforced by the smallness of σ_n. This will be true of all subsequent steps unless at the rth step all the quantities $(\frac{\lambda_i-\lambda}{\lambda_j-\lambda})^r$ are almost equal, when the normalised value will have large components of each of the x_i in the same ratios as in the first vector. In this case every rth iteration will give a large growth and consequently a satisfactory vector. This situation will usually occur when A has an elementary divisor of degree r. Varah has effectively used this behaviour of the iterates to given information on the structure of the J.c.f. of A [21].

The analysis may be carried out in an alternative way which is also instructive. We observe first that if λ_i is an exact eigenvalue of A then $A-\lambda_i I$ is singular and if

$$(A-\lambda_i I) = U\Sigma V^H \qquad (15.14)$$

then $\sigma_n = 0$. Consequently

$$(A-\lambda_i I)v_n = 0\,, \quad u_n^H(A-\lambda_i I) = 0 \qquad (15.15)$$

and v_n and u_n are normalised right-hand and left-hand eigenvectors of A, with $u_n^H v_n = s_i$.

Now suppose λ_i is an exact eigenvalue of $A+E$; *then* $\sigma_n(A-\lambda_i I) \leq \|E\|_2$. If we now write

$$A-\lambda_i I = U\Sigma V^H \text{ and } (A-\lambda_i I)v_s = \sigma_s u_s \qquad (15.16)$$

then $\sigma_n \leq \|E\|$. An arbitrary unit vector y may now be expanded in the form

$$y = \sum \alpha_j u_j \text{ with } \|\alpha\|_2 = 1 \qquad (15.17)$$

and

$$z = (A-\lambda I)^{-1}y = \sum \frac{\alpha_j v_j}{\sigma_j}\,. \qquad (15.18)$$

The coefficient of v_n is α_n/σ_n and

$$\left|\alpha_n/\sigma_n\right| \geq \alpha_n/\|E\|\,. \qquad (15.19)$$

Unless y is accidentally deficient in u_n the full growth takes place in the first iteration. The normalised z is essentially of the form

$$v_n + \sum_n^n \gamma_i v_i\,, \qquad (15.20)$$

where the γ_i are small. To see the effect of the second iteration one requires an expansion in terms of the u_i rather than the v_i and we now show that in this expansion the coefficient of u_n is small. Indeed since $u_n^H v_n$ is roughly s_i from the previous argument, and all the γ_i are small. this is immediately obvious. The normalised z is therefore an unfortunate vehicle for inverse iteration since it is deficient in u_n.

16 IMPROVEMENT OF AN INVARIANT SUBSPACE

Suppose A has been reduced to upper triangular form T by a unitary similarity X with a group of associated λ_j in the s leading diagonal positions of T. We then have for the computed X and T

$$AX - XT = E . \tag{16.1}$$

The error analysis guarantees that E will be almost negligible to working accuracy. Each element of the matrix E may be determined in practice by accumulating the whole of the inner-product involved in doable-precision before rounding. If $F = X^{-1}E$ then

$$X^{-1}AX = T + X^{-1}E = T + F \tag{16.2}$$

and since the computed X is almost exactly orthogonal one can compute F via $X^{T}E$. From an invariant subspace of T+F one can improve the corresponding subspace of A itself. We partition T+F in the form

$$\begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix} + \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix}, \tag{16.3}$$

where T_{11} contains the grouped eigenvalues. The relevant invariant subspace of T is spanned by the first s columns of I and hence if we write

$$\left(\begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix} + \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix}\right)\begin{bmatrix} I \\ \hline Y \end{bmatrix} = \begin{bmatrix} I \\ \hline Y \end{bmatrix}\left[T_{11} + G_{11}\right] \tag{16.4}$$

$\begin{bmatrix} I \\ Y \end{bmatrix}$ gives the improved subspace. From (16.4) neglecting second order quantities

$$T_{11} + F_{11} + T_{12}Y - T_{11} + G_{11}$$

$$T_{22}Y + F_{21} = YT_{11} \tag{16.5}$$

and Y is the solution of

$$\left[T_{22}Y - YT_{111} = -F_{21}\right] . \tag{16.6}$$

The matrix Y may be determined column by column via the relations

$$T_{22}y_1 - t_{11}y_1 = -f_1 , \qquad (T_{22}-t_{11}I)y_1 = -f_1 \tag{16.7}$$

$$T_{22}y_2 - t_{12}y_1 - t_{22}y_2 = f_2 , \quad (T_{22}-t_{22}I)y_2 = -f_2 + t_{12}y_1 . \tag{16.8}$$

In general the rth column of Y is the solution of a triangular system of equations with matrix $(T_{22}-t_{rr}I)$.

From Y one can determine G_{11} via (16.5).

If one includes the second order terms then (16.6) becomes

$$\left[T_{22}Y - YT_{11}\right] = -F_{21} + \left[-F_{22}Y + Y\left(T_{12}Y + F_{11} + F_{12}Y\right)\right] \tag{16.9}$$

and after solving (16.6) an improved right-hand side is that in (16.9) in which the computed Y is used. In this way Y may be repeatedly improved by iteration.

However there is little point in this. The matrix F is not known exactly. There are errors made in computing E in the first place and further errors in computing $X^{-1}E$, and here no purpose is served in computing Y accurately. In (16.3) we have purposely refrained from writing

$$\begin{bmatrix} \bar{T}_{11} & \bar{T}_{12} \\ F_{21} & \bar{T}_{22} \end{bmatrix} \quad \text{where } \bar{T}_{11} = T_{11} + F_{11},\ \bar{T}_{12} = T_{12} + F_{12},\ \bar{T}_{22} = T_{22} + F_{22} \qquad (16.10)$$

although this would have simplified the expressions. This is because it is necessary to keep the F matrix separate from the T matrix on the computer. The information in F would promptly be lost if the addition were carried out. However there are obviously slight advantages in replacing (16.6) by

$$(\bar{T}_{22} Y - Y_{11}\bar{T}_{11}) = -F_{21} . \qquad (16.11)$$

The improved subspace X_1 is now $\tilde{X}_1 = X_1 + X_2Z$. It is no longer quite orthogonal but this is of no importance. If one wishes to continue with the refinement of the subspace one should return to the computation of the residual via the relation

$$A\left[\tilde{X}_1 \mid X_2\right] - \left[\tilde{X}_1 \mid X_2\right]\left[\tilde{T}\right] = \tilde{E} , \qquad (16.12)$$

where

$$\tilde{T} = \begin{bmatrix} T_{11} + G_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix} \qquad (16.13)$$

The new $\tilde{E}$ will not be smaller than E in general but the next correction to the subspace will be. If $\tilde{X} = \left[\tilde{X}_1 \mid X_2\right]$ then

$$\tilde{X}^{-1} A \tilde{X} - \tilde{T} = (\tilde{X})^{-1} \tilde{E} = \tilde{F} \qquad (16.14)$$

and one can still use the approximation $(\tilde{X})^{-1} = X^T$. When computing the new correction Y the equations corresponding to (16.6)' will be

$$\left[T_{22}\tilde{Y} - \tilde{Y}\tilde{T}_{11}\right] = -\tilde{F}_{21} \qquad (16.15)$$

and $\tilde{T}_{11} = T_{11} + S_{11}$ is no longer upper triangular. However we may use T_{11} in

place of $\tilde{T}_{11}$ since $Y S_{11}$ will be of second order.

The process of iterative refinement is wholly analogous to that used with linear equations [see eg [25] Chapter 4]. In general we can continue until we have a basis which is correct to working accuracy. Indeed at the time when the process terminates the new X_1 will be obtained in terms of the old X_1 and the original X_2 by the relation

$$X_1 \text{ (new)} = X_1 \text{ (old)} + X_2Z \qquad (16.16)$$

where X_2Z is small. The true sum on the right-hand side will be accurate to more than the working precision.

The final X_1 will not have orthogonal columns but they will be almost orthogonal. If true orthogonality is required no appreciable loss of accuracy will occur when the Schmidt orthogonalisation process is performed.

ACKNOWLEDGEMENTS

The authors wish to thank Mrs G Peters of the National Physical Laboratory and Dr Peter Businger of the Bell Telephone Laboratories for the computational work on invariant subspsces and Dr B Heap and Mr Clive Hall also of the National Physical Laboratory for the production of Figure 1 on the computer KDF9. The work of G H Golub was supported in part by NSF, GJ35135X and AEC, AT(04-3)-326PA # 30.

[1] C DAVIS and W M KAHAN, The rotation of eigenvectors by a perturbation III. SIAM J Numer. Anal. 7, 1-46 (1970).

[2] D K FADDEEV and V N FADDEEVA, Computational Methods of Linear Algebra. W H Freeman, San Francisco and London (1963).

[3] G H GOLUB and W M KAHAN, Calculating the singular values and pseudo-inverse of a matrix. J SIAM Numer. Anal. Ser. B 2, 205-224 (1965).

[4] G H GOLUB and C REINSCH, Singular value decomposition and least squares solutions. Numer. Math. 14, 403-420 (1970).

[5] R T GREGORY, Defective and derogatory matrices. SIAM Review, 2, 134-139 (1960).

[6] A S HOUSEHOLDER, The Theory of Matrices in Numerical Analysis. Ginn (Blaisdell) Boston, Mass (1964).

[7] W M KAHAN, Conserving confluence curbs ill-condition. Computer Science Technical Report 6 of the University of California (1972).

[8] T KATO, PerturbationTheory for Linear Operators. Springer-Verlag, Perlin (1966).

[9] G KREISSELMEIER, On the dotermination of the Jordan form of a matrix. IEEE Transactions on automatic control AC-18, No 6, 686-687 (1973).

[10] V N KUBLANOVSKAYA, On a method of solving the complete eigenvalue problem for a degenerate matrix. Z. Vycisl. Mat. mat. Fiz. 6, 611-620 (1966). Translation in USSR Comp. Math & math. Phys. 6, 4, 1-14 (1968).

[11] P LANCASTER, Theory of Matrices. Academic Press, New York (1969).

[12] B ii PARLETT and C REINSCH, Balancing a matrix for calculation of eigenvalues and eigenvectors. Numer. Math. 13, 293-304 (1969).

[13] G PETERS and J H WILKINSON, Studies in Numerical Analysis. Edited by B K P Scaife. Academic Press New York and London (1974).

[14] A RUHE, An algorithm for numerical determination of the structure of a general matrix. BIT, 10, 196-216 (1970).

[15] A RUHE, Perturbation bounds for means of eigenvalues and invariant subspaces. BIT, 10 343-354 (1970).

[16] A RUHE, Properties of a matrix with a very ill-conditioned eigenproblem. Numer. Math. 15, 57-60 (1970).

[17] G W STEWART, On the sensitivity of the eigenvalue problem. SIAM J. Numer Anal. 9, 669-686 (1974).

[18] G W STEWART, Error bounds for invariant subspaces of closed operators. SIAM J. Numer. Anal. 8, 796-808 (1972).

[19] G W STEWART, Introduction to Matrix Computations. Academic Press, New York and London (1973).

[20] J M VARAH, Rigorous machine bounds for the eigensystem of a general complex matrix. Math. Comp. 22, 793-801 (1968).

[21] J M VARAH, Computing invariant subspaces of a general matrix when the eigensystem is poorly conditioned. Math. Comp. 24, 137-149 (1970).

[22] J M VARAH, Invariant subspace perturbations for a non-normal matrix. Presented at IFIP Conference (1971).

[23] H WIELANDT, Bestimmung höherer eigenwerte durch gebrochene iteration. Ber, B44/J/37 der Aerodynamischen Versuchanstalt Göttingen (1944).

[24] J H WILKINSON, Rounding Errors in Algebraic Processes. Her Majesty's Stationery Office, Prentice-Hall, New Jersey (1963).

[25] J H WILKINSON, The Algebraic Eigenvalue Problem. Clarendon Press, Oxford (1969).

[26] J H WILKINSON, Note on matrices with a very ill-conditioned eigenproblem. Numer. Math. 19, 176-178 (1972).

[27] J H WILKINSON, Inverse iteration in theory and in practice. Istituto Nazionale di Alta Matematica Symposia Mathematica, 10 (1972).

[28] J H WILKINSON and C REINSCH, Handbook far Automatic Computation, Vol II, Linear Algebra. Springer.-Verlag, Berlin and New York (1971).

[29] A WINTNER and F D MURNAGHAN, A canonical form for real matrices under orthogonal transformations. Proc. USA Nat. Academy 17, 417-420 (1931).