

## SOLUTION OF LINEAR EQUATIONS BY DIAGONALIZATION OF COEFFICIENTS MATRIX\*

BY

E. G. KOGBETLIANTZ

*International Business Machines Corporation*

**Introduction.** To solve a system of linear equations  $Gx = g$ , where  $G$  is a complex matrix, usually an equivalent system  $Hx = h$  is formed in which  $H = *G'G$  is Hermitian and  $h = *G'g$ . The solution of  $Hx = h$  is performed again by replacing its complex equations by real ones

$$Au - Bv = m, \quad Bu + Av = n,$$

where  $H = A + iB$ ,  $x = u + iv$ ,  $h = m + in$ . Thus, finally the order of  $G$  is doubled which is undesirable even when electronic computers are used.

It is preferable to solve  $Gx = g$  without doubling the order of  $G$  and this can be done inverting the complex matrix  $G$  by diagonalization. If such two unitary matrices  $U$ ,  $T$  can be formed, that

$$*U'GT = K,$$

where  $*U'$  denotes the conjugate transpose of  $U$  and  $K$  is a diagonal matrix, then  $G^{-1} = TK^{-1}*U'$  and the direct solution of  $Gx = g$  is obtained:

$$x = G^{-1}g = TK^{-1}*U'g.$$

This form is very convenient since the multiplication of matrices is performed by an electronic computer in almost no time.

The unitary matrices  $U$  and  $T$  are known to exist (for example modal matrices of  $G*G'$  and  $*G'G$ ), but the question how to form them for a given complex matrix  $G$  of large order is a difficult one. In what follows they are formed as infinite convergent products of simple unitary matrices  $u_n(z_n)$  which represent plane rotations through complex angles  $z_n = \theta_n + i\phi_n$ . Therefore, their practical computation is easy for electronic equipment and it yields a new method of solution of linear systems.

When this method is applied to a Hermitian system  $Hx = h$  or to a symmetric system of real equation (particular case of  $Hx = h$ ) the diagonal matrix  $K$  into which  $H$  is transformed has as its elements the characteristic roots (Eigenvalues) of  $H$ , while  $U = T$  yields the characteristic vectors, because  $K = *U'HU$ . The same holds for skew symmetric Hermitian matrices  $[(a_{rr} + ib_{rs})]$  with  $b_{rs} = b_{sr}$ ,  $a_{rs} + a_{sr} = 0$ .

To achieve the diagonalization of a given complex matrix  $A = A_0$  a sequence  $[A_n]$  of its transforms is constructed, the transform  $A_i$  being formed from  $A_{i-1}$  by right and left multiplications:

$$A_i = *u'_i(\zeta_i)A_{i-1}u_i(z_i). \quad (1)$$

Thus

$$A_n = *U'_n A T_n \quad \text{and} \quad K = \lim_{n \rightarrow \infty} A_n,$$

with

$$T = \lim_{n \rightarrow \infty} T_n = \lim_{n \rightarrow \infty} \prod_1^n u_k(z_k); \quad U = \lim_{n \rightarrow \infty} U_n = \lim_{n \rightarrow \infty} \prod_1^n u_k(\zeta_k).$$

---

\*Received March 29, 1954.



practically impossible to reduce the moduli of off-diagonal elements below a prescribed quantity.

It is a favorable circumstance that the speed of convergence of our method does not depend on the number of steps already performed. On the average it remains constant however large the number of cycles becomes: as will be shown later, the moduli of off-diagonal elements decrease in the ratio of one to  $e^{-\frac{1}{2}}$  per cycle.

Thus after, at most, 46 cycles the average modulus of all off-diagonal elements of  $A_N$ , where  $N = 23n(n - 1)$ , will be less than  $Me^{-23} \approx M10^{-10}$ ,  $M$  denoting the largest modulus of off-diagonal elements of  $A$ .

**2. Computation of  $C$ .** We formulate now the rules for the numerical computation of elements  $c_r$ , of  $C = {}^*u_{km}(z')Au_{km}(z)$ , omitting the derivation of these rules. For any  $z$  and  $\zeta$  we have the following invariants:

$$c_{kk}c_{mm} - c_{km}c_{mk} = a_{kk}a_{mm} - a_{km}a_{mk},$$

$$|c_{jk}|^2 + |c_{im}|^2 = |a_{ik}|^2 + |a_{jm}|^2, \quad |c_{kj}|^2 + |c_{mi}|^2 = |a_{ki}|^2 + |a_{mi}|^2, \quad (2)$$

$$|c_{kk}|^2 + |c_{km}|^2 + |c_{mk}|^2 + |c_{mm}|^2 = |a_{kk}|^2 + |a_{km}|^2 + |a_{mk}|^2 + |a_{mm}|^2. \quad (3)$$

Equations (2) and (3) prove the invariance of the norm:  $\|C\| = \|A\|$ . Equation (3) for  $c_{km} = c_{mk} = 0$  becomes

$$|a_{kk}|^2 + |a_{km}|^2 + |a_{mk}|^2 + |a_{mm}|^2 = |c_{kk}|^2 + |c_{mm}|^2$$

which shows that the amount  $|a_{km}|^2 + |a_{mk}|^2$  of off-diagonal terms which vanish in  $C$  is transferred into the principal diagonal. Denoting the real and imaginary components of a complex number  $z$  by  $z'$  and  $iz''$  we first form 33 real numbers in six steps.

- 1)  $M = (a'_{ik} + a'_{mm})/2, \quad r = (a''_{ik} + a''_{mm})/2,$   
 $P = (a'_{ik} - a'_{mm})/2, \quad p = (a''_{ik} - a''_{mm})/2,$   
 $N = (a'_{km} + a'_{mk})/2, \quad n = (a''_{km} + a''_{mk})/2,$   
 $Q = (a'_{km} - a'_{mk})/2, \quad q = (a''_{km} - a''_{mk})/2.$
  - 2)  $A = PM + pr - (NQ + nq), \quad B^* = MN + rn - (PQ + pq),$   
 $A^* = PM + pr + (NQ + nq), \quad C = rQ - qM + (pN - nP),$   
 $B = MN + rn + (PQ + pq), \quad C^* = rQ - qM - (pN - nP).$
  - 3)  $w = (A^2 + B^2)^{1/2}, \quad w^* = (A^{*2} + B^{*2})^{1/2}, \quad R = (w^2 + C^2)^{1/2} = (w^{*2} + C^{*2})^{1/2};$   
 $\cos 2\theta = \delta = A/w, \quad \cos 2\xi = \delta^* = A^*/w^*,$
  - 4)  $\operatorname{sech} 2\phi = \gamma = w/R, \quad \operatorname{sech} 2\eta = \gamma^* = w^*/R.$
  - 5)  $\sin \theta = s_1 = \sigma(B)[(1 - \delta)/2]^{1/2}, \quad \sin \xi = s_1^* = \sigma(B^*)[(1 - \delta^*)/2]^{1/2},$   
 $\cos \theta = c_1 = [(1 + \delta)/2]^{1/2} > 0, \quad \cos \xi = c_1^* = [(1 + \delta^*)/2]^{1/2} > 0,$
- (4)
- $$\gamma^{1/2}\sinh\phi = s_2 = \sigma(C)[(1 - \gamma)/2]^{1/2}, \quad \gamma^{*1/2}\sinh\eta = s_2^* = \sigma(C^*)[(1 - \gamma^*)/2]^{1/2},$$
- $$\gamma^{1/2}\cosh\phi = c_2 = [(1 + \gamma)/2]^{1/2} > 0, \quad \gamma^{*1/2}\cosh\eta = c_2^* = [(1 + \gamma^*)/2]^{1/2} > 0.$$

Here the symbol  $\sigma(v)$  denotes the sign of a real number  $v$ , so that  $\sigma(v) = v/|v|$ , if  $v \neq 0$ , while  $\sigma(0) = 0$  by definition. All eight numbers computed in (4) are less in absolute value than one. The two cosines  $c_1, c_1^*$  are positive, acute angles being chosen as  $\theta$  and  $\xi$ .

$$\begin{aligned}
 6) \quad G &= (M + P)c_1c_1^* + (M - P)s_1s_1^* + (N + Q)s_1c_1^* + (N - Q)c_1s_1^*, \\
 g &= (r + p)c_1c_1^* + (r - p)s_1s_1^* + (n + q)s_1c_1^* + (n - q)c_1s_1^*, \\
 H &= (M - P)c_1c_1^* + (M + P)s_1s_1^* - (N - Q)s_1c_1^* - (N + Q)c_1s_1^*, \\
 h &= (r - p)c_1c_1^* + (r + p)s_1s_1^* - (n - q)s_1c_1^* - (n + q)c_1s_1^*.
 \end{aligned}
 \tag{5}$$

Now  $c_{kk}$  and  $c_{mm}$  are:

$$\begin{aligned}
 (\gamma + \gamma^*)c_{kk} &= 2[(G + ig)c_2c_2^* - (H + ih)s_2s_2^*], \\
 (\gamma + \gamma^*)c_{mm} &= 2[(H + ih)c_2c_2^* - (G + ig)s_2s_2^*].
 \end{aligned}
 \tag{6}$$

Among the off-diagonal elements of  $C$  only those in the  $k$ th and  $m$ th rows and columns are to be computed since  $c_{ij} = a_{ij}$  for  $i, j \neq k, m$ . For each value of  $j \neq k, m, 1 \leq j \leq n$ , the elements  $c_{jk}, c_{jm}, c_{ki}, c_{mi}$  are computed at a time as a group. First eight real numbers  $\alpha, \lambda; \alpha^*, \lambda^*; \beta, \mu; \beta^*, \mu^*$  are formed:

$$\begin{aligned}
 \alpha + i\lambda &= c_1a_{jk} + s_1a_{jm}, & \alpha^* + i\lambda^* &= c_1^*a_{kj} + s_1^*a_{mj}, \\
 \beta + i\mu &= c_1a_{jm} - s_1a_{jk}, & \beta^* + i\mu^* &= c_1^*a_{mi} - s_1^*a_{ki},
 \end{aligned}$$

and then

$$\begin{aligned}
 c_{jk} &= (\alpha + i\lambda)c_2 + (\beta + i\mu)is_2, & c_{ki} &= (\alpha^* + i\lambda^*)c_2^* - (\beta^* + i\mu^*)is_2^*, \\
 c_{jm} &= (\beta + i\mu)c_2 + (\alpha + i\lambda)is_2, & c_{mi} &= (\beta^* + i\mu^*)c_2^* - (\alpha^* + i\lambda^*)is_2^*.
 \end{aligned}
 \tag{7}$$

To form the matrices  ${}^*U'_n$  and  $T_n$  the elements of  $u_{km}(z)$ , and  ${}^*u_{km}(\zeta')$  should also be computed. They are given by

$$\begin{aligned}
 \gamma^{1/2} \sin z &= s_1c_2 + ic_1s_2, & \gamma^{*1/2} \sin \zeta &= s_1^*c_2^* + ic_1^*s_2^*, \\
 \gamma^{1/2} \cos z &= c_1c_2 - is_1s_2, & \gamma^{*1/2} \cos \zeta &= c_1^*c_2^* - is_1^*s_2^*.
 \end{aligned}$$

**3. Exceptional cases.** The rules of Sec. 2 hold in the general cases when  $ww^* \neq 0$ . If  $ww^* = 0$ , some of them are modified as indicated below. We discuss five exceptional cases:

I)  $w^* = 0, w \neq 0$ ; II)  $w = 0, w^* \neq 0$ ; III)  $w = w^* = 0, C^2 = C^{*2} \neq 0$ ; IV)  $w = w^* = C = C^* = 0$ , but  $n^2 + Q^2 > 0$ ; V)  $w = w^* = C = C^* = n = Q = 0$ .

Case I.  $|C^*| = R > w > 0$ , so that  $\sigma(C^*)$  has a meaning. We take  $s_1, c_1, s_2, c_2$  as in (4), but  $s_2^* = \sigma(C^*)/2^{1/2}, c_2^* = 1/2^{1/2}$  since  $\gamma^* = 0$ . Now  $B^* = 0$  and  $\sigma(B^*)$  is meaningless, but in this case  $\xi = 0$ , so that  $s_1^* = 0$  and  $c_1^* = 1$ . The elements  $c_{jk}, c_{jm}$  are given by (7), but

$$\begin{aligned}
 c'_{ki} &= 2^{-1/2}[a'_{ki} + a''_{mi}\sigma(C^*)], & c''_{ki} &= 2^{-1/2}[a''_{ki} - a'_{mi}\sigma(C^*)], \\
 c'_{mi} &= 2^{-1/2}[a'_{mi} + a''_{ki}\sigma(C^*)], & c''_{mi} &= 2^{-1/2}[a''_{mi} - a'_{ki}\sigma(C^*)],
 \end{aligned}
 \tag{8}$$

as well as

$$\begin{aligned} \gamma c'_{kk} &= \mathbf{1}2^{1/2}[Gc_2 - Hs_2\sigma(C^*)], & \gamma c''_{kk} &= 2^{1/2}[gc_2 - hs_2\sigma(C^*)], \\ \gamma c'_{mm} &= 2^{1/2}[Hc_2 - Gs_2\sigma(C^*)], & \gamma c''_{mm} &= 2^{1/2}[hc_2 - gs_2\sigma(C^*)]. \end{aligned}$$

Case II. This is similar to I:  $\gamma = 0$  and  $|C| = R > w^* > 0$ .

We take  $s_1^*, c_1^*, s_2^*, c_2^*$  as in (4), but  $s_1 = 0, c_1 = 1, s_2 = \sigma(C)/2^{\frac{1}{2}}$  and  $c_2 = 1/2^{\frac{1}{2}}$ . The expressions of elements  $c_{ik}, c_{im}, c_{kk}, c_{mm}$  are:

$$\begin{aligned} c'_{ik} &= 2^{-1/2}[a'_{ik} - a''_{im}\sigma(C)], & c''_{ik} &= 2^{-1/2}[a''_{ik} + a'_{im}\sigma(C)], \\ c'_{im} &= 2^{-1/2}[a'_{im} - a''_{ik}\sigma(C)], & c''_{im} &= 2^{-1/2}[a''_{im} + a'_{ik}\sigma(C)], \\ \gamma^* c'_{kk} &= 2^{1/2}[Gc_2^* - Hs_2^*\sigma(C)], & \gamma^* c_{kk} &= 2^{1/2}[gc_2 - hs_2^*\sigma(C)], \\ \gamma^* c'_{mm} &= 2^{1/2}[Hc_2^* - Gs_2^*\sigma(C)], & \gamma^* c_{mm} &= 2^{1/2}[hc_2 - gs_2^*\sigma(C)]. \end{aligned} \tag{9}$$

Case III. Here  $\gamma = \gamma^* = 0$  and  $c_1 = c_1^* = 1, s_1 = s_1^* = 0, c_2 = c_2^* = 1/2^{\frac{1}{2}}$ , while  $s_2 = \sigma(C)/2^{\frac{1}{2}}, s_2^* = \sigma(C^*)/2^{\frac{1}{2}}$ . The expressions of diagonal elements  $c_{kk}, c_{mm}$  are:

$$\begin{aligned} c_{kk} &= M + P + (M - P)\sigma(CC^*) - (n + q)\sigma(C) + (n - q)\sigma(C^*), \\ c'_{kk} &= r + p + (r - p)\sigma(CC^*) + (N + Q)\sigma(C) - (N - Q)\sigma(C^*), \\ c'_{mm} &= M - P + (M + P)\sigma(CC^*) - (n - q)\sigma(C) + (n + q)\sigma(C^*), \\ c''_{mm} &= r - p + (r + p)\sigma(CC^*) + (N - Q)\sigma(C) - (N + Q)\sigma(C^*). \end{aligned}$$

The off-diagonal elements are computed as in (8) and (9).

Case IV. The conditions  $c_{km} = 0, c_{mk} = 0$  yield one and the same equation, so that we add the condition  $\zeta = z$ . We subdivide this case  $n^2 + Q^2 > 0$  into: IVa)  $n \neq 0$  and IVb)  $n = 0$ , but  $Q \neq 0$ .

In the first case IVa):

$$\delta^* = \delta = p(n^2 + p^2)^{-1/2}; \quad \gamma^* = \gamma = (n^2 + p^2)^{1/2}/R,$$

where  $R = (n^2 + p^2 + Q^2)^{\frac{1}{2}}$ . Formulae (4) hold for  $c_1, c_2, |s_1|$  and  $|s_2|$ , the signs of  $s_1$  and  $s_2$  being now  $\sigma(n)$  and  $\sigma(Q)$  respectively. If  $Q = 0$ , then  $s_2 = 0$ . The off-diagonal terms in  $C$  are computed as in (7), while

$$c'_{kk} = M + NR/n; \quad c''_{kk} = r + R; \quad c'_{mm} = M - NR/n; \quad c''_{mm} = r - R. \tag{10}$$

IVb). Now  $n = 0, Q \neq 0$ . Interpreting  $\sigma(0)$  as zero, we have  $\delta = \sigma(p)$  for any  $p \geq 0$ , so that

$$s_1^* = s_1 = \{[1 - \sigma(p)]/2\}^{1/2}, \quad c_1^* = c_1 = \{[1 + \sigma(p)]/2\}^{1/2},$$

while  $c_2^* = c_2 = [(1 + |p|/R)/2]^{\frac{1}{2}}, s_2^* = s_2 = \sigma(Q)[(1 - |p|/R)/2]^{\frac{1}{2}}$ . The elements of  $C$  are computed as in IVa), replacing in (10)  $N/n$  by  $-q/Q$ , that is:

$$c'_{kk} = M - qR/Q; \quad c''_{kk} = r + R; \quad c'_{mm} = M + qR/Q; \quad c''_{mm} = r - R.$$

Case V. In this case the vanishing of  $n$  and  $Q$  entails also  $p = M = 0$  since  $w = w^* = C = C^* = 0$ . Thus,  $c_{km} = 0$  and  $c_{mk} = 0$  reduce to only one equation, namely

$$P \sin 2z - N \cos 2z + iq = 0, \tag{11}$$

if we add the condition  $\zeta = z$ . This condition  $\zeta = z$  presupposes  $N^2 + P^2 > 0$ , since if  $N = P = 0$ , then  $q$  cannot vanish because of  $|a_{mk}|^2 + |a_{km}|^2 = N^2 + q^2 > 0$ .

Va). If  $w^2 = N^2 + P^2 > 0$ , then, denoting  $w^2 + q^2$  by  $R^2$ , we have:  $\delta = P/w, \gamma = w/R, \sigma(\sin \theta) = \sigma(N), \sigma(\sinh \varphi) = -\sigma(q)$ . Since  $\theta = \xi, \varphi = \eta$  we have  $c_1^* = c_1$  etc., and  $c_1 = [(1 + \delta)/2]^\frac{1}{2}; s_1 = \sigma(N)[1 - \delta]/2]^\frac{1}{2}; c_2 = [(1 + \gamma)/2]^\frac{1}{2}; s_2 = -\sigma(q)[1 - \gamma]/2]^\frac{1}{2}$ . Now  $G = -H = w, g = h = r$  and thus  $c_{kk} = R + ir$ , while  $c_{mm} = -R + ir$ .

Vb). If  $N = P = 0$ , we take  $\zeta + z' = 0$  that is  $\varphi = \eta, \theta + \xi = 0$ . The condition  $c_{km} = c_{mk} = 0$  becomes  $\tan 2\theta = q/r$  so that  $\varphi = \eta$  is arbitrary. We take  $\varphi = \eta = 0$ . Since  $\delta = r(r^2 + q^2)^{-\frac{1}{2}}$  we have  $c_1^* = c_1 = [(1 + \delta)/2]^\frac{1}{2}, s_1^* = s_1 = \sigma(q)[1 - \delta]/2]^\frac{1}{2}$ . The elements of  $C$  are:  $c_{kk} = c_{mm} = i(r^2 + q^2)^\frac{1}{2}; c_{jk} = \alpha + i\lambda; c_{jm} = \beta + i\mu; c_{ki} = \alpha^* + i\lambda^*$  and  $c_{mi} = \beta^* + i\mu^*$ .

**4. Convergence.** Let us consider two consecutive transformations of  $A$  into  $C$  with  $c_{kj} = c_{jk} = 0$  and then of  $C$  into  $D$  with  $d_{km} = d_{mk} = 0$ . Thus, the elements  $a_{km}, a_{mk}, a_{ki}, a_{mj}, a_{jk}, a_{jm}$  of  $A$  are transformed according to the scheme:

$$\begin{aligned} a_{ki} &\rightarrow 0 \rightarrow d_{ki}; & a_{km} &\rightarrow c_{km} \rightarrow 0; & a_{jm} &\rightarrow c_{jm} \rightarrow d_{jm}; \\ a_{jk} &\rightarrow 0 \rightarrow d_{jk}; & a_{mk} &\rightarrow c_{mk} \rightarrow 0; & a_{mi} &\rightarrow c_{mi} \rightarrow d_{mi}. \end{aligned}$$

The relations (2) yield  $|d_{jk}|^2 + |d_{jm}|^2 = |c_{jm}|^2; |d_{ki}|^2 + |d_{mi}|^2 = |c_{mi}|^2$ . Now

$$\begin{aligned} |c_{jm}|^2 &= \gamma^* |a_{jm} \cos \zeta' + a_{km} \sin \zeta'|^2, \\ |c_{mi}|^2 &= \gamma |a_{mi} \cos z + a_{mk} \sin z|^2. \end{aligned}$$

Using polar forms so that  $a_{rs} = |a_{rs}| e^{i\omega_{rs}}$  etc., we have:

$$\begin{aligned} 2 |c_{mi}|^2 &= |a_{mi}|^2 + |a_{mk}|^2 + [|a_{mi}|^2 - |a_{mk}|^2] \cos 2\theta \\ &+ 2 |a_{mi} a_{mk}| [\gamma \sin 2\theta \cos (\omega_{mi} - \omega_{mk}) + \tanh 2\varphi \sin (\omega_{mi} - \omega_{mk})] \end{aligned} \tag{12}$$

and a similar expression for  $2 |c_{jm}|^2$  in which  $a_{mi}, a_{mk}, \theta$  and  $\varphi$  are replaced by  $a_{jm}, a_{km}, \xi$  and  $-\eta$  respectively. Each pair of symmetric elements  $a_{rs}, a_{sr}, s \neq r$ , vanishes and reappears once in a cycle. When the number of cycles  $p$  increases the various values of  $\theta, \varphi, \xi, \eta, \omega_{rs}, \omega_{sr}$  related to a fixed pair  $a_{rs}, a_{sr}$  can be considered as random values. Therefore, when  $p$  increases the average values of sine and cosine of  $2\theta, 2\xi, \omega_{rs} - \omega_{ri}, \omega_{rs} - \omega_{jr}$ , and of  $\tanh 2\varphi$  tend to zero. It means that when  $p$  is large the two last terms in (12) and in the corresponding expression  $2 |c_{jm}|^2$  do not contribute to final values of moduli and these values are to be derived from the approximate expressions

$$|c_{jm}|^2 \approx [|a_{jm}|^2 + |a_{km}|^2]/2; \quad |c_{mi}|^2 \approx [|a_{mi}|^2 + |a_{mk}|^2]/2$$

which represent the average effect of one cycle, when a large number of cycles is considered. This gives, for one cycle,

$$|d_{jk}|^2 + |d_{jm}|^2 + |d_{ki}|^2 + |d_{mi}|^2 \approx [|a_{jm}|^2 + |a_{km}|^2 + |a_{mi}|^2 + |a_{mk}|^2]/2.$$

Consider now the global effect of a cycle on the average modulus  $\nu$  of all  $n(n - 1)$  off-diagonal elements, where

$$\nu^2 = \left( \sum_{i \neq j} |a_{ij}|^2 \right) / n(n - 1).$$

Since for each pair of symmetric elements the sum of four moduli squared of elements of  $D$  is approximately equal to half the sum of four moduli squared of elements of  $A$ ,

the average modulus  $\nu$  is multiplied  $(1/2)n(n - 1)$  times per cycle by the factor  $[1 - 2/(n - 1)n]^{\frac{1}{2}}$ .

At each step, indeed, two among  $n(n - 1)$  squares of moduli of off-diagonal terms are lost, the sum of four new ones being equal to the sum of only two old ones on the average. After the completion of a cycle with its  $n(n - 1)/2$  steps, the average modulus  $\nu$  becomes

$$\nu[1 - 2/(n - 1)n]^{n(n-1)/4} \sim \nu e^{-1/2}.$$

Thus, after  $p$  cycles the average modulus of off-diagonal terms decreases  $e^{p/2}$  times. The speed of convergence as characterized by a constant factor of decrease, namely  $e^{-\frac{1}{2}}$  per cycle, does not depend on the number of cycles. The results of numerical computations performed with the aid of IBMs new electronic data processing machine type 701 agree completely with the conclusions of this study of convergence. The type 701 performs with extreme rapidity: a  $32 \times 32$  matrix is diagonalized in 19 minutes and the numerical results are printed in four minutes, the total time being 23 minutes. Since the errors do not accumulate the procedure is very accurate: reconstructing the original matrix from the diagonal matrix and the Eigenvectors, if the matrix is normal, or from the diagonal matrix and unitary matrices  $U$  and  $T$  in the general case, very small differences are obtained between the elements of the original and the corresponding elements of the reconstructed matrix. Thus, for instance, in the case of a real symmetric matrix of order  $32 \times 32$  these differences were less than  $10^{-9}$ . Since the elements of the original matrix were of the order of  $10^{-2}$  the relative differences were less than  $10^{-7}$ .

**5. Special matrices.** Our diagonalization method does not preserve the skew symmetry. A skew symmetric matrix is diagonalized as a general complex matrix, its skew symmetry being lost from the first step. Since for a skew symmetric matrix  $a_{km} + a_{mk} = 0$ ,  $a_{kk} = 0$  its transformation is characterized by

$$c_{kj} + c_{jk} = ua_{kj} + va_{mj}; \quad c_{mj} + c_{jm} = -v'a_{kj} + u'a_{mj}, \quad \text{where}$$

$u = \gamma^{*\frac{1}{2}} \cos \zeta' - \gamma^{\frac{1}{2}} \cos z$  and  $v = \gamma^{*\frac{1}{2}} \sin \zeta' - \gamma^{\frac{1}{2}} \sin z$ . If we want the skew symmetry preserved, we must have  $ua_{kj} + va_{mj} = 0$  and  $v'a_{kj} - u'a_{mj} = 0$  for any  $a_{kj}$ ,  $a_{mj}$  so that  $u = v = 0$  and this gives  $\gamma^* = \gamma$ ,  $\zeta' = z$ . But then  $\zeta' = z$  entails together with  $a_{kk} = a_{mm} = 0$  the corollary  $-c_{mk} = c_{km} = a_{km}$  and we cannot make vanish in  $C$  the elements  $c_{mk}$ ,  $c_{km}$ . But the symmetry of a complex matrix as well as the Hermitian and skew Hermitian character of matrices are preserved. We consider in this paragraph the symmetric complex, general real, skew Hermitian and Hermitian matrices. In each case the general procedure is somewhat simplified thanks to special properties of the class considered.

*Symmetric complex matrices.* If  $a_{km} = a_{mk}$ , then  $Q = q = 0$ ,  $C = C^* = 0$ ,  $w = w^*$ . Therefore  $\zeta = z'$  and  $c_1^* = c_1$ ,  $s_1^* = s_1$ ,  $c_2^* = c_2$ ,  $s_2^* = -s_2$  which entails also  $\alpha^* = \alpha$ ,  $\beta^* = \beta$ ,  $\lambda^* = \lambda$  and  $\mu^* = \mu$ . Formulae (7) hold and they prove that  $c_{jk} = c_{kj}$ . Formulae (6) are simplified:

$$\begin{aligned} c_{kk} &= c'_{kk} + ic''_{kk}, & c_{mm} &= c'_{mm} + ic''_{mm}, & \text{with} \\ c'_{kk} &= \delta P + (RM + DN)/w, & c''_{mm} &= -\delta P + (RM - DN)/w, \\ c''_{kk} &= \delta p + (Rr + Dn)/w, & c'_{mm} &= -\delta p + (Rr - Dn)/w. \end{aligned}$$

For real general matrices the numbers  $n$ ,  $p$ ,  $q$ ,  $r$  vanish, so that  $C = C^* = \varphi = \eta = 0$ ,  $\gamma = \gamma^* = 1$ ,  $z = \theta$ ,  $\zeta = \xi$ ,  $c_2 = c_2^* = 1$ ,  $s_2 = s_2^* = 0$ ,  $w = w^* = R = R_1 R_2$ , where

$R_1^2 = M^2 + Q^2$  and  $R_2^2 = N^2 + P^2$ . We have also  $\delta = \cos 2\theta = (MP - NQ)/R_1R_2$ ;  $\delta^* = \cos 2\xi = (MP + NQ)/R_1R_2$  as well as

$$\begin{aligned} s_1 &= \sigma(s_1)[(1 - \delta)/2]^{1/2}; & c_1 &= \sigma(c_1)[(1 + \delta)/2]^{1/2}; \\ s_1^* &= \sigma(s_1^*)[(1 - \delta^*)/2]^{1/2}; & c_1^* &= \sigma(c_1^*)[(1 + \delta^*)/2]^{1/2}. \end{aligned} \tag{13}$$

The signs in (13) must agree with the relations  $M \sin(\theta - \xi) = Q \cos(\theta - \xi) = MQ/R_1$ ;  $P \sin(\theta + \xi) = N \cos(\theta + \xi) = NP/R_2$  and thus, for  $|MN| \neq |PQ|$ , the signs are:

		$\sigma(s_1)$	$\sigma(s_1^*)$	$\sigma(c_1)$	$\sigma(c_1^*)$	
For	$ MN  >  PQ  \geq 0:$	$\sigma(MN)$	$\sigma(N)$	+1	$\sigma(M)$	(14)
For	$ PQ  >  MN  \geq 0:$	$\sigma(PQ)$	$-\sigma(Q)$	+1	$\sigma(P)$	

Suppose  $|MN| = |PQ| > 0$ , the singular case  $MN = PQ = 0$  being discussed later. If  $|MN| = |PQ| > 0$  then both rules (14) hold since this is a limiting case when  $|MN| \rightarrow |PQ|$  for  $|MN| \leq |PQ|$ . No contradiction can arise between the two rules (14):  $\sigma(s_1^*)$  becomes meaningless, if  $MN = PQ$ , because then  $\sin \xi = s_1^*$  vanishes, while for  $MN = -PQ$ ,  $s_1 = \sin \theta$  vanishes so that  $\sigma(s_1)$  is meaningless.

In the singular case  $MN = PQ = 0$  the parameters  $N$  and  $Q$  cannot vanish simultaneously because  $2(N^2 + Q^2) = a_{km}^2 + a_{mk}^2 > 0$ , otherwise no transformation is applied to  $a_{km} = a_{mk} = 0$  and this pair is skipped. At least two of four parameters  $M, N, P, Q$  vanish since  $MN = PQ = 0$ , but it may also happen that three of them vanish in which case either  $N \neq 0$ , or  $Q \neq 0$ . In all we distinguish five subcases:

- 1)  $a_{kk} = a_{mm}, a_{km} + a_{mk} = 0$ , so that  $N = P = 0$ , but  $MQ \neq 0$ ;
- 2)  $a_{kk} + a_{mm} = 0, a_{km} = a_{mk}$ , so that  $M = Q = 0$ , but  $NP \neq 0$ ;
- 3)  $a_{kk} = a_{mm} = 0, a_{km} \pm a_{mk} \neq 0$ , so that  $M = P = 0$ , but  $NQ \neq 0$ ;
- 4)  $a_{kk} = a_{mm} = 0, a_{km} + a_{mk} = 0$ , so that  $M = P = N = 0, Q \neq 0$ ;
- 5)  $a_{kk} = a_{mm} = 0, a_{km} - a_{mk} = 0$ , so that  $M = P = Q = 0, N \neq 0$ .

In all these five cases  $s_1^* = 0$  and  $c_1^* = 1$ , while the values of  $s_1 = \sin \theta$  and  $c_1 = \cos \theta$  are found to be:

Case:	1)	2)	3)	4)	5)
$s_1 =$	$Q/R_1$	$N/R_2$	$\sigma(Q)$	$\sigma(Q)$	$\sigma(N)$
$c_1 =$	$M/R_1$	$P/R_2$	0	0	0

The elements of  $C$  are computed using the following expressions:

$$\begin{aligned} c_{jk} &= c_1 a_{jk} + s_1 a_{jm}, & c_{jm} &= -s_1 a_{jk} + c_1 a_{jm}, \\ c_{ki} &= c_1^* a_{ki} + s_1^* a_{mi}, & c_{mi} &= -s_1^* a_{ki} + c_1^* a_{mi}, \\ c_{kk} &= (M^2 + Q^2)^{1/2} + (N^2 + P^2)^{1/2}, & c_{mm} &= (M^2 + Q^2)^{1/2} - (N^2 + P^2)^{1/2}. \end{aligned}$$



If a real matrix is symmetric, then the procedure is simplified further and for a *real symmetric matrix* we have  $\delta = P(N^2 + P^2)^{-\frac{1}{2}}$  so that

$$c_1^* = c_1 = [(1 + \delta)/2]^{1/2}; \quad s_1^* = s_1 = \sigma(N)[(1 - \delta)/2]^{1/2}.$$

The elements of  $C$  are as follows:

$$c_{kk} = M + (N^2 + P^2)^{1/2}, \quad c_{mm} = M - (N^2 + P^2)^{1/2},$$

$$c_{ik} = c_1 a_{ik} + s_1 a_{im}, \quad c_{im} = -s_1 a_{ik} + c_1 a_{im}.$$

The diagonalization yields the characteristic roots and vectors of the original matrix since  $u'(\xi) = u'(\theta) = [u(\theta)]^{-1}$ .

*Skew Hermitian matrices.* The skew symmetry of a Hermitian matrix is preserved. In this case  $M = P = N = q = 0$  and  $a_{kk} = i(r + p)$ ,  $a_{mm} = i(r - p)$ ,  $a_{km} = Q + in$  and  $a_{mk} = -Q + in$ , with  $Q^2 + n^2 > 0$ . Here we find  $H = G = 0$  so that  $c_{kk}$  and  $c_{mm}$  are purely imaginary. We take  $\zeta = z$  since the vanishing of  $c_{km}$  entails also  $c_{mk} = 0$ . It is easy to verify that  $\alpha^* + \alpha = 0$ ,  $\beta^* + \beta = 0$ ,  $\lambda^* = \lambda$ ,  $\mu^* = \mu$  so that  $c_{ki} + c_{ik} = 0$  and  $c_{mi} + c_{im} = 0$ . This completes the proof that the skew symmetry is preserved. Since  $\zeta = z$  we have  ${}^*u(\zeta') = u^{-1}(z)$  so that our procedure yields the characteristic roots and vectors of skew symmetric Hermitian matrices. It proceeds as usual in the case of a complex matrix, a half only of off-diagonal elements being computed because of the skew symmetry, while the diagonal elements are:  $c_{kk} = i(r + R)$  and  $c_{mm} = i(r - R)$  with  $R = (n^2 + p^2 + Q^2)^{\frac{1}{2}}$ .

*Hermitian matrices.* Now  $a_{mk} = a'_{km}$  and  $a_{kk}$ ,  $a_{mm}$  are real. Therefore,  $m = n = p = Q = 0$  and  $\zeta = z$  since the vanishing of  $c_{km}$  entails that of  $c_{mk} = c'_{km}$ , the character of a Hermitian matrix being preserved. We have, indeed, in this case  $\alpha^* = \alpha$ ,  $\beta^* = \beta$ ,  $\lambda^* + \lambda = 0$ ,  $\mu^* + \mu = 0$  so that  $c_{ki} = c'_{ik}$  and  $c_{mi} = c'_{im}$ . On the other hand,  $g = h = 0$  and  $c_{kk}$ ,  $c_{mm}$  are real. Denoting  $(N^2 + P^2)^{\frac{1}{2}}$  and  $(N^2 + P^2 + q^2)^{\frac{1}{2}}$  by  $w$  and  $R$  respectively, it is found that  $\delta = P/w$  and  $\gamma = w/R$  which gives  $c_1^* = c_1$ ,  $s_1^* = s_1$ ,  $c_2^* = c_2$  and  $s_2^* = s_2$  in their usual form with  $\sigma(s_1) = \sigma(N)$  and  $\sigma(s_2) = -\sigma(q)$ , while  $c_1 > 0$ ,  $c_2 > 0$ . These results hold in the general case, when  $N \neq 0$ . When  $q = 0$ , we have  $s_2 = 0$  so that  $\sigma(q)$  is not needed. The elements of  $C$  are computed by (7), while the diagonal elements are simply  $c_{kk} = M + R$ ,  $c_{mm} = M - R$ .

In the case when  $N = 0$ , we have  $q \neq 0$  because  $|a_{km}|^2 = |a_{mk}|^2 = N^2 + q^2$  cannot vanish. The numbers  $c_2^* = c_2$  and  $s_2^* = s_2$  are given by the formulas  $s_2 = -\sigma(q)[(1 - |P|/R)2]^{\frac{1}{2}}$ ,  $c_2 = [(1 + |P|/R)/2]^{\frac{1}{2}}$  for  $P \geq 0$ , while  $s_1^* = s_1 = [1 - \sigma(P)]/2$  and  $c_1^* = c_1 = [1 + \sigma(P)]/2$  with  $\sigma(0) = 0$ , as usual. The diagonal elements are again  $c_{kk} = M + R$ ,  $c_{mm} = M - R$  where  $R = (p^2 + q^2)^{\frac{1}{2}}$  since  $N = 0$ . The off-diagonal elements are computed as in (7) in which this time, if  $P \neq 0$

$$2(\alpha + i\lambda) = a_{ik} + a_{im} + (a_{ik} - a_{im})\sigma(P),$$

$$(P \neq 0)$$

$$2(\beta + i\mu) = -a_{ik} + a_{im} + (a_{ik} + a_{im})\sigma(P).$$

For  $P = 0$  we have  $(\alpha + i\lambda)2^{\frac{1}{2}} = a_{ik} + a_{im}$  and  $(\beta + i\mu)2^{\frac{1}{2}} = a_{im} - a_{ik}$ .

6. It is interesting to observe that the diagonalization of the most general complex matrix  $G$  yields the upper and lower bounds for the absolute values  $|\lambda_j|$  of its characteristic roots.

A known theorem (H. Weyl, Proc. Natl. Acad. Sci. U.S. 1949, pp. 408-411) states that the common characteristic roots  $\alpha_i^2$  of two matrices  $G^{**}G$  and  $GG^{**}$ ,  $\alpha_1^2 \geq \alpha_2^2 \geq \alpha_3^2 \geq \dots \geq \alpha_n^2 \geq 0$  are related to those  $\lambda_1, \lambda_2, \dots, \lambda_n$  of  $G$ ,  $|\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|$ , by an equality  $\prod_1^n |\lambda_i| = \prod_1^n \alpha_i$  and by  $n - 1$  inequalities  $\prod_1^m |\lambda_i| \leq \prod_1^m \alpha_i$  ( $m = n - 1, n - 2, \dots, 3, 2, 1$ ), where  $\alpha_i$  denotes the positive square root  $(\alpha_i^2)^{\frac{1}{2}}$ .

Therefore,  $\alpha_1$  and  $\alpha_n$  are the bounds for the moduli of  $\lambda_i$  :

$$\alpha_1 \geq |\lambda_1| \geq \dots \geq |\lambda_n| \geq \alpha_n .$$

On the other hand it is easy to prove that the moduli of elements  $k_i$  (these elements are complex numbers in general) of the diagonal matrix  $K$  into which our diagonalization transforms  $G$  are precisely equal to  $\alpha_i$  :  $|k_i| = \alpha_i$  .

Thus, although the diagonalization in general does not yield the characteristic roots  $\lambda_i$  of  $G$ , it gives the upper and lower bounds for their moduli:  $|k_1| \leq |\lambda_1| \leq |\lambda_n| \leq |k_n|$  .

All the proofs were omitted for the sake of brevity and also because of their elementary character.