UNCLASSIFIED

OK

OAK RIDGE NATIONAL LABORATORY

Operated By

UNION CARBIDE NUCLEAR COMPANY

UCC

POST OFFICE BOX P
OAK RIDGE, TENNESSEE ORNL
CENTRAL FILES NUMBER

CF- 55-10- 95

0

DATE: October 20, 1955

SUBJECT: SOME INVERSE CHARACTERISTIC VALUE PROBLEMS
WHICH ARISE IN THE STUDY OF SIMPLE MOLECULES

TO: Listed Distribution

FROM: A. C. Downing and A. S. Householder

LEGAL NOTICE

This report was prepared as an account of Government sponsored work. Neither the United States, nor the Commission, nor any person acting on behalf of the Commission:

A. Makes any warranty or representation, express or implied, with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use of any information, apparatus, method, or process disclosed in this report may not infringe privately owned rights; or

B. Assumes any liabilities with respect to the use of, or for damages resulting from the use of any information, apparatus, method, or process disclosed in this report.

As used in the above, "person acting on behalf of the Commission" includes any employee or contractor of the Commission to the extent that such employee or contractor prepares, handles or distributes, or provides access to, any information pursuant to his employment or contract with the Commission.

Photostat Price \$ 3.30

Microfilm Price \$ 2.40

Available from the
Office of Technical Services
Department of Commerce
Washington 25, D. C.

UNCLASSIFIED

1 2501

292 1

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

1

SOME INVERSE CHARACTERISTIC VALUE PROBLEMS

A. C. Downing, Jr. and A. S. Householder
Mathematics Panel
Oak Ridge National Laboratory

1. Introduction: This paper discusses the numerical solution of certain matrix problems which are inverses of the familiar characteristic value problem.

Problem I: Given a Hermitian matrix A , find a real diagonal matrix D such that the equation $|A - \lambda D^2| = 0$ has prescribed roots. An equivalent formulation is to require that the Hermitian matrix $D^{-1} A D^{-1}$ have prescribed characteristic values.

Problem II: Given a Hermitian matrix A , find a real diagonal matrix D such that the matrix $A + D$ has prescribed characteristic values.

Mathematically these problems appear to be distinct. The first arises in the study of certain dynamical systems [1]; the second has applications in statistics. In the final section of this paper, a generalization of these problems is considered in which the unknown matrix D is not restricted to be diagonal, but rather is only assumed to be Hermitian.

Criteria for the existence of solutions do not appear to be known. For the simple case when A is a 2×2 matrix, it is easily shown that solutions do not exist for all prescribed λ 's, even when A is positive definite. In practical applications, however, circumstantial evidence may be sufficient to justify the assumption that at least one solution exists.

Let D_i be an approximation to such a solution. In the next section, a second-order iteration will be developed for improving D_i . This iteration requires that the characteristic values and vectors of $D_i^{-1} A D_i^{-1}$ (or $A + D_i$) be computed. A first-order iteration which is applicable when A has a dominant diagonal is also presented. This has the advantage that only characteristic values need be computed.

Since the convergence proofs are similar for both problems, the method will only be demonstrated for Problem I. The first- and second-order iterations for this problem are described by equations (11) and (10), respectively. The notation is defined in the first paragraph of the next section.

2. Let U be a unitary matrix whose columns are the characteristic vectors of $D_i^{-1} A D_i^{-1}$. Thus

$$(1) \quad D_i^{-1} A D_i^{-1} = U \Lambda (I + \epsilon E) U^*,$$

where Λ is the diagonal matrix of prescribed proper values, and $\epsilon \in E$ is a real diagonal matrix. The auxiliary parameter ϵ is introduced to facilitate the identification of matrices whose elements tend to zero as D_i converges to a solution. ϵ belongs to the real field as do the elements of all the diagonal matrices to be introduced. If a better approximation, D_{i+1} , is defined in terms of D_i and a diagonal matrix G by the relation

$$D_{i+1} = (I - \frac{1}{2} \epsilon G)^{-1} D_i ,$$

then Problem I may be rephrased as follows: How shall G be chosen in order that the diagonalized form of the matrix

$$(I - \frac{1}{2} \epsilon G) D_i^{-1} A D_i^{-1} (I - \frac{1}{2} \epsilon G)$$

more nearly approximates the required Λ ?

Let $U(\epsilon)$ be the unitary matrix associated with the diagonalization of this new matrix. Since $U^{-1} U(\epsilon)$ is unitary and tends to the identity as ϵ approaches zero, it is convenient to express $U^{-1} U(\epsilon)$ in a form explicitly displaying this property.

Lemma: The matrix $V = (I + i \epsilon H)(I - i \epsilon H)^{-1}$ is unitary if and only if ϵH is Hermitian.

It is easily verified that $V V^* = I$ for any Hermitian H .
(The magnitude of ϵ plays no role in this lemma.) Moreover, for

any V , H is uniquely determined. In fact,

$$\epsilon H = i(I - V)(I + V)^{-1} = i(I + V)^{-1} (I - V),$$

and hence,

$$\epsilon H^* = -i(I + V^*)^{-1} V^{-1} V (I - V^*) = \epsilon H.$$

It is possible, therefore, to represent $U(\epsilon)$ in the form

$$(2) \quad U(\epsilon) = U(I + i\epsilon H)(I - i\epsilon H)^{-1}.$$

If the diagonal form of $D_{i+1}^{-1} A D_{i+1}^{-1}$ is $\bigwedge (I + \epsilon F)$, where F is diagonal, then applying equations (1) and (2), one obtains

$$(3) \quad (I - \frac{1}{2}\epsilon G) U \bigwedge (I + \epsilon E) U^* (I - \frac{1}{2}\epsilon G) \\ = U(I + i\epsilon H)(I - i\epsilon H)^{-1} \bigwedge (I + \epsilon F)(I + i\epsilon H)^{-1}(I - i\epsilon H) U^*.$$

This is an identity in ϵ . It may be expanded and first-order terms in ϵ collected and set equal to zero. The result, after rearrangement and use of the commutativity of diagonal matrices, is

$$(4) \quad (-GU + UE - UF - 4iUH) \bigwedge U^* + U \bigwedge (-U^*G + EU^* - FU^* + 4iHU^*) = 0.$$

Multiplying from the left by U^* and from the right by U yields

$$(5) \quad (-U^*GU + E - F - 4iH)\Lambda + \Lambda(-U^*GU + E - F + 4iH) = 0.$$

The second term is the conjugate transpose of the first. Hence, the first term is $\sqrt{-1}$ times some Hermitian matrix K . The resultant relation is readily solved for F in terms of K :

$$(6) \quad F = E - U^*GU - 4iH - iK\Lambda^{-1}.$$

Since F is diagonal, it is equal to its conjugate transpose:

$$(7) \quad F = E - U^*GU + 4iH + i\Lambda^{-1}K.$$

Adding and subtracting these two equations yields

$$(8) \quad F = E - U^*GU + i(\Lambda^{-1}K - K\Lambda^{-1}),$$

$$(9) \quad 8H = -(\Lambda^{-1}K + K\Lambda^{-1}).$$

The diagonal elements of $\Lambda^{-1}K - K\Lambda^{-1}$ are identically zero, however, so that it follows from equation (8) that the significant

elements of F are the diagonal elements of $E - U^*GU$. This provides several possible criteria for selecting G . It must be demonstrated, however, that any G which is chosen by considering diagonal elements alone will be consistent with the off-diagonal elements of equation (5) or with the equivalent pair of equations (8) and (9).

Note first that the diagonal elements of H play no role in equation (5) since they cancel identically. In addition, it can be seen that if the λ 's are not all distinct, say if $\lambda_i = \lambda_j$, then the elements k_{ij} and k_{ji} also cancel in equation (5). But for any fixed G , the off-diagonal elements of K are uniquely determined by equation (8), provided the λ 's are distinct; and H may be determined from K by equation (9). Therefore, once G is chosen, the elements of H which are significant in equation (5) may be determined by the first-order terms in ϵ in a manner consistent with the choice of G .

In general, G may be selected so that $F = 0$, at least to within terms of order ϵ . To accomplish this, the diagonal elements g_i of G must satisfy the linear system of equations

$$(10) \quad \sum_{i=1}^n |u_{ji}|^2 g_j = e_i \quad (i = 1, \dots, n)$$

where the u_{ji} are elements of U and the e_i are diagonal elements of E . Since all terms of order ϵ are strictly accounted for and the diagonal form of $D_{i+1}^{-1} A D_{i+1}^{-1}$ is $\Lambda + O(\epsilon^2)$, this is a second-order iteration.

It may happen that the matrix of coefficients $|u_{ji}|^2$ is singular, even though a solution does exist. To obtain an improvement in the approximate solution, however, it is sufficient to select a G satisfying the inequalities

$$|f_i| \equiv \left| e_i - \sum_{j=1}^n |u_{ji}|^2 g_j \right| < \max_j |e_j| \quad (i = 1, \dots, n).$$

In some problems, the choice $g_i = e_i$ will satisfy these conditions. For

$$\begin{aligned} \left| e_i - \sum_{j=1}^n e_j |u_{ji}|^2 \right| &= \left| e_i(1 - |u_{ii}|^2) - \sum_{\substack{j=1 \\ j \neq i}}^n e_j |u_{ji}|^2 \right| \\ &\leq |e_i| (1 - |u_{ii}|^2) + \left(\max_{j \neq i} |e_j| \right) \sum_{\substack{j=1 \\ j \neq i}}^n |u_{ji}|^2 \\ &\leq 2 \left(\max_j |e_j| \right) (1 - |u_{ii}|^2). \end{aligned}$$

Hence, setting

$$(11) \quad G = E$$

yields a convergent first-order iteration*, provided the magnitudes of all the diagonal elements of U exceed $1/\sqrt{2}$.

* This iteration was originally suggested in discussions with Professor R. C. F. Bartels, of the University of Michigan. The plausibility argument he used was based upon the geometric interpretation of the matrix D^{-1} as a dilatation of the Hermitian form associated with A .

But when A has a sufficiently dominant diagonal, i.e.,
 $A = \Lambda + \epsilon K$, with $\epsilon \|K\| < \|A\|$, then it can be shown with the aid
of the lemma above that the diagonalizing unitary matrix has a dominant
diagonal.

3. Analogous results hold for Problem II. Indeed, if

$$A + D_i = U(\Lambda + \epsilon E) U^*,$$

and, if it is assumed that

$$D_{i+1} = D_i - \epsilon G,$$

then selecting G according to equation (10) or (11) will yield a
second- or a first-order iterative procedure, respectively. The proof
of this assertion proceeds exactly as in the analysis of Problem I,
except for the simplification that an auxiliary matrix K need not be
introduced.

4. The two problems already discussed can be generalized by
not restricting the unknown matrix to be diagonal. Corresponding to
Problem I, there is the following problem:

Problem III: Given a family of $n \times n$ Hermitian matrices

A_μ ($\mu = 1, \dots, m$), and with each matrix A_μ let
there be associated a set of real numbers $\lambda_v^{(\mu)}$

$$(1 \leq v \leq v_\mu \leq n) \text{ where } \sum_{\mu=1}^m v_\mu = \frac{1}{2} (n^2 + n);$$

the problem then is to find a Hermitian matrix B such that each of the polynomials $P_\mu(\lambda) \equiv |A_\mu - \lambda B^2|$ will have the prescribed zeros $\lambda_v^{(\mu)}$.

Note that when $v_\mu < n$, then all the zeros of $P_\mu(\lambda)$ are not specified. Also, if the total number of prescribed zeros for all the polynomials exceeds the number of unknowns in B , i.e., if

$$\sum_{\mu=1}^m v_\mu > \frac{1}{2} (n^2 + n),$$

then the problem is to find a best matrix B according to some criterion such as least square deviations of the roots.

Let B_1 be an approximate solution such that

$$B_i^{-1} A_\mu B_i^{-1} = U_\mu \bigwedge_\mu (I + \epsilon E_\mu) U_\mu^*, \quad (\mu = 1, \dots, m).$$

If a better approximation is

$$B_{i+1} = (I - \frac{1}{2} \epsilon G) B_i$$

where G is Hermitian but not necessarily diagonal, then for each μ the equation describing the diagonalization of the matrix $B_{i+1}^{-1} A_\mu B_{i+1}^{-1}$ leads to a relation which is identical with equation (3), except for

0-9

THIS PAGE
WAS INTENTIONALLY
LEFT BLANK