JOURNAL OF COMPUTATIONAL PHYSICS 65, 107–119 (1986)

The Local Evaluation of the Derivative of a Determinant*

THOMAS J. BRIDGES

Mathematics Research Center, 610 Walnut Street, Madison, Wisconsin 53705

AND

LEONID N. VASERSTEIN

Institute for Advanced Study, Princeton, New Jersey 08540

Received April 13, 1984; revised November 27, 1985

In this paper, the local evaluation of the derivative of a determinant of a λ-matrix is considered. The entries of a λ-matrix are scalar polynomials, of finite degree, in the independent variable λ. The existing methods with $O(N^3)$ operation counts, developed for non-singular matrices, are reviewed, and extended, where possible, to singular matrices. An alternative approach, similar in nature to the previous methods, based on direct selection of the necessary matrix entries, is suggested. A general expression, valid at both singular and non-singular points, is derived and then the simplifications to be found in special cases are discussed, and applications where the algorithms might be useful are given. © 1986 Academic Press, Inc.

1. INTRODUCTION

This paper is concerned with evaluating the derivative of the determinant of square matrices whose elements are polynomials in the relevant parameter. Matrices of this type may be represented in the form of a scalar polynomial with matrix coefficients,

$$\mathbf{D}_p(\lambda) = \sum_{n=0}^{p} \mathbf{B}_n \lambda^{p-n} \tag{1}$$

where the $\mathbf{B}_n$ are square matrices with complex entries, λ is a scalar, and p is the degree of the polynomial matrix. These matrices are often referred to as *lambda matrices* [1; Chap. 6, 2]. The lambda matrix is regular if $\mathbf{B}_0$ is invertible, and monic if $\mathbf{B}_0 = \mathbf{I}$, the identity matrix. Matrices of this type arise in the analysis of dynamical systems [2], the stability of fluid flows [3], and in bifurcation theory [4].

* Sponsored by the United States Army under Contract DAAG29-80-C-0041. This material is based upon work supported by the National Science Foundation under Grant DMS-8210950, Mod. 1.

0021-9991/86 $3.00

The eigenvalue problem associated with (1) is to find the roots of the determinant, $\Delta(\lambda)$, of $\mathbf{D}_p(\lambda)$

$$\Delta(\lambda) = |\mathbf{D}_p(\lambda)| = 0. \tag{2}$$

When $p = 1$, the expression (1) is the form for a linear eigenvalue problem, and when $p > 1$ it is the form for a *non-linear, in the parameter, eigenvalue problem.* If the order of the matrix is N, then the determinantal equation (2) results in a scalar polynomial of degree $\leqslant pN$. If $\mathbf{D}_p$ is regular the degree of the resulting scalar polynomial is exactly pN. Corresponding to each latent root, λ_k, of (2) there is a right and left eigenvector x_k and y_k of $\mathbf{D}_p(\lambda_k)$ which satisfy

$$\mathbf{D}_p(\lambda_k) \cdot x_k = \mathbf{0} \qquad \text{and} \qquad y_k^H \cdot \mathbf{D}_p(\lambda_k) = \mathbf{0}^H \tag{3}$$

where the superscript H denotes complex conjugate transpose.

A review of numerical methods for eigenvalue problems of this type is given by Lancaster [5]. Applications of these methods, as well as some recent developments, can be found in [6], and [7].

The major reason for this work was the necessity to evaluate the derivative of a determinant at points where the determinant is zero with a reasonable operation count, say $O(N^3)$ operations. The available methods for evaluating the derivative of a determinant (which will be discussed in the next section) had been developed for evaluation at non-singular points. Consequently, we developed a simple method for the expressed purpose of efficiently evaluating the derivative of a determinant at singular points. It was successfully used in a statistical analysis of the transition to turbulence in [8]. It was then generalized to include non-singular points as well, and applied to various problems. At non-singular points the approach is similar to other proposed methods.

Algorithms, of this type, for the local evaluation of the derivative of a determinant at non-singular points are useful as a basis for a Newton iteration scheme to find the eigenvalues of a lambda matrix. When inverting Fourier or Laplace transforms with matrix integrands, using residue theory, the poles occur at the zeros of a determinant. Evaluation of each residue requires the derivative of the determinant at singular points. The sum of these residues provides the necessary inversion of the integrals.

In Section 2, known methods for evaluating the derivative of a determinant are reviewed. These methods are extended where possible to the evaluation at singular points. Then, in Section 3 our alternative approach is presented. In Section 4 the special case of the first derivative at singular points is considered with applications. Then in Section 5 the special case of the first and second derivative at non-singular points, with applications, is considered.

2. Review of Existing Methods

The global derivative of a determinant can be constructed in the folowing way. Define $\mathbf{D}_{*j}(\lambda)$ as the jth column of $\mathbf{D}_{\mathbf{p}}(\lambda)$, and $\mathbf{D}_{\mathbf{p}}{}^{(1)}(\lambda)$ as the derivative of $\mathbf{D}_{\mathbf{p}}(\lambda)$ with respect to λ. Then, if N is the order of the matrix $\mathbf{D}_{\mathbf{p}}(\lambda)$, the derivative of the determinant of $\mathbf{D}_{\mathbf{p}}(\lambda)$ can be expressed as

$$\frac{d}{d\lambda}\Delta(\lambda)=\Delta^{(1)}(\lambda)=\sum_{j=1}^{N}\Delta_j(\lambda) \tag{4}$$

where $\Delta_j(\lambda)$ is a determinant whose jth column is $\mathbf{D}_{*j}{}^{(1)}(\lambda)$ and the remaining columns are those of $\mathbf{D}_{\mathbf{p}}(\lambda)$. With this expression, the evaluation of the derivative would require the ealuation of N determinants for each point λ_k, resulting in an operation count of order N^4.

However for the local evaluation of the derivative of a determinant methods with an operation count of $O(N^3)$ have been developed. The first of these methods is the method of Lancaster based on the trace theorem [2, p. 99]. The essential result of this theorem is that the local derivative of a determinant, at points where $\Delta(\lambda_k)\neq 0$, can be expressed as

$$\Delta^{(1)}(\lambda_k)=\Delta(\lambda_k)\cdot\mathrm{Trace}\{\mathbf{D}^{-1}(\lambda_k)\,\mathbf{D}^{(1)}(\lambda_k)\} \tag{5}$$

where $\mathrm{Trace}\{\mathbf{A}\}=\sum_{i=1}^{N}a_{ii}$. The major amount of work in evaluating (5) is the $O(N^3)$ operations required to form the inverse of $\mathbf{D}_{\mathbf{p}}(\lambda_k)$. The remaining work is in construcing $\mathbf{D}^{(1)}(\lambda_k)$, which is of $O(N^2)$ operations, and forming the trace of the product which requires and additional N^2 operations. However the $\mathrm{Trace}\{\mathbf{M}\}$, where $\mathbf{M}=\mathbf{D}^{-1}\mathbf{D}^{(1)}$ can be constructed with greater efficiency (although the operation count is still $O(N^3)$) by using the **LU** decomposition of **M**. Then solve $\mathbf{LUM}=\mathbf{D}^{(1)}$ in the usual way. **UM** is constructed by forward substitution with **L** and **M** is constructed by back substitution with **U**. With this procedure the lower part of **M** only is computed.

When $\Delta(\lambda_k)=0$, this expression may be cast in a different form,

$$\Delta^{(1)}(\lambda_k)=\mathrm{Trace}\{\mathrm{cof}(\mathbf{D}(\lambda_k))\,\mathbf{D}^{(1)}(\lambda_k)\} \tag{6}$$

where $\mathrm{cof}(\mathbf{A})$ denotes the cofactor matrix of **A**. The construction of a cofactor matrix for a singular matrix may be performed in a number of ways. If the singularity is due to a distinct eigenvalue, and the associated left and right eigenvectors are available, they can be used to construct the cofactor matrix. An example of this approach, for computing the cofactor matrix of a singular matrix, is given in Section 4. Otherwise elementary operations may be used to construct a zero row and column, then the cofactor matrix is a single entry matrix which is pre- and post-multiplied by the matrices containing the elementary operations. Lancaster has

also extended this result to the second derivative at non-singular point. The result, from [2, p. 84], is

$$\Delta^{(2)}(\lambda_k) = \Delta(\lambda_k) \cdot \text{Trace}\{\mathbf{D}^{-1}(\lambda_k)\,\mathbf{D}^{(2)}(\lambda_k) - (\mathbf{D}^{-1}(\lambda_k))\,\mathbf{D}^{(1)}(\lambda_k))^2\} + \{\Delta^{(1)}(\lambda_k)/\Delta(\lambda_k)\}^2. \tag{7}$$

Evaluation of this expression will require, in addition to the evaluation of $\Delta^{(1)}(\lambda_k)$, the formation of the product $\mathbf{D}^{-1}(\lambda_k)\,\mathbf{D}^{(1)}(\lambda_k)$, requiring N^3 operations, and the formation of $\mathbf{D}^{(2)}$ and the summing of the trace, both of which require $O(N^2)$ operations. The expressions (5), and (7) were used by Lancaster to construct two efficient Newton based iteration schemes for finding the eigenvalues of lambda matrices. The extension of the result (7) to singular points is not, however, obvious.

Another approach has been suggested by Kublanovskaya [9]. It is based on the *QR* algorithm. The essence of the method is as follows. The matrix $\mathbf{D}_p(\lambda)$ is decomposed into the product of a unitary matrix $\mathbf{Q}$ and an upper triangular matrix $\mathbf{R}$,

$$\mathbf{D}_p(\lambda) = \mathbf{Q}(\lambda) \cdot \mathbf{R}(\lambda) \tag{8}$$

where $\mathbf{Q}(\lambda)$ satisfies

$$\mathbf{Q}^H \cdot \mathbf{Q} = \mathbf{I}. \tag{9}$$

Taking the determinant of $\mathbf{Q}$ to be unity, the determinant of (8) is

$$\Delta(\lambda) = \prod_{j=1}^{N} r_{jj}(\lambda) \tag{10}$$

with derivative

$$\Delta^{(1)}(\lambda) = \sum_{i=1}^{N} \frac{dr_{ii}}{d\lambda} \prod_{\substack{j=1 \\ i \neq j}}^{N} r_{jj} \tag{11}$$

but dividing (11) by (10)

$$\Delta^{(1)}(\lambda) = \Delta(\lambda) \sum_{i=1}^{N} \frac{1}{r_{ii}} \frac{dr_{ii}}{d\lambda}. \tag{12}$$

It remains now to find an expression for $dr_{ii}/d\lambda$. Taking the derivative of (8), premultiplying by $\mathbf{Q}^H$ and post-multiplying by $\mathbf{R}^{-1}$ results in

$$\mathbf{Q}^H \mathbf{D}_p{}^{(1)}(\lambda)\,\mathbf{R}^{-1}(\lambda) = \mathbf{Q}^H(\lambda)\,\mathbf{Q}^{(1)}(\lambda) + \mathbf{R}^{(1)}(\lambda)\,\mathbf{R}^{-1}(\lambda). \tag{13}$$

However, from (9), the product $\mathbf{Q}^H\mathbf{Q}^{(1)}$ is skew-symmetric. Therefore the trace of (13) results in

$$\sum_{i=1}^{N} \frac{1}{r_{ii}} \frac{dr_{ii}}{d\lambda} = \text{Trace}\{\mathbf{Q}^H \mathbf{D}_p{}^{(1)}(\lambda)\,\mathbf{R}^{-1}(\lambda)\}. \tag{14}$$

This may then be used in (12) to evaluate the derivative. This is the essence of Kublanovskaya's approach. There are additional details and the reader is referred to her paper [9] for complete details. The operation count for this method is dominated by the orthogonal decomposition in (8). It requires $O(N^3)$ operations. However there is the additional work of forming the trace of the matrix triple product in (14), which requires $O(N^3)$ operators. Ruhe [6] gives a description of this method, as well as a variation of it, as a basis for Newton iteration for obtaining the roots of lambda matrices. Although the approach of Kublanovskaya is potentially interesting it has the drawback that it fails when $\mathbf{D}_p(\lambda)$ is singular, and no analogy to this method for singular $\mathbf{D}_p(\lambda)$ has been found. When $\mathbf{D}_p(\lambda)$ is singular, the orthogonal decomposition may not be unique and therefore not differentiable with respect to λ.

3. An Alternative Approach for the Derivative

An alternative approach for the local evaluation of the derivative of a determinant may be constructed. This approach is based on expanding the lambda matrix in a Taylor series about the desired point,

$$\mathbf{D}_p(\lambda) = \sum_{n=0}^{p} \mathbf{F}_n(\lambda - \lambda_k)^n. \tag{15}$$

No approximation has been made, merely a recasting, since the degree is finite. The matrices $\mathbf{F}_n$ are the Taylor matrices

$$\mathbf{F}_n = \frac{\mathbf{D}_p{}^{(n)}(\lambda_k)}{n!}. \tag{16}$$

When the matrix is cast in this form, the determinant of the matrix will result in a scalar polynomial in $(\lambda - \lambda_k)$,

$$\Delta(\lambda) = d_0 + d_1(\lambda - \lambda_k) + d_2(\lambda - \lambda_k)^2 + \cdots \tag{17}$$

where the d_j are the scalars,

$$\begin{aligned} d_0 &= \Delta(\lambda_k) \\ d_1 &= \Delta^{(1)}(\lambda_k) \\ d_2 &= \tfrac{1}{2}\Delta^{(2)}(\lambda_k) \end{aligned} \tag{18}$$

etc.. Therefore, the object is to arrange the matrices $\mathbf{F}_0$ and $\mathbf{F}_1$ (and $\mathbf{F}_2$, etc. for derivatives higher than the first) in a convenient form such that the scalars d_j may be obtained in a simple manner. Here a general procedure for finding the scalar d_1

will be given, and in the subsequent sections the special case of singular and non-singular points and higher derivatives will be considered.

For the first derivative only the matrices $\mathbf{F}_0$ and $\mathbf{F}_1$ are necessary. The matrix $\mathbf{D}_p(\lambda)$ is operated on by elementary operations, represented by the matrices $\mathbf{E}_1$ and $\mathbf{E}_2$,

$$\mathbf{E}_1\mathbf{D}_p(\lambda)\ \mathbf{E}_2 = \mathbf{E}_1\mathbf{F}_0\mathbf{E}_2 + \mathbf{E}_1\mathbf{F}_1\mathbf{E}_2(\lambda - \lambda_k) + \cdots. \tag{19}$$

The matrices $\mathbf{E}_1$ and $\mathbf{E}_2$ are chosen such that $\mathbf{F}_0$ is brought to a diagonal form,

$$\mathbf{A} = \mathbf{E}_1\mathbf{F}_0\mathbf{E}_2 = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{NN} \end{pmatrix} \tag{20}$$

with the same operations applied to $\mathbf{F}_1$. Defining $\mathbf{B} = \mathbf{E}_1\mathbf{F}_1\mathbf{E}_2$, then

$$\mathbf{E}_1\mathbf{D}_p(\lambda)\,\mathbf{E}_2 = \mathbf{A} + \mathbf{B}(\lambda - \lambda_k) + \cdots. \tag{21}$$

If $\Delta(\lambda_k) = 0$, the matrix $\mathbf{A}$ will be singular and the diagonal form will be modified by a zero entry. This case, in which the procedure may be simplified further, is discussed in Section 4. With the matrix in the modified form, (21), it is necessary to find the proper combination of entries of $\mathbf{A}$ and $\mathbf{B}$ that are linear in $(\lambda - \lambda_k)$ when the determinant is formed. It is straightforward to show that the result is

$$\Delta^{(1)}(\lambda_k) = \sum_{i=1}^{N} b_{ii} \prod_{\substack{j=1 \\ i \neq j}}^{N} a_{ij}. \tag{22}$$

This is the general expression for the derivative valid at both singular and non-singular points. When $\Delta(\lambda_k) = 0$ there will be at most one non-zero term in the sum. Consequently the expression can be simplified for this and other special cases. There is an obvious similarity between this expression and (10), derived using Kublanovskaya's approach. However, in her method, an expression for $dr_{ii}/d\lambda$ is not obtainable. Therefore one must use the ratio $(1/r_{ii})(dr_{ii}/d\lambda)$, derived in (14), which fails at singular points.

4. The Local Derivative When $\Delta(\lambda_k) = 0$

A simplification of the above derived procedure is possible at points where $\Delta(\lambda_k) = 0$. When $\Delta(\lambda_k) = 0$ the determinant of the matrix $\mathbf{F}_0$ is zero. Therefore it is not necessary to bring this matrix to a diagonal form. If the singularity is due to a

simple root, an approach is to modify $\mathbf{E}_1$ and $\mathbf{E}_2$, the elementary matrices, such that $\mathbf{A}$ becomes

$$\mathbf{A}=\mathbf{E}_1\mathbf{F}_0\mathbf{E}_2=\begin{pmatrix}0 & 0 & 0 & \cdots & 0\\ 0 & a_{22} & a_{23} & \cdots & a_{2N}\\ 0 & 0 & a_{33} & \cdots & a_{3N}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \cdots & a_{NN}\end{pmatrix}. \tag{23}$$

In this case the derivative will be

$$\Delta^{(1)}(\lambda_k)=b_{11}\prod_{j=1}^{N}a_{jj}. \tag{24}$$

however if the normalized eigenvectors are available the following general expression may be derived. $\mathbf{E}_1$ and $\mathbf{E}_2$ can be chosen such that $\mathbf{A}$ becomes

$$\mathbf{A}=\mathbf{E}_1\mathbf{F}_0\mathbf{E}_2=\begin{pmatrix}a_{11} & a_{12} & \cdots & 0 & \cdots & a_{1N}\\ a_{21} & a_{22} & \cdots & 0 & \cdots & a_{2N}\\ \vdots & \vdots & & \vdots & & \vdots\\ 0 & 0 & \cdots & 0 & \cdots & 0\\ \vdots & \vdots & & \vdots & & \vdots\\ a_{N1} & a_{N2} & \cdots & 0 & \cdots & a_{NN}\end{pmatrix} \tag{25}$$

with the Ith row and the Jth column zero. With the same operations applied to $\mathbf{F}_1$ resulting in $\mathbf{B}$, then $\Delta^{(1)}(\lambda_k)$ can be expressed as

$$\Delta^{(1)}(\lambda_k)=b_{IJ}\,|\mathbf{A}_{\mathbf{IJ}}| \tag{26}$$

where $|\mathbf{A}_{\mathbf{IJ}}|$ is the determinant of the matrix $\mathbf{A}$ neglecting the zero row and column. Since λ_k is an eigenvalue of $\mathbf{D}_{\mathbf{p}}(\lambda)$, the corresponding right and left eigenvectors can be used to create the zero row and column. If the eigenvectors are normalized such that Ith entry of $\mathbf{y}_\mathbf{k}$, the left eigenvector, is unity and the Jth entry of $\mathbf{x}_\mathbf{k}$, the right eigenvector is unity, then

$$\mathbf{E}_1=\begin{pmatrix}1 & 0 & \cdots & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots & & \vdots\\ y_1^* & y_2^* & \cdots & 1 & \cdots & y_N^*\\ \vdots & \vdots & & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 0 & \cdots & 1\end{pmatrix} \tag{27a}$$

and

$$\mathbf{E}_2 = \begin{pmatrix} 1 & 0 & \cdots & x_1 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_N & \cdots & 1 \end{pmatrix} \tag{27b}$$

where the starred Ith row in $\mathbf{E}_1$ is the complex conjugate of $\mathbf{y}_\mathbf{k}$ and the Jth column of $\mathbf{E}_2$ is $\mathbf{x}_\mathbf{k}$. If the eigenvalue is known the eigenvectors can be obtained using inverse iteration in $O(N^2)$ operations. However in the absence of the eigenvectors a row echelon form (with at least one zero row at the bottom) can be found using only row operations. Then in (25) $\mathbf{E}_2 = \mathbf{I}$ and $\mathbf{E}_1$ would contain the set of row operations.

An algorithm of this sort may be used to invert Fourier transforms with matrix integrands. For example, consider the following initial-boundary value problem,

$$\left(\frac{\partial^2}{\partial x^2} EI(x) \frac{\partial^2}{\partial x^2} + C\frac{\partial}{\partial t} + \rho A(x) \frac{\partial^2}{\partial t^2}\right) \phi = f(x, t) \tag{28}$$

$$\phi(x, 0) = 0, \qquad \phi_t(x, 0) = 0 \tag{29a}$$

$$\phi(-1), t) = 0, \qquad \phi_x(-1, t) = 0 \tag{29b}$$

$$\phi_{xx}(+1, t) = 0, \qquad \phi_{xxx}(+1, t) = 0. \tag{29c}$$

This the governing equation for the non-uniform cantilevered beam under distributed loading [10, p. 389]. The loading $f(x, t)$ may be deterministic or random. E is Youngs modulus, $I(x)$ is the distributed moment of inertia, C is a damping parameter, ρ is the material density, $A(x)$ is the cross-sectional area, and $\phi(x, t)$ is the transverse displacement of the beam. The length of the beam is 2 and $x \in [-1, +1]$. To solve this problem, a Fourier transform is used in time, where the transform mates are defined as

$$\hat{\phi}(x; \omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \phi(x, t)\, e^{i\omega t}\, dt \tag{30a}$$

$$\phi(x, t) = \int_{-\infty}^{+\infty} \hat{\phi}(x; \omega)\, e^{-i\omega t}\, d\omega. \tag{30b}$$

Taking the Fourier transform of (28) results in the following parameter dependent ordinary differential equation,

$$\frac{\partial^2}{\partial x^2} EI(x) \frac{\partial^2}{\partial x^2} \hat{\phi} - i\omega C\hat{\phi} - \omega^2 \rho A(x)\, \hat{\phi} = \hat{f}(x; \omega) \tag{31}$$

with the boundary conditions (29a), (29b) in the frequency domain. The coefficients $I(x)$ and $A(x)$, the forcing function $\hat{f}(x;\omega)$, and the solution $\hat{\phi}(x;\omega)$ are expanded in finite Chebyshev series. For example,

$$\hat{\phi}(x;\omega) = \tfrac{1}{2}\hat{\phi}_0(\omega) + \sum_{n=1}^{N-1} \hat{\phi}_n(\omega)\, T_n(x) \tag{32a}$$

$$\hat{f}(x;\omega) = \tfrac{1}{2}\hat{f}_0(\omega) + \sum_{n=1}^{N-1} \hat{f}_n(\omega)\, T_n(x) \tag{32b}$$

where the $T_n(x)$ are the Chebyshev polynomials. Substituting (32) and the expansions for $I(x)$ and $A(x)$ into (31), results in a matrix equation for the vector $\{\hat{\phi}\}$,

$$[\mathbf{A_0}\omega^2 + \mathbf{A_1}\omega + \mathbf{A_2}]\{\hat{\phi}_n\} = [\mathbf{A_3}]\{\hat{f}_n\}. \tag{33}$$

Details of this conversion from the differential equation to the matrix equation, as well as further details about Chebyshev polynomials, can be found in [7, 11]. Equation (33) may be written as

$$[\mathbf{D_2}(\omega)]\{\hat{\phi}_n\} = [\mathbf{A_3}]\{\hat{\phi}_n\} \tag{34}$$

inverting $\mathbf{D_2}(\omega)$ and using the inverse Fourier transform, the solution $\phi(x, t)$ is

$$\phi(x, t) = \tfrac{1}{2}\phi_0(t) + \sum_{n=1}^{N-1} \phi_n(t)\, T_n(x) \tag{35}$$

where the vector ϕ_n is given by

$$\{\phi_n\} = \int_{-\infty}^{+\infty} \frac{[\mathbf{C}(\omega)][\mathbf{A_3}]\{\hat{f}_n\}}{\Delta(\omega)}\, e^{-i\omega t}\, d\omega \tag{36}$$

where $\mathbf{C}(\omega)$ and $\Delta(\omega)$ are the cofactor matrix and determinant respectively of $\mathbf{D_2}(\omega)$. To invert the integral in (36), a large semi-circle with infinite radius is used in the lower half plane. Therefore the integral is just the sum of the residues of the integrand. If the vector $\{\hat{f}_n\}$ is taken to be entire in ω (if not, the residues due to $\{\hat{f}_n\}$ are also included in the sum), and the roots, $\Delta(\omega_k) = 0$, have unit multiplicity, then the solution vector takes the form

$$\{\phi_n\} = 2\pi i \sum_{k=1}^{M} \frac{[\mathbf{C}(\omega_k)][\mathbf{A_3}]\{\hat{f}_n\}}{\Delta^{(1)}(\omega_k)}\, e^{-i\omega_k t}. \tag{37}$$

Since the matrix equation (33) has order N, there are $M \leqslant 2N$ residues. In most practical cases only the first few terms are significant due to the exponential weighting function. It remains to consider the form for the cofactor matrix. However, if $\mathbf{D_2}(\omega)$ is converted to the Taylor form and the $\mathbf{E_1}$ and $\mathbf{E_2}$ matrices,

appropriate for the singular case, are applied, the leading matrix will have the form of **A**, defined in Eq. (25). Then in the limit as $\omega \to \omega_k$, the cofactor matrix becomes

$$\mathbf{C}(\omega_k) = \mathbf{E}_1 \mathbf{\Upsilon}(\omega_k) \mathbf{E}_2 \tag{38}$$

where,

$$\mathbf{\Upsilon}(\omega_k) = \begin{pmatrix} 0 & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & v_{IJ} & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}. \tag{39}$$

This solo entry is $v_{IJ} = |\mathbf{A}_{\mathbf{IJ}}|$, which is defined in Eq. (26).

This model problem has been used as a simple example to illustrate contextually the utility of the result derived in Section 3. In a statistical analysis of the transition to turbulence, Bridges and Morris [8] use the above approach to efficiently invert multidimensional Fourier transforms.

5. The Local Derivative at Points Where $\Delta(\lambda_k) \neq 0$

At points where $\Delta(\lambda_k) \neq 0$, the matrix $\mathbf{D}_\mathbf{p}(\lambda)$ is brought to the form

$$\mathbf{E}_1 \mathbf{D}_\mathbf{p}(\lambda) \mathbf{E}_2 = \mathbf{A} + \mathbf{B}(\lambda - \lambda_k) + \cdots \tag{40}$$

where $\mathbf{E}_1$ and $\mathbf{E}_2$ are chosen such that **A** is the diagonal form given in (20). However, the determinant of $\mathbf{D}_\mathbf{p}(\lambda)$ at $\lambda = \lambda_k$ may be used in this case. It is the product of the diagonal entries in **A**,

$$\Delta(\lambda_k) = \prod_{j=1}^{N} a_{jj}. \tag{41}$$

Factoring this out of the general expression (22) results in

$$\Delta^{(1)}(\lambda_k) = \Delta(\lambda_k) \sum_{i=1}^{N} \frac{b_{ii}}{a_{ii}} \tag{42}$$

which is similar to the result found using the trace theorem (5) as well as the approach using the orthogonal decomposition (12) and (14). The evaluation of this expression requires the matrix setup, which is of $O(N^2)$, the application of the elementary operations, which is of $O(N^3)$, and the sum (42), which is of $O(N)$. It is therefore equivalent to the operation count for the trace theorem method and the orthogonal decomposition method.

In order to obtain the second derivative of the determinant, the same elementary operations are applied to $\mathbf{D}_p(\lambda)$ but it is necessary to retain the next order in the Taylor expansion,

$$\mathbf{E}_1\mathbf{D}_p(\lambda)\,\mathbf{E}_2 = \mathbf{A} + \mathbf{B}(\lambda - \lambda_k) + \mathbf{C}(\lambda - \lambda_k)^2 + \cdots \tag{43}$$

where $\mathbf{C} = \mathbf{E}_1\mathbf{F}_2\mathbf{E}_2$ and $\mathbf{F}_2$ is defined in (16), and $\mathbf{A}$ is diagonal and non-singular. However if $p = 1$, the matrix $\mathbf{F}_2$ is zero. In order to determine $\Delta^{(2)}(\lambda_k)$ the combination of entries in the modified matrix (43) which results in $2d_2$, defined in Eqs. (17) and (18), is needed. It is readily shown, by induction, that the second derivative of the determinant of $\mathbf{D}_p(\lambda)$ at $\lambda = \lambda_k$ is

$$\Delta^{(2)}(\lambda_k) = 2\Delta(\lambda_k) \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \frac{b_{ii}b_{jj} - b_{ij}b_{ji}}{a_{ii}a_{jj}} + 2\Delta(\lambda_k) \sum_{i=1}^{N} \frac{c_{ii}}{a_{ii}} \tag{44}$$

which may be contrasted with the form (7) derived using the trace theorem. The expression (44) is interesting in that once the elementary operations $\mathbf{E}_1$ and $\mathbf{E}_2$ have been performed, the only additional work over that for the first derivative, is the double sum in (44) which requires *less* than N^2 operations since it is a progressive double sum. In addition if the λ-matrix is linear in λ, the c_{ii} are zero and the first sum only is required.

Although the evaluation of the first derivative using this approach is about equivalent in operation count to the methods in Section 2, the evaluation of the second derivative requires less operations because of the direct selection of the matrix entries which are second order in $(\lambda - \lambda_k)$.

The expression (44) also has an analogue for the case $\Delta(\lambda_k) = 0$. Although a general expression may be derived, assume the matrix $\mathbf{A}$ has been brought to a diagonal form with $a_{11} = 0$. Then the analogue of (44) when $\Delta(\lambda_k) = 0$ is

$$\Delta^{(2)}(\lambda_k) = 2 \prod_{i=2}^{N} a_{ii} \left\{ c_{11} + \sum_{m=2}^{N} \frac{b_{11}b_{mm} - b_{1m}b_{m1}}{a_{mm}} \right\}. \tag{45}$$

Following Lancaster the expressions for the derivative of the determinant at non-singular points may be used to derive Newton iteration schemes for calculating the eigenvalues of λ-matrices.

The eigenvalues of the matrix $\mathbf{D}_p(\lambda)$ are given by the zeros of its determinant, and Newton's method may be used on the equation $\Delta(\lambda) = 0$. If an initial estimate is available, then Newton's method can be used to refine this value,

$$\lambda_{s+1} = \lambda_s - \frac{\Delta(\lambda_s)}{\Delta^{(1)}(\lambda_s)} \qquad \text{for} \quad s = 0, 1, 2, \ldots. \tag{46}$$

However, using Eq. (42) to define

$$d_1 = \frac{\Delta^{(1)}(\lambda_s)}{\Delta(\lambda_s)} = \sum_{i=1}^{N} \frac{b_{ii}}{a_{ii}} \tag{47}$$

then

$$\lambda_{s+1} = \lambda_s - \frac{1}{d_1} \qquad \text{for} \quad s = 0, 1, 2, \ldots \tag{48}$$

where the a_{ii} and b_{ii} are as previously defined. This algorithm is quadratically convergent to regular roots. It can be used to determine the eigenvalues of both linear and nonlinear matrix eigenvalue problems.

An improvement of Newton's method, with better convergence rates, was first suggested by Taylor [12]. It involves using the second derivative of the function whose zero is sought. For instance, expanding $\Delta(\lambda)$ in a Taylor series about $\lambda = \lambda_k$,

$$\Delta(\lambda) = \Delta(\lambda_k) + \Delta^{(1)}(\lambda_k)(\lambda - \lambda_k) + \tfrac{1}{2}\Delta^{(2)}(\lambda_k)(\lambda - \lambda_k)^2 + \cdots \tag{49}$$

and truncating the exansion after the first term results in the usual Newton algorithm,

$$\lambda - \lambda_k = -\frac{\Delta(\lambda_k)}{\Delta^{(1)}(\lambda_k)}. \tag{50}$$

However, truncating after the second term results in

$$\lambda - \lambda_k = -\frac{\Delta(\lambda_k)}{\Delta^{(1)}(\lambda_k) + (1/2)\Delta^{(2)}(\lambda_k)(\lambda - \lambda_k)} \tag{51}$$

which is implicit in the unknown $\lambda - \lambda_k$. However, the $\lambda - \lambda_k$ on the right-hand side of (51) can be approximated by (50). This substitution results in the formula first derived by Taylor,

$$\lambda = \lambda_s - \frac{1}{d_1 - d_2/d_1} \qquad \text{for} \quad s = 0, 1, 2, \ldots \tag{52}$$

where d_1 is given in (47) and

$$d_2 = \frac{1}{2}\frac{\Delta^{(2)}(\lambda_s)}{\Delta(\lambda_s)} \tag{53}$$

and $\Delta^{(2)}(\lambda_k)$ is given in (44). Lancaster [2] points out that this modified Newton algorithm is cubically convergent to regular roots. For linear eigenvalue problems, the second term in (44) is zero. One startling fact is that for linear eigenvalue problems, the cubically convergent method requires virtually the same amount of work as the quadratically convergent method. The only additional work involves the summing of the first term in the expression (44), which is less than N^2 operations. These two methods can be compared to the methods of Lancaster based on the trace theorem [2, p. 83]. The quadratically convergent method here is about equal to Lancasters quadratically convergent method in operation count, but this

cubically convergent method may be slightly faster due to the direct evaluation of the second derivative using (44).

Instead of approximating the quadratic term in (49) using (50), the quadratic formula may be used on

$$\Delta(\lambda_k)+\Delta^{(1)}(\lambda_k)(\lambda-\lambda_k)+\tfrac{1}{2}\Delta^{(2)}(\lambda_k)(\lambda-\lambda_k)^2=0 \tag{54}$$

resulting in the iteration scheme,

$$\lambda=\lambda_s+\frac{2}{-d_1\pm\sqrt{d_1^2-4d_2}} \qquad \text{for} \quad s=0,\,1,\,2,\ldots. \tag{55}$$

This is, in fact, a variant of Mullers method (Conte [13, p. 65] has an elaboration of Mullers method). Conte suggests that the sign before the radical in (55) should be chosed such that the denominator is largest in magnitude. Traub [14, p. 211] shows that the convergence of Mullers method is "almost quadratic" and also suggests some improvements over Mullers original formulation. It is possible that (55) converges at a better rate than Mullers method because in Mullers method the derivatives are not calculated exactly; they are approximated using divided differences.

ACKNOWLEDGMENTS

The authors would like to thank the reviewer for valuable suggestions concerning the paper; in particular the use of the LU decomposition for Eq. (5) and [5, 6, 9].

REFERENCES

1. F. R. GANTMACHER, *The Theory of Matrices, Vol. I* (Chelsea, New York, 1959).
2. P. LANCASTER, *Lambda Matrices and Vibrating Systems* (Permagon, Oxford, 1966).
3. T. J. BRIDGES AND P. J. MORRIS, AIAA Paper 84-0437, Aerospace Sciences Meeting, Jan. 1984 (unpublished).
4. T. J. BRIDGES, Math. Res. Center Report 2839, 1985 (unpublished).
5. P. LANCASTER, *Int. Schr. Numer. Math. (Birkhauser)*, **38**, 43 (1977).
6. A. RUHE, *SIAM J. Numer. Anal.* **10**, 674 (1973).
7. T. J. BRIDGES AND P. J. MORRIS, *J. Comput. Phys.* **55**, 437 (1984).
8. T. J. BRIDGES AND P. J. MORRIS, A statistical analysis of the effect of freestream turbulence on the Blasius boundary layer, in preparation.
9. V. N. KUBLANOVSKAYA, *SIAM J. Numer. Anal.* **7**, 532 (1970).
10. I. ELISHAKOFF, *Probabilistic Methods in the Theory of Structures* (Wiley-Interscience, New York, 1983).
11. L. FOX AND I. PARKER, *Chebyshev Polynomials in Numerical Analysis* (Oxford Univ. Press, London, 1968).
12. B. TAYLOR, *Philos. Trans. Royal Soc. London* **30**, 610 (1717).
13. S. D. CONTE, *Elemantary Numerical Analysis* (McGraw-Hill, New York, 1965).
14. J. F. TRAUB, *Iterative Methods for the Solution of Equations* (Prentice-Hall, Englewood Cliffs, N.J., 1964).