Partial Derivatives of Repeated Eigenvalues and Their Eigenvectors

Uwe Prells* and Michael I. Friswell†
University of Wales, Swansea SA2 8PP, Wales, United Kingdom

The analysis of inverse problems in linear modeling often require the sensitivities of the eigenvalues and eigenvectors. The calculation of these sensitivities is mathematically related to the corresponding partial derivatives, which do not exist for any parameterization. Inasmuch as eigenvalues and eigenvectors are coupled by the constitutive equation of the general eigenvalue problem, their derivatives are coupled, too. Conditions on the parameterization are derived and formulated as theorems, which ensure the existence of the partial derivatives of the eigenvalues and eigenvectors with respect to these parameters. The application of the theorems is demonstrated by examples.

I. Introduction

ANY engineering optimization problems, for instance, optimal design or model updating, lead to a sensitivity analysis of the eigenvalue problem. In the case of distinct eigenvalues, the partial derivatives of the eigenvalues and eigenvectors with respect to a prechosen parameterization can be calculated. The situation is more complicated in the case of multiple eigenvalues because the associated subspace of eigenvectors is only defined up to an arbitrary orthonormal matrix. This case has been the subject of many studies in the recent years (see, for instance, Refs. 1–13). As demonstrated by Haug and Rousselet, the (Fréchet) derivatives of the multiple eigenvalues do not exist, in general, for any parameterization in the case of more than one parameter. They suggest the use of the directional (Gateaux) derivatives, which exist for some directions. The question of which parameterization is permissible to ensure the existence of the partial derivatives of the eigenvalues and eigenvectors has not been discussed. A remark on that issue is the purpose of this paper. In Sec. II the problem of calculating the partial derivatives of repeated eigenvalues and the partial derivatives of the corresponding eigenvectors is recalled. Conditions on the existence of the partial derivatives of the eigenvalues and eigenvectors are investigated in Sec. III, accompanied by simple two-dimensional examples. To demonstrate the application of the theorems derived, in the fourth section an example of a three-dimensional elastomechanical model is presented.

II. Recalling the Problem

Consider the general eigenvalue problem

\[ AX\lambda = BX \in \mathbb{R}^{N \times N} \] (1)

with the normalization of the eigenvectors

\[ X^T AX = I_N \] (2)

Both equations are equivalent to

\[ A = (X^T)^{-1} \] (3)

\[ B = (X^T A^{-1} X)^{-1} \] (4)

with the real-valued symmetric and positive definite \( N \times N \) matrices \( A \) and \( B \) and the diagonal matrix \( \Lambda \) of the eigenvalues \( \lambda_i, i = 1, \ldots, N \). If \( A(q) \) and \( B(q) \) are given functions of the parameter vector \( q \in \mathbb{R}^r \) such that for each \( q \in \mathbb{R}^r \) decompositions (3) and (4) of \( A(q) \) and \( B(q) \) exist, then, of course, the eigenvectors and eigenvalues will depend on \( q \). If at parameter vector \( q^* \) the first \( n \leq N \) eigenvalues are equal, i.e.,

\[ \Lambda(q^*) = \begin{bmatrix} \ell_1 & 0 & \cdots & 0 \\
0 & \ell_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \ell_n \end{bmatrix} \] (5)

then the primary partition \( X_1 = [X_1, X_2] \) of the matrix of eigenvectors \( X(q^*) \) is defined only up to postmultiplication by an \( n \times n \) orthogonal matrix \( \Theta \). Indeed, from Eqs. (3) and (4), substituting \( X_1 \rightarrow X_1 \Theta \) gives

\[
X(q^*)^T(q^*) = X(q^*) \begin{bmatrix} \Theta & 0 \\
0 & I_{N-n} \end{bmatrix} \begin{bmatrix} \Theta^T & 0 \\
0 & I_{N-n} \end{bmatrix} X^T(q^*) \]

\[ X(q^*) A^{-1}(q^*) X(q^*) = X(q^*) \begin{bmatrix} \Theta & 0 \\
0 & I_{N-n} \end{bmatrix} \begin{bmatrix} \Theta^T & 0 \\
0 & I_{N-n} \end{bmatrix} X^T(q^*) \] (6)

For some parameterizations this redundancy can be used to ensure the existence of the partial derivatives of the eigenvalues at \( q^* \). Partial differentiation of Eqs. (3) and (4) with respect to the \( r \)th component \( q_r \) of \( q \) leads to

\[
\mathcal{A}_r := X^T A_r X = -(Z_r, Z_r^T) \] (8)

\[
\mathcal{B}_r := X^T B_r X = -(AZ_r - A_r Z_r^T A) \] (9)

where the subscript \( r \) denotes the partial derivative with respect to \( q_r \) and the matrix \( Z_r \) is defined by

\[ X_r(q) = X(q) Z_r(q), \quad \forall r = 1, \ldots, m \] (10)

Note that Eqs. (8) and (9) imply that, for all \( r = 1, \ldots, m \),

\[ \mathcal{A}_r \text{ symmetric } \Rightarrow A_r \text{ symmetric} \] (11)

\[ \mathcal{B}_r \text{ symmetric } \Rightarrow B_r \text{ symmetric} \] (12)

Using Eq. (8) to eliminate \( Z_r \) in Eq. (9) leads to

\[ [Z_r; A] + A_r = \mathcal{B}_r - \mathcal{A}_r A \equiv C \] (13)

where the commutator product is defined by

\[ [A; B] := AB - BA \] (14)

Considering only the diagonal of Eq. (8) leads to

\[ (Z_{r})_{\text{diag}} = -\frac{1}{2} \mathcal{A}_{r} (\Lambda_{r})_{\text{diag}} \] (15)

Inserting this result into the diagonal of Eq. (9), \( \Lambda_r \) can be calculated, yielding

\[ \Lambda_r = (\mathcal{B}_r - \mathcal{A}_r \Lambda_{r})_{\text{diag}} = (C)_{\text{diag}} \] (16)
This can be done for all \( r = 1, \ldots, m \) and also at \( q \equiv q' \). The problem is that, at \( q \equiv q' \), Eq. (13) requires the first \( n \times n \) diagonal block \( C^1 \) of
g
to be diagonal for all \( r = 1, \ldots, m \), i.e.,

\[
C^1_r(q') = \begin{bmatrix}
C^1_{i,j} \\
C^1_{j,i}
\end{bmatrix}
\tag{17}
\]

Only if Eq. (18) holds can the associated part of \( \Lambda \) be understood as the corresponding partial derivative. Note that, at \( q \equiv q' \), first the first \( n \times n \) diagonal block \( Z^{11} \) of \( Z \), the first partial derivative of the eigenvectors can be calculated from Eq. (13). Partitioning of \( Z(q') \) according to the partition of \( \Lambda(q') \) in Eq. (5) and of \( C(q') \) in Eq. (17),

\[
Z_r(q') = \begin{bmatrix}
Z^{11}_{r} \\
Z^{12}_{r} \\
Z^{21}_{r} \\
Z^{22}_{r}
\end{bmatrix}
\tag{19}
\]

and using Eq. (13) leads to

\[
Z^{12}_{r} = C^{12}_r (\Gamma - \ell I_{N - s})^{-1}
\tag{20}
\]

\[
Z^{21}_{r} = - (\Gamma - \ell I_{N - s}) C^{11}_r
\tag{21}
\]

and for \( N - n > 1 \)

\[
(Z^{ij}_{r})_{ij} = \frac{C^{ij}_r}{\gamma_i - \gamma_j}, \quad i \neq j, \quad i, j \in \{1, \ldots, N - n \}
\tag{22}
\]

where \( \gamma = (\Gamma)_{ii} \). The expression in Eq. (22) is equivalent to the coefficients of the standard series approach for nonrepeated eigenvalues. Note that the diagonal of \( Z^{11}_{r} \) and of \( Z^{21}_{r} \) is given by Eq. (15). To calculate the off-diagonal part \( (Z^{ij}_{r})_{off} \) of \( Z^{11}_{r} \), the second partial derivatives of the eigenvalue problem are needed. Here the subscript \( o \) denotes a matrix with a zero diagonal.

Differentiating Eqs. (8) and (9) with respect to \( q \), yields

\[
A_r : = X^\top A_r X = -(Z_r + Z_{r}^\top) - Z_r A_r + A_r Z_r
\tag{23}
\]

\[
B_r : = X^\top B_r X = -(A_r Z_r + A Z_r + A_{r}^\top A_{r}) - Z_r B_r - B_r Z_r
\tag{24}
\]

where according to Eq. (10)

\[
X_{r,s} : = X(Z_r, Z_s)
\tag{25}
\]

Again using Eq. (8) to eliminate \( Z_r \) and \( Z_{r}^\top \) leads to

\[
A_r = -(Z_r + Z_{r}^\top) + [Z_r; A_r] + A_r A_r
\tag{26}
\]

\[
B_r = -(A_{r} Z_r + Z_{r}^\top A_r + A_{r} Z_r + A_r A_r) + [Z_r; A_r] + [Z_r; B_r]
\tag{27}
\]

Inserting \( Z_{r}^\top \) from Eq. (26) into Eq. (27) and using Eq. (13) finally yields

\[
Q_{r,s} : = B_r - A_r A_r - A_r C_r - A_r C_r
\tag{28}
\]

On the right-hand side of this equation the diagonal of the first commutator vanishes at \( q \equiv q' \) and the diagonal of the second commutator is zero because \( \Lambda \) is diagonal. Only the off-diagonal elements of \( Z^{11}_{r} \), \( r = 1, \ldots, m \), are unknown. But this block does not occur in the diagonal of the last commutator on the right-hand side of Eq. (28). That can be verified by recalling that with reference to Eq. (18) the first diagonal block of \( C \) is diagonal, i.e.,

\[
C^{11}_r = \Lambda^1_r
\tag{29}
\]

with the first partition \( \Lambda^1_r \) of the eigenvalue derivatives given by

\[
\Lambda^1_r = \begin{bmatrix}
\Lambda^1_r & 0 \\
0 & \Lambda^2_r
\end{bmatrix}
\tag{30}
\]

Only the off-diagonal blocks of the last commutator depend on \( Z^{11}_{r} \). Thus, the second partial derivatives of the eigenvalues can be calculated from the diagonal of Eq. (28), yielding

\[
\Lambda_{r,s} = (Q_{r,s})_{diag}
\tag{31}
\]

\[
\left[
\begin{array}{c}
0 \\
(Z^{12}_r C^{11}_r - C^{11}_r Z^{12}_r)_{diag}
\end{array}
\right]
\tag{32}
\]

It remains to derive conditions that enable the calculation of the unknown off-diagonal of \( Z^{11}_{r} \). Because the first partition of Eq. (28) does not depend on \( Z^{11}_{r} \), it can be used to calculate \( (Z^{11}_{r})_{off} \), i.e.,

\[
(Q^{11}_{r,s})_{off} = \left[ ((Z^{11}_r)_{off} \Lambda^1_r) + ((Z^{11}_r)_{off} \Lambda^1_r) + ((Z^{12}_r C^{11}_r - C^{11}_r Z^{12}_r)_{off}) \right]
\tag{33}
\]

For distinct eigenvalue derivatives, Eq. (32) allows the calculation of \( (Z^{11}_r)_{off} \) using, for instance, the diagonal of \( S \), which yields with reference to Eq. (21) for \( i \neq k \)

\[
Z^{11}_{i,k} = \frac{1}{2(\Lambda^1_{ii} - \Lambda^1_{kk})} [Q^{11}_{i,k} - 2 C^{12}_r (\Gamma - \ell I_{N - s})^{-1} C^{11}_r]_{ik}
\tag{34}
\]

Of course, the consistency of Eq. (32) with this result has to be checked for the remaining equations resulting from \( r \neq s \) because the coupling of the sensitivities of the eigenvalues and eigenvectors means that their existence is coupled, too. In the following section, conditions for the existence of the first partial derivatives will be inferred and their continuity is investigated.

### III. Existence of Eigenvalue and Eigenvector Sensitivities

#### A. Two Theorems on the Existence of the Partial Derivatives of Repeated Eigenvalues

As already pointed out, the eigenvectors related to the multiple eigenvalues, in general, are not unique. In some cases it is possible to choose the orthonormal matrix \( \Theta \) in such a way that it diagonalizes the matrices \( C^1 \) for all \( r = 1, \ldots, m \). Thus, the existence of the partial derivatives of the eigenvalues depends on the choice of the parameterization.

**Theorem 1**: A necessary and sufficient condition of the existence of the partial derivatives of \( \Lambda \) at \( q' \) is that

\[
[C^1_s; C^1_r] = 0 \quad \forall r, s = 1, \ldots, m
\tag{35}
\]

From basic algebra (for further reading see, for instance, Refs. 14–16) it is clear that for two matrices having the same orthonormal eigenvectors the commutator vanishes. On the other hand, if the commutator of two matrices vanishes they have the same diagonalizing matrix. If both matrices are symmetric [see Eqs. (11) and (12)], the diagonalizing matrix is orthonormal. The fact that there exist a maximum \( n \) linearly independent commuting \( n \times n \) matrices leads to the following theorem.

**Theorem 2**: Denoting \( cs(C^1_s) \) the \( n^2 \)-dimensional vector containing the sequence of the column vectors of \( C^1_s \), then a necessary condition for the existence of the first partial derivative of the eigenvalues at \( q \equiv q' \) is

\[
\text{rank}(cs(C^1_s), \ldots, cs(C^1_m)) \leq n
\tag{36}
\]

The conditions of Theorems 1 and 2 enable a given parameterization to be tested. If either theorem is violated, further computations
are in vain because the derivatives will not exist at \( q = q' \). Moreover, if the parameterization is permissible in the specified sense, at \( q = q' \) the eigenvalues and eigenvectors and their first partial derivatives are continuous if \( A(q) \) and \( B(q) \) and their first derivatives are continuous. As a consequence of the continuity of the first partial derivatives, the order of differentiation within the second partial derivatives is arbitrary. Interchanging \( r \rightarrow s \), for example, in Eq. (28) and using \( B_r = B_s \), as well as \( A_r = A_s \), lead to

\[
0 = Q_s - Q_r
\]

\[
= [Z_s - Z_r; A] + [Z_s; A_r] - [Z_s; A_r] + [Z_r; A] - [Z_r; A_r] + [Z_r; A] 
\]

\[
+ [Z_r; A_r] - [Z_r; A_s] + [Z_r; A_s] + A_{r,s} - A_{r,s} - A_{r,s} - A_{r,s} 
\]

\[
= [Z_s - Z_r; A] + A_{r,s} - A_{r,s} - [Z_r; A] + [Z_r; A] 
\]

\[
(36)
\]

where for \( C \) and \( C_s \) the corresponding left-hand sides of Eq. (13) have been inserted. Interchanging the arguments in the last commutator of Eq. (36) and using the Bianci identity

\[
[A; [B; C]] + [B; [C; A]] + [C; [A; B]] = 0 
\]

(37)
to rewrite the last two commutators in Eq. (36) yields

\[
[Z_s - Z_r; A] + A_{r,s} - A_{r,s} 
\]

(38)

In general, this equation holds if

\[
\Lambda_r = \Lambda_s 
\]

(39)

\[
Z_r = Z_s 
\]

(40)

The first equation expresses the arbitrariness of the order of differentiation of the eigenvalues and is equivalent with the continuity of their first partial derivative. The second equation is, with reference to Eq. (25), equivalent with the continuity of the first partial derivatives of the eigenvectors. Before corresponding conditions for the existence of the partial derivative of the eigenvectors will be derived the following academical example of Seyranian et al.\(^5\) (see also Ref. 7, p. 158) shows that not for any parameterization do the partial derivatives of the eigenvalues exist.

For the example with a linear parameterization, let \( A = I_1 \), and for \( q = (q_1, q_2) \), let

\[
B(q) := I_2 + \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} q_1 + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} q_2 
\]

(41)

which leads to the repeated eigenvalue \( 1 = \lambda_1 = \lambda_2 \) at \( q = 0 \). Theorem 2 does hold in this case because

\[
\text{rank} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} = 2 = n 
\]

(42)

Checking Theorem 1 by using \( C = B = B_r \) leads to

\[
[B; B_r] = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \neq 0 
\]

(43)

Thus, the given parameterization (41) is not permissible. Indeed, the calculation of the eigenvalues leads to

\[
\dot{\lambda}(q) = 1 + \frac{1}{2} q_1 \pm w(q) 
\]

(44)

where

\[
w(q) := \sqrt{(q_1/2)^2 + q_2^2} 
\]

(45)

The first partial derivatives are

\[
\frac{\partial \lambda}{\partial q} = \frac{1}{2} (0 \pm \frac{1}{4w} q_1) 
\]

(46)

and the second partial derivatives turn out to be

\[
\frac{\partial^2 \lambda}{\partial q^2} q = \pm \frac{1}{4w} q_1 q_2 - \frac{q_2^2}{q_1^2} 
\]

(47)

Obviously the limits of the derivatives as \( q \rightarrow 0 \) do not exist.

B. Two Theorems on the Existence of the Partial Derivatives of Eigenvectors

Although it is popular (see, for instance, Ref. 5 or 9) to use the diagonal \( r = s \) of Eq. (32) only to calculate \( (Z_s^{11})_{\text{off}} \), it may be that this solution violates the remaining equations resulting from \( r \neq s \). Thus, it remains to investigate Eq. (32) and to assure its consistency. It is sufficient to show that the system of equations

\[
(G_{r,s})_{\text{off}} = \begin{pmatrix} (Z_{r}^{11})_{\text{off}}^{A} \\ (Z_{s}^{11})_{\text{off}}^{A} \end{pmatrix} 
\]

\[
= \begin{pmatrix} (Z_{r}^{11})_{\text{off}}^{A} - (Z_{r}^{11})_{\text{off}}^{A} + (Z_{s}^{11})_{\text{off}}^{A} - (Z_{s}^{11})_{\text{off}}^{A} \end{pmatrix} 
\]

\[
\forall r, s = 1, \ldots, m 
\]

(48)

has unique solutions \( (Z_{r}^{11})_{\text{off}}^{A} \), \( r = 1, \ldots, m \), for given diagonal matrices \( A_{r,s} \), having distinct elements, i.e., \( A_{r,s} \neq A_{s,r}, i \neq k \), for \( i, k \in \{1, \ldots, n\} \) for each \( r, s = 1, \ldots, m \). The matrix \( G_{r,s} \) in Eq. (48) represents all known terms of Eq. (32), i.e.,

\[
(G_{r,s})_{\text{off}} := (Q_{r,s})_{\text{off}} - (Z_{r}^{12} C_{r}^{2} + Z_{s}^{12} C_{s}^{2})_{\text{off}} 
\]

(49)

where \( Z_{r}^{12} \) and \( Z_{s}^{12} \) are known from Eqs. (20) and (21). Because the diagonal part of Eq. (48) is identically zero it represents a coupled system of \( m(n-1) \) equations. Writing one equation for the element \( g_{i,s,s} \) of \( G_{s,s} \) in row \( i \) and column \( k \), where \( i \neq k \) for \( i, k \in \{1, \ldots, n\} \) and denoting the corresponding element of \( Z_{s}^{11} \) by \( z_{s,k} \), Eq. (48) reads

\[
g_{i,s,s} = (A_{s,s} - A_{s,s}) z_{s,k} + (\lambda_{s,s} - \lambda_{s,s}) z_{s,k} 
\]

(50)

Expanding this equation to \( m^2 \) equations resulting from \( r, s = 1, \ldots, m \), leads to

\[
\begin{pmatrix} g_{1,1} & \cdots & g_{1,m} \\ \vdots & \ddots & \vdots \\ g_{m,1} & \cdots & g_{m,m} \end{pmatrix} = \begin{pmatrix} z_{1,1} \\ \vdots \end{pmatrix} + \begin{pmatrix} (A_{s,s} - A_{s,s}) \cdots (A_{s,m} - A_{s,m}) + b_{s} h_{s}^{T} \end{pmatrix} = \begin{pmatrix} z_{1,1} \cdots z_{m,m} \end{pmatrix} + b_{s} h_{s}^{T} 
\]

(51)

Because the right-hand side of this equation is symmetric, the left-hand side has to be symmetric, too. Moreover, the ranks of both sides have to be the same, which is equivalent to the condition

\[
\text{rank}(g_{s}) = \text{rank}(z_{s} h_{s}^{T} + b_{s} z_{s}^{T}) 
\]

(52)

Of course, Eq. (51) has a unique solution only if

\[
N_{s} g_{s} N_{s} = 0 
\]

(53)

where the symmetric matrix \( N_{s} \) is the orthogonal projector (see, for instance, Ref. 17) into the \((m-1)\)-dimensional orthogonal complement of the one-dimensional subspace spanned by \( h_{s} \), i.e.,

\[
N_{s} := I_{m} - P_{s} 
\]

(54)

where

\[
P_{s} := \frac{h_{s} h_{s}^{T}}{||h_{s}||^{2}} 
\]

(55)
is the orthogonal projector into the subspace spanned by \( h_k \in \mathbb{R}^m \). To see this, one can multiply Eq. (51) by \( h_k \), yielding

\[
G_k h_k = z_k h_k^2 + h_k \varepsilon^T_k h_k
\]

\[
= (\| h_k \|^2 I_m + h_k h_k^T) z_k
\]

For \( h_k \neq 0 \), the matrix on the right-hand side of Eq. (56) is non-singular. The unique solution of Eq. (56) is given by

\[
z_{ik} = \frac{1}{\| h_k \|^2}(I_m - \frac{1}{2} h_k h_k^T) G_k h_k
\]

Of course, in general this solution is not a solution of Eq. (51). Inserting \( z_{ik} \) from Eq. (57) into Eq. (51) yields

\[
G_k (I_m - \frac{1}{2} P_k) G_k + P_k G_k (I_m - \frac{1}{2} P_k) = G_k P_k + P_k G_k - P_k G_k P_k
\]

The latter equation is equivalent to

\[
0 = G_k - G_k P_k - P_k G_k + P_k G_k P_k = N_k G_k N_k
\]

Summarizing what has been said in this section leads to the following.

**Theorem 3:** Necessary conditions for the existence of a unique solution of Eq. (51) are

\[
G_k \equiv G_k^{\top} \geq 2
\]

\[
\text{rank}(G_k) \leq 2
\]

Moreover, Eq. (32) is consistent with the result given in Eq. (33) if and only if

\[
N_k G_k N_k = 0 \quad \forall i \neq k, \quad i, k \in \{1, \ldots, n\}
\]

where \( G_k \) and \( N_k \) are defined in Eqs. (51–55).

Of course, this leads to restrictions on the given parameterization. Inserting the expressions given in Eqs. (20) and (21) into Eq. (49), and using Eq. (28) with partitions for \( A_r, B_r, \) and \( B_r \), corresponding to that of Z and \( C \) defined in Eqs. (19) and (17), respectively, yields

\[
G_r \equiv B_r^{11} - \varepsilon A_r^{11} - A_r^{11} A_r^{11} - A_r^{11} \varepsilon A_r + C_r^{11} D C_r^{11} - C_r^{11} D C_r^{11}
\]

where \( D = (I - \varepsilon T N - \varepsilon)^{-1} \). Considering the element in row \( i \) and column \( k \) of Eq. (63) leads to

\[
\varepsilon_{ikrs} = b_{ikrs} - a_{ikrs} \varepsilon - a_{ikrs} b_{ikrs} - a_{ikrs} C_r D \varepsilon - C_r D \varepsilon_{ikrs}
\]

where \( b_{ikrs}, a_{ikrs}, \) and \( a_{ikrs} \) denoting the elements in the \( i \)th row and in the \( k \)th column of the matrices \( B_r^{11}, A_r^{11}, \) and \( A_r^{11} \), respectively, and \( \varepsilon \) is the \( i \)th column vector of \( C_r^{11} \). Writing Eq. (64) for all rows \( r = 1, \ldots, m \) and for all columns \( s = 1, \ldots, m \) yields

\[
G_k = \begin{bmatrix}
\mathbf{b}_{ik11} & \cdots & \mathbf{b}_{iklm} \\
\vdots & \ddots & \vdots \\
\mathbf{b}_{ikml} & \cdots & \mathbf{b}_{ikmm}
\end{bmatrix} - \varepsilon
\begin{bmatrix}
\mathbf{a}_{ik11} & \cdots & \mathbf{a}_{iklm} \\
\vdots & \ddots & \vdots \\
\mathbf{a}_{ikml} & \cdots & \mathbf{a}_{ikmm}
\end{bmatrix}
\]

\[
= : A_k
\]

\[
\mathbf{A}_k = \begin{bmatrix}
\mathbf{a}_{ik11} & \cdots & \mathbf{a}_{iklm} \\
\vdots & \ddots & \vdots \\
\mathbf{a}_{ikml} & \cdots & \mathbf{a}_{ikmm}
\end{bmatrix}
\]

\[
= : \mathbf{a}_k
\]

\[
\mathbf{f}_k = \begin{bmatrix}
\mathbf{a}_{ik11} & \cdots & \mathbf{a}_{iklm} \\
\vdots & \ddots & \vdots \\
\mathbf{a}_{ikml} & \cdots & \mathbf{a}_{ikmm}
\end{bmatrix} \mathbf{e}_k - \mathbf{a}_k \mathbf{f}_k^T - \mathbf{f}_k \mathbf{a}_k^T - \mathbf{f}_k \mathbf{DC}_1 - \mathbf{C}_1^T \mathbf{DC}_k
\]

Obviously, the matrices \( \mathbf{G}_k \) are symmetric for all \( i \neq k \in \{1, \ldots, n\} \). Thus, the first of the necessary conditions of Theorem 3 does hold. Although Eq. (65) is related to Eq. (53) via \( h_k = f_k - f \), neither the rank condition nor the sufficient condition formulated in Theorem 3 does necessarily hold for any parameterization. Even in the case of a linear parameterization of the matrix \( B \), Eq. (65) becomes

\[
G_k \equiv -C_1^T \mathbf{DC}_k - C_k^T \mathbf{DC}_k
\]

which may be of rank \( m > 2 \) for some \( i \neq k, i, k \in \{1, \ldots, n\} \). The sufficient condition formulated in Theorem 3 should be checked numerically rather than analytically. The analytical expression of how the projector \( N_k \) acts on \( G_k \) is rather difficult. A better insight can be achieved by translating the effect of \( N_k \) on \( G_k \) rather than on \( G_k \). Of course, in the light of Eq. (53) the effect of \( N_k \) on \( G_k \) is related to the effect of the projector \( P_k \):

\[
0 = N_k G_k N_k \equiv G_k - G_k P_k - P_k G_k + P_k G_k P_k
\]

Considering only the element in row \( r \) and in column \( s \) of Eq. (67) and expanding this expression for all \( i \neq k \in \{1, \ldots, n\} \) leads to

\[
0 = (G_{rs})_{al} - H \odot ([S_r; A_r^1] + [S_r; A_r^1]) + H \odot ([T^1; A_r^1]; A_r^1)
\]

where \( \odot \) denotes the Hadamard product, which is the component-wise product of two matrices having the same size, \( l_l \) and

\[
S_r := \sum_{i=1}^m [G_{ir}; A_r^1]
\]

\[
T := \sum_{i=1}^m [A_{ir}; S_r]
\]

\[
(H_{rs})_{al} := \left[ \sum_{i=1}^m (A_{is} - \lambda_{is})^2 \right]^{-1}
\]

This leads to the formulation of the following theorem.

**Theorem 4:** Equation (32) is consistent with the result given in Eq. (33) if and only if

\[
(G_{rs})_{al} = H \oplus ([S_r; A_r^1] + [S_r; A_r^1]) + H \oplus ([T^1; A_r^1]; A_r^1)
\]

\forall r, s, \ldots, m \), \( S, T, \) and \( H \) are defined in Eqs. (69–71). Of course, Theorem 3 is equivalent to Theorem 4 but for a numerical check either one of them may be used. Before a three-dimensional example is presented (in the next section), the following academical example of a nonlinear parameterization shows that, though the partial derivatives of the eigenvalues exist, the partial derivatives of the eigenvectors do not exist in general. For the example with a nonlinear parameterization, let \( A = L_2 \) and define \( B(q) \) with \( q \in \mathbb{R}^2 \) by

\[
B(q) := f_1 + \begin{bmatrix} 0 & 2 \\ 1 & -q_1 \\ 0 & -q_2 \end{bmatrix} q_1 + \begin{bmatrix} 0 & -q_1 \\ 1 & -q_2 \end{bmatrix} q_2
\]

which leads at \( q = 0 \) to eigenvalues \( \lambda_1 = \lambda_2 = 1 \). A brief calculation shows

\[
C_1 = B_1 = 0 \quad \Rightarrow \quad \lambda_1 = 0 \quad \Rightarrow \quad \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}
\]

\[
C_2 = B_2 = 0 \quad \Rightarrow \quad \lambda_2 = 0 \quad \Rightarrow \quad \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}
\]

Thus, due to Theorems 1 and 2, the partial derivatives of the eigenvalues exist. To check Theorems 3 or 4, one has to calculate the following quantities:

\[
B_1 = 0 \quad \Rightarrow \quad \lambda_1 = 0 \quad \Rightarrow \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}
\]

\[
B_2 = 0 \quad \Rightarrow \quad \lambda_2 = 0 \quad \Rightarrow \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}
\]

\[
B_2 = 0 \quad \Rightarrow \quad \lambda_2 = 0 \quad \Rightarrow \quad \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}
\]
A brief calculation shows that $G_k = 0$ for all but one pair of indices $(i, k) = (1, 2)$, yielding
\[ G_{12} = -2 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \]

On the other hand,
\[ h_{12} = \left( \frac{\lambda_{12} - \lambda_{22}}{\lambda_{22} - \lambda_{11}} \right) \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \]

and, therefore,
\[ N_{12} = 1 - \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \]

Finally, to check Theorem 3, one has to calculate
\[ N_{12}G_{12}N_{12} = -\frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \] (81)

Thus, the partial derivatives of the eigenvectors do not exist. A direct calculation by using the diagonal [see Eq. (33)] without checking the consistency would lead to the incorrect results
\[ Z_1 = 0 \] (84)
\[ Z_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \] (85)

In the next section, a three-dimensional example is investigated concerning permissible linear parameterizations.

IV. Example

To demonstrate the application of the theorems, the spring-mass model depicted in Fig. 1, which corresponds to the example presented by Friswell, is used. The mass matrix
\[ A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix} \] (86)

is assumed to be constant and the stiffness matrix is parameterized by
\[ B(q) = \begin{bmatrix} e_1 \otimes e_1 q_1 + 8 e_1 \otimes e_2 q_2 + e_1 \otimes e_1 q_3 \\ e_2 \otimes e_1 q_1 + 8 e_2 \otimes e_2 q_2 + e_2 \otimes e_1 q_3 \\ 2(e_2 - e_3)(e_2 - e_3)^T q_4 + 2(e_2 - e_3)(e_2 - e_3)^T q_5 \end{bmatrix} \]
\[ =: B_1 =: B_2 =: B_3 \]
\[ + \begin{bmatrix} (e_1 - e_2)(e_1 - e_2)^T q_1 + (e_1 - e_2)(e_1 - e_2)^T q_3 \end{bmatrix} \]
\[ =: B_4 \]
\[ + \begin{bmatrix} (e_1 - e_2)(e_1 - e_2)^T q_6 \end{bmatrix} \]
\[ =: B_6 \] (87)

which corresponds (see Fig. 1) to
\[ (k_1, k_2, k_3, k_4, k_5) = (q_1, 8q_2, q_3, 2q_4, 2q_5, q_6) \] (88)

Fig. 1 Simple discrete three-degree-of-freedom model.

At $q = (0, 1, 0, 1, 1)^T$, a repeated eigenvalue $\lambda_4 = \lambda_5 = 4$ occurs. The associated eigenvectors are
\[ x_4 = a \cdot (1, 0, 0, 1, 1)^T, \quad x_5 = b \cdot (1, 0, 0, 1, 1)^T \] (89)

The remaining eigenvalue is $\lambda_3 = 1$ with the eigenvector
\[ x_3 = c \cdot (2, 1, 2)^T \] (90)

The normalization constants are $a := 1 / \sqrt{2}$, $b := 1 / \sqrt{8}$, and $c := 1 / \sqrt{2}$. Because $A$ does not depend on the parameters $C \in \mathbb{R}^4$, and because all submatrices $B_r$ are symmetric generated by one vector, each $C_r$ also is symmetric and generated by one vector only. The generating vectors are listed in Table 1. For example,
\[ C_1 \equiv (a, b, 2c)^T (a, b, 2c) \] (91)

To check Theorems 1 and 2, the matrices $C^{(1)} \in \mathbb{R}^{2 \times 2}$ have to be calculated for all $r = 1, \ldots, 6$. They are generated by the twodimensional vectors containing the first two components of the generators listed in Table 1, for instance,
\[ C_1^{(1)} = (a, b)^T (a, b) \] (92)

Theorem 2 yields
\[ \text{rank} \begin{bmatrix} a^2 & 0 & a^2 & 1 & 1 & 2 \\ ab & -ab & 4ab & -4ab & 0 & 0 \\ ab & 0 & -ab & 4ab & -4ab & 0 \\ b^2 & 8b^2 & b^2 & 8b^2 & 0 & 0 \end{bmatrix} = 4 \geq 2 = n \] (93)

Thus, the partial derivatives of the eigenvalues will not exist for the complete parameterization as defined by Eq. (87). To answer the question of a permissible parameterization Theorem 1 has to be checked. This requires the calculation of 15 commutators. The resulting matrices are skew symmetric. The corresponding upper off-diagonal elements are listed in Table 2 for all $r < s$. Of the six matrices only the two corresponding to parameters $q_2$ and $q_4$ commute. Thus, an orthogonal matrix $\Theta \in \mathbb{R}^{2 \times 2}$ will only exist for the parameterization
\[ B(q_1, q_2) = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} + 8 e_1 \otimes q_1 + (e_1 - e_3)(e_1 - e_3)^T q_2 \]
\[ 0 \leq q_2 \leq 1 \] (94)

where the parameters have been changed according to $(q_2, q_4) \rightarrow (q_1, q_2)$. Using this new parameterization to calculate the first derivative of the eigenvalues from Eq. (13) at $q_1 = q_2 = 0$ it turned out
that $\Theta^1$ is already diagonal, i.e., $\Theta = I_2$. The eigenvalue derivatives are

$$\frac{\partial \Lambda}{\partial q_1} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ \frac{\partial \Lambda}{\partial q_2} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

(95) (96)

Because the inertia matrix $A$ does not depend on the parameters from Eq. (8), it follows that $Z_r$ is skew symmetric and, thus, $(Z_r)_i = 0$ for $r = 1, 2$ and for $i = 1, \ldots, N$. Equations (20) and (21) lead to

$$(Z_r)_{11} = 0 \quad \forall r, 1, 2 \\ (Z_1)_{23} = \frac{2}{3} h_{12}$$

(97) (98)

$$(Z_r)_{23} = 0$$

(99)

To determine the off-diagonal part of $(Z_r)_{12}$, Eq. (32) can be used. For the example considered here, $Q_2$, and $\Theta$ are zero and Eq. (32) is consistent with the result

$$(Z_2)_{12} = 0 \quad \forall r, 1, 2$$

(100)

The result $X_2 = XZ_2 = 0$ means that a first-order approximation change in the stiffness $K_2$ does not affect the eigenvectors.

Using Eq. (31) to calculate the second derivatives of the eigenvalues this particular example leads to

$$\frac{\partial^2 \Lambda}{\partial q_1^2} = \begin{bmatrix} 16 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

(102)

All other derivatives turned out to be zero.

To compare these results with those presented by Friswell, the derivatives of the eigenvalues and eigenvectors with respect to parameter $\theta$ will be calculated, where

$$q_1 = \theta - \frac{h}{2}$$

(103)

$$q_2 = \theta$$

(104)

At $q_1 = q_2 = 1 \Rightarrow \theta = 1$ the operators of the partial differentiation are related via

$$\frac{\partial}{\partial \theta} = \frac{3}{2} \frac{\partial}{\partial q_1} + \frac{\partial}{\partial q_2}$$

(105)

$$\frac{\partial^2}{\partial \theta^2} = \frac{9}{4} \frac{\partial^2}{\partial q_1^2} + \frac{3}{2} \frac{\partial^2}{\partial q_1 \partial q_2} + \frac{\partial^2}{\partial q_2^2}$$

(106)

Using the first operator equation together with Eqs. (95) and (96) leads to

$$\Lambda_{\theta} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(107)

From Eq. (102) the second operator equation yields

$$\Lambda_{\theta, \theta} = \frac{4}{3} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

(108)

To calculate the derivatives of the eigenvectors with respect to $\theta$ at $\theta = 1$, the operator in Eq. (105) leads to

$$X_{\theta, \theta} = \frac{3}{2} XZ_{1, \theta} = \begin{bmatrix} 0 & -2b/3 & 2c/3 \\ 0 & -b/3 & -2c/3 \\ 0 & -2b/3 & 2c/3 \end{bmatrix}$$

(109)

The results given by Friswell match the results in Eqs. (107–109).

V. Conclusions

The existence of the derivatives of eigenvalues is investigated. In the case of multiple eigenvalues, the partial derivatives do not exist for any parameterization. Two conditions are deduced, which enable a given parameterization to be tested to determine its permis-

sibility. For any continuous permissible parameterization, the partial derivatives of the eigenvalues and eigenvectors are continuous, too.

The application of the theorems presented has been demonstrated by examples. In preparation is a method to calculate the partial derivatives of repeated eigenvalues independent of the existence of eigenvectors.

Acknowledgment

M. I. Friswell gratefully acknowledges the support of the Engineering and Physical Science Research Council through the award of an Advanced Fellowship.

References

18. A. D. Belegundu

Associate Editor