Eigenvector sensitivity when tracking modes with repeated eigenvalues

D. Ruiz, J.C. Bellido*, A. Donoso

Departamento de Matemáticas, ETSII Universidad de Castilla-La Mancha, 13071 Ciudad Real, Spain

Received 24 March 2017; received in revised form 2 June 2017; accepted 26 July 2017
Available online 12 August 2017

Abstract

While eigenvalue optimization problems have been widely studied during the last three decades, particularly for structural and topology optimization problems, there are very few examples of problems involving mode shapes either in the cost or the constraints. In this paper, we propose a general framework for computing eigenvector sensitivity whenever tracking specific mode shapes selected beforehand. Of course, the approach is valid for either non-repeated or repeated eigenvalues, but here the emphasis is placed on the multiple eigenvalues case. Both mathematical validity of the algorithm and numerical corroboration through several examples are included.

Keywords: Eigenvector derivative; Repeated eigenvalues; Mode shape derivative; Eigenmode optimization; Eigenmode selection; Modal assurance criterion

1. Introduction

Eigenvalue optimization problems appear very frequently when dealing with structural stability and vibration analysis, and nowadays it is a very well understood topic in topology optimization [1]. Eigenvalue sensitivity analysis for this kind of problems is a classic issue dating back to the sixties. For eigenvalue sensitivity calculation there are two different cases: simple, non-repeated, or multiple, repeated, eigenvalues, being this last case much more difficult and subtle than the former one, since multiple eigenvalues are not differentiable. There are many references where this has been addressed, and among those we cite [2,3]. Sensitivity analysis has been applied for optimization of eigenvalues for free vibrations together with proper model formulations like the bound formulation (see [1,4] and the references therein).

Optimization problems involving eigenmodes, i.e. such that either cost or constraints depends on eigenmodes are much more scarce in the literature. However, in the last years several models requiring to optimize eigenmodes or constraining functionals depending on eigenmodes have been studied. In [5], a mode shape corresponding to a simple
eigenvalue of a fiber laser package is designed in order to minimize the elongation of the fiber under dynamic excitation. In [6], a multi-objective function is formulated in order to find optimal configurations that simultaneously satisfy (simple) eigenfrequency, eigenmode, and stiffness requirements at certain points of a vibrating structure. A similar problem is treated in [7]. The novelty there is to include the electromechanical coupling coefficient in the objective function so that the energy conversion is maximized for a specific mode. In [8], eigenmodes appear in the constraints only. One of the objectives of that work is to determine the material distribution of a structure that maximizes the fundamental frequency and at the same time synthesizes the first two modes. Recently, in [9] an optimization problem in which both cost and constraints depend on eigenmodes has been studied, and furthermore, all the pathologies appearing when dealing with eigenproblems are present: spurious modes, mode switching and multiple eigenfrequencies. The objective in that paper is the optimal design of piezoelectric modal sensor/actuators, in which we simultaneously designed the ground structure and polarization profiles of the piezoelectric transducers (see the recent survey [10]).

One of the main difficulties when dealing with eigenmode optimization is the sensitivity analysis. Eigenvector sensitivity for structural analysis problems has been addressed in [11–19]. The first issue is about differentiability, while eigenmodes corresponding to non-repeated eigenvalues are differentiable, eigenmodes corresponding to repeated ones are not even continuous. This is due to the fact that there are infinite eigenvector basis of the eigenspace associated to a multiple eigenvalue, so that this ambiguity spoils continuity and therefore differentiability. However, differentiability can be guaranteed for a concrete eigenvector basis, the so-called adjacent basis, and for it explicit derivative computations can be performed. Further to this, the eigenvector derivative computation leads to solve degenerate linear systems for which specific methods have been developed [11,13,14,16,17].

Another crucial point is eigenmode selection. When a functional involved in the optimization problem, either the cost or a constraint, depends on a mode shape it must be very clear on which specific mode shape is such a dependence, since mode shapes change and mode switching (change of order of eigenvalues in the spectrum) happens during the optimization process when updating variables. Therefore cost and constraints must clearly depend on selected mode shapes. In [9], a procedure for selecting the closest mode shapes to given reference mode shapes was developed. This procedure follows the idea and spirit of the well-known modal assurance criterion [20] but including the possibility of repeated eigenvalues appearance. The difficulty now is that once eigenmodes have been selected, the optimization algorithm requires derivatives for those specific eigenmodes. This is not an issue for eigenmodes associated with non-repeated eigenvalues since those are differentiable, but if they are associated with repeated ones, then available methods only provide derivatives for the adjacent basis of eigenvectors. For the specific situation considered by the authors in [9], a heuristic method working out for obtaining these derivatives was developed. The aim of this paper is to deepen those ideas whenever tracking specific mode shapes with three objectives in mind: establish a method of eigenmode selection suitable for a wide range of situations of eigenmode optimization models, give a general algorithm for computing derivatives of selected eigenmodes valid for the case of repeated and non-repeated eigenvalues, and finally, analyze the mathematical foundations of the algorithm in order to show its validity. These three objectives have been reached in this paper.

The paper is organized as follows. In Section 2 we review the main algorithms for eigenvector derivatives computation and comment on differentiability, clarifying some misleading ideas and claims present in the literature. In Section 3 we present a method for eigenmode selection and for the calculation of derivatives of selected eigenmodes. The algorithm validity is numerically corroborated with two structural analysis examples. Finally, in Section 4 we consider an example of an eigenmode optimization problem inspired by [9] in order to numerically validate the previous algorithm in a real optimization problem.

2. Review of current methods for eigenvectors sensitivity analysis

Given $K, M \in \mathbb{R}^{n \times n}$ real, symmetric and positive definite matrices. We are interested in the eigenproblem

$$(K - \lambda M)U = 0. \quad (1)$$

In structural dynamics, $K$ stands for the stiffness matrix, $M$ the mass matrix, $\lambda$ the eigenvalues or eigenfrequencies, i.e. the square of the natural frequencies, and $U$ stands for the mode shape, or eigenvector, corresponding to the eigenfrequency $\lambda$.

It is well known that under these conditions all the eigenfrequencies are real, $\lambda_1 \leq \cdots \leq \lambda_n \in \mathbb{R}$, and there exists a $M$-orthonormal basis of $\mathbb{R}^n$ made of eigenvectors $U_1, \ldots, U_n$, i.e.

$$(K - \lambda_j M)U_j = 0 \quad (2)$$
\[ U_j^T M U_j = \delta_{ij}, \quad i, j = 1, \ldots, n, \] (3)

where \( \delta_{ij} \) is the Kronecker’s delta.

We assume that both mass and stiffness matrices depend smoothly on a vector of real variables \( x = (x_1, \ldots, x_L) \in \mathbb{R}^L \), that usually represents a physical magnitude for the different elements of a finite element mesh, as Young’s modulus, Poisson’s ratio, thickness, moment of inertia, etc. We are interested in the computation of the eigenproblem sensitivity with respect to \( x \). In particular, our main interest is the eigenvectors sensitivity.

Let us begin by doing some formal computations for obtaining the derivatives with respect to the single variable \( x_i \), the \( i \)th component of \( x \). Differentiating (2) with respect to \( x_i \),

\[ \frac{\partial}{\partial x_i} \left( (K - \lambda_j M) U_j \right) = 0, \] (4)

and operating we arrive at

\[ (K - \lambda_j M) \frac{\partial U_j}{\partial x_i} = - \left( \frac{\partial K}{\partial x_i} - \lambda_j \frac{\partial M}{\partial x_i} \right) U_j + \frac{\partial \lambda_j}{\partial x_i} M U_j. \] (5)

Therefore, the derivative of the eigenvector with respect to the variable, \( \frac{\partial U_j}{\partial x_i} \), if it exists, is the solution of a degenerate system (\( \det (K - \lambda_j M) = 0 \)). Furthermore, the right hand side of (5) depends on the eigenfrequency derivative, so that it must exist for the eigenvector derivative to exist. At this point, to go further in this computation we should distinguish between two cases depending on whether the eigenfrequency \( \lambda_j \) is simple or multiple, i.e. \( \lambda_j \) a non-repeated or repeated eigenvalue.

### 2.1. Simple eigenfrequencies

In the case that \( \lambda_j \) is a simple eigenvalue for (2), eigenvectors associated to it are unique up to a multiplicative constant. In this case, it is well known that both eigenvalues and eigenvectors are differentiable with respect to \( x \), provided matrices \( K \) and \( M \) so are with respect to \( x \). See, for instance, [2] or [21] where a nice analytic proof of this fact based on the Implicit Function Theorem is given.

From now on (‘) stands for the derivative with respect to the parameter \( x_i \). For obtaining the derivatives we premultiply both sides of Eq. (5) by \( U_j^T \) and, taking into account (2) and (3), the expression for the partial derivative of \( \lambda_j \) with respect to \( x_i \) is

\[ \lambda_j' = U_j^T \left( K' - \lambda_j M' \right) U_j. \] (6)

Notice that the computation of the sensitivity of the \( j \)th eigenvalue only requires the \( j \)th eigenvector. This is an important fact regarding other methods based on spectral or singular values expansions requiring all the eigenvectors in order to compute any eigenvalue derivative [12,19]. Indeed, this is also an important advantage when working in structural problems where the size of the matrices \( K \) and \( M \) is typically large and then the computation of all eigenvalues and eigenvectors would be computationally prohibitive.

Regarding the computation of the eigenvector derivative, as we mentioned above, we have to solve system (5), and the matter is that this is a degenerate system as the coefficients matrix has rank \( n - 1 \). There are several ways to overcome this difficulty, being the most popular one the Nelson’s method [16]. The idea is to replace a redundant equation in (5) with Eq. (3), or more properly, the derivative of this identity,

\[ U_j^T M U_j' = -\frac{1}{2} U_j^T M' U_j. \] (7)

To this end, the eigenvector derivative with respect to \( x_i \) is decomposed as the sum of a particular solution, called \( V_j \) and the homogeneous solution, \( c_j U_j \), with \( c_j \) a constant to be determined,

\[ U_j' = V_j + c_j U_j. \] (8)

A summary of the process is shown in Table 1.
Table 1
The procedure of Nelson’s method.

1. Compute the eigenvalues derivatives \( \lambda_j' = U_j^T (K' - \lambda_j M') U_j \).
2. Set \( G_j \equiv K - \lambda_j M \) and \( f_j \equiv \lambda_j' M U_j - (K' - \lambda_j M') U_j \).
3. Identify the largest element in \( U_j \), placed in the \( k \)th position.
4. Construct \( \tilde{G}_j \) by zeroing out the \( k \)th row and the \( k \)th column in \( G_j \) and setting the diagonal element equal to 1.
5. Construct \( \tilde{f}_j \) by zeroing out the \( k \)th element of \( f_j \).
6. Compute \( V_j \) by solving the system \( \tilde{f}_j = \tilde{G}_j V_j \).
7. Set \( c_j = -V_j^T M U_j - 0.5 U_j^T M' U_j \), \( c_j \) is obtained from (7).
8. Finally \( U_j' = V_j + c_j U_j \).

Fig. 1. Evolution of eigenvalues with the variable.

Again, a remarkable fact on Nelson’s Method is that computations of the partial derivatives of the \( j \)th eigenvalue and the \( j \)th eigenvector only depend on the own \( j \)th eigenpair itself. Other related methods use a linear combination of eigenvectors, \( \sum_{j=1}^{m} c_j U_j \), as the homogeneous solution of (5).

When we are dealing with optimization problems depending on modes, we typically have to compute derivatives of linear functions of the eigenmodes rather than eigenmodes derivatives themselves. In [22], a procedure for computing derivatives with respect to all the variables at once for this kind of functionals was developed. It is valid for the case of eigenvectors corresponding to simple eigenvalues only, but this is still very remarkable from a computational point of view.

2.2. Multiple eigenfrequencies

The typical situation, but not the only one, in which multiple eigenvalues appear is that two different eigenvalues coalesce at a certain value of the variables vector \( x \), and split out under a small perturbation on any \( x_i \). For instance, for a double eigenvalue (also the most common situation when repeated eigenvalues occur) this can be understood in terms of non-smooth analysis [23]. The subderivative with respect to a single, real variable, is defined as the set of slopes of any tangent straight line to the graph of the eigenvalue, and so in this situation the subderivative is a closed interval whose extreme points are the derivatives of the two coalescent eigenvalues (see Fig. 1). There are well established techniques both for computing multiple eigenvalues derivatives and for dealing with eigenvalues optimization problems (see [1] and the references therein).

For a multiple eigenvalue, \( \lambda_j \), there are infinitely many \( M \)-orthonormal basis of eigenvectors associated to \( \lambda_j \). This fact makes eigenvectors differentiability, and even continuity, to fail. To understand this issue we include, for the interest of the reader, a very simple and clarifying example taken from [18].

Let us consider the eigenvalue problem

\[
\begin{bmatrix}
1 & x \\
x & 1
\end{bmatrix} U = \lambda U,
\] (9)
where $x \in \mathbb{R}$ is the single variable. Clearly, eigenvalues are $1 - x$ and $1 + x$, and an orthonormal basis of eigenvectors written in matrix form is

\[
\begin{bmatrix}
-1 & 1 \\
1 & 1
\end{bmatrix}.
\]  

(10)

For $x = 0$, simple eigenvalues coalesce at a double one, and any basis in $\mathbb{R}^2$ is an eigenvectors basis, spoiling continuity. However, the previous basis would be valid as eigenvector basis when $x = 0$, and for it, continuity would be preserved. This is the idea in computing eigenvector derivatives, that there exists a $\mathbf{M}$-orthonormal basis of eigenvectors which is differentiable with respect to single variables. This basis is sometimes called in the literature adjacent basis, keeping in mind the idea that eigenvalues split out when a perturbation is applied to the variable, and then being both eigenvalues and eigenvectors differentiable.

Let $\lambda$ be an eigenvalue of multiplicity $m$. For computing the right basis of eigenvectors we proceed in the following way. Let us consider an initial matrix $\mathbf{U}$ whose columns, $\mathbf{U}_1, \ldots, \mathbf{U}_m$, are $\mathbf{M}$-orthonormal eigenvectors associated to $\lambda$, i.e.

\[
\mathbf{KU} = \mathbf{MU}\Lambda,
\]  

(11)

and

\[
\mathbf{U}^\top\mathbf{MU} = \mathbf{I}_m,
\]  

(12)

where $\Lambda = \lambda \mathbf{I}_m$, and $\mathbf{I}_m$ is the identity matrix of order $m$. Any other set of $m$ $\mathbf{M}$-orthonormal vectors can be obtained from $\mathbf{U}$ through an orthogonal transformation

\[
\mathbf{Z} = \mathbf{U}\Gamma,
\]  

(13)

with $\Gamma$ an orthogonal matrix (i.e. $\Gamma^\top\Gamma = \mathbf{I}$). Imposing the eigenproblem to $\mathbf{Z}$ we get,

\[
\mathbf{KZ} = \mathbf{MZ}\Lambda,
\]  

(14)

and differentiating and arranging terms we arrive at

\[
(\mathbf{K} - \lambda\mathbf{M})\mathbf{Z}' = -(\mathbf{K}' - \lambda\mathbf{M}')\mathbf{Z} + \mathbf{MZ}'\Lambda'.
\]  

(15)

$\Lambda'$ is the $m \times m$ diagonal matrix whose diagonal elements are the $m$ derivatives of $\lambda$ [3]. Multiplying now by $\mathbf{U}^\top$ and taking into account (11)–(13) we get that $\Gamma'$ and $\Lambda'$ are the solutions of the eigenproblem

\[
\mathbf{D}\Gamma = \Gamma\Lambda',
\]  

(16)

where $\mathbf{D}$ is the $m \times m$ matrix

\[
\mathbf{D} = \mathbf{U}^\top(\mathbf{K}' - \lambda\mathbf{M}')\mathbf{U}.
\]  

(17)

$\mathbf{D}$ is symmetric and therefore $\Gamma$ is orthogonal.

For solving system (15), Nelson’s method is not operative since it is based on the fact that $\text{rank}(\mathbf{K} - \lambda\mathbf{M}) = n - 1$, that is not true for multiple eigenvalues. Several modifications of Nelson’s method have been introduced to overcome this difficulty. The most extended extension of Nelson’s method to multiple eigenvalues is the Dailey’s method [11] (that improved the previous Ojalvo’s method [17]). This method is valid for the case in which eigenvalues derivatives are different, being the most usual situation when dealing with structural or topology optimization problems where, on the one hand, due to symmetry one may have multiple eigenfrequencies in the initial layout that separate when variables are updated, or on the other hand, two eigenvalues coalesce at a certain iteration of the algorithm. In Table 2 the summary of the algorithm for an eigenvalue of multiplicity $m$ is presented. For technical details and the justification of the algorithm readers are referred to [11].

Notice that the main idea is, as in Nelson’s method, to write eigenvectors derivatives as the sum of a solution of a linear transformation of the eigenvectors (solution of the homogeneous system) and a particular solution of (15).

Obviously, construction of matrix $\mathbf{C}$ (Table 2) in the Dailey’s algorithm is not compatible with equal eigenvalues derivatives. As we mentioned above, this is an uncommon situation in structural optimization, but of course it may happen sometimes. In [13], this situation was analyzed, and academic examples were included, as for instance, a circular shaft symmetric in both axes and such that variables are chosen so that symmetry is retained. Friswell solved
Table 2
The procedure of Dailey’s method.

1. Given an initial set of $M$-orthonormal $m$ eigenvectors of $\lambda$ arranged in a matrix $U$, compute, as explained above, the adjacent eigenvectors: $Z = U'V$.
2. Set $G \equiv K - \lambda M$ and $f \equiv A'MZ - (K' - \lambda M')Z$.
3. Identify the $m$ largest elements in matrix $Z$, placed in the $k_1$, $\ldots$, $k_m$ positions.
4. Construct $G$ by zeroing out the $k_l$th rows and the $k_l$th columns, $l = 1, \ldots, m$, in $G$ and setting the diagonal elements equal to 1.
5. Construct $f$ by zeroing out the $k_l$th elements of $f$, $l = 1, \ldots, m$.
6. Compute $V$ by solving the system $f = GV$.
8. Compute $R = Z^T (K' - \lambda M')V + Z^T (M' Z + M V)A' + 0.5Z^T (K'' - \lambda M'')Z$.
9. Let $C$ be the $m \times m$ matrix given by:
   
   \[
   c_{i k} = \begin{cases} 
   r_{i k} (\lambda_i' - \lambda_i') & \text{if } \lambda_i' \neq \lambda_i' \\
   0.5 \delta_{i k} & \text{otherwise}
   \end{cases}
   \]  

(18)
10. Finally, $Z' = V + Z C$.

Table 3
The procedure of Friswell’s method.

1. Compute the matrix $D = U^T (K' - \mu M')U$.
2. Solve the eigenproblem $D\Gamma = \Gamma A'$. Note than in this case matrix $\Gamma$ is not unique, since the eigenvalues of $D$ are repeated.
3. Compute the adjacent eigenvectors, $Z = U\Gamma'$.
4. Set $G \equiv K - \lambda M$ and $f \equiv A'MZ - (K' - \lambda M')Z$.
5. Identify the $m$ largest elements in matrix $Z$, placed in the $k_1$, $\ldots$, $k_m$ positions.
6. Construct $G$ by zeroing out the $k_l$th row and the $k_l$th column, $l = 1, \ldots, m$, in $G$ and setting the diagonal elements equal to 1.
7. Construct $f$ by zeroing out the $k_l$th elements of $f$, $l = 1, \ldots, m$.
8. Compute $V$ by solving the system $f = GV$.
9. Compute $V_\phi = V^T \Gamma\phi$.
10. Compute $\Gamma\phi$ by solving the next eigenproblem: $2U^T (K' - \lambda M' - \lambda M')V_\phi \Gamma\phi + U^T (K'' - \lambda M'' - 2\lambda M')U \Gamma\phi = \Lambda^T \Gamma\phi$. $\Lambda''$ stands for the second eigenvalues derivatives.
11. Since the eigenvalues of the previous eigenproblem are single, $\Gamma\phi$ is unique. If $\Gamma \neq \Gamma\phi$, set $\Gamma = \Gamma\phi$ and $Z = U\Gamma'$.
12. Set $H \equiv K - \lambda M$ and $J_i \equiv -2(K' - \lambda M')V_i + Z C_i - (K'' - \lambda M'' - \lambda_i''M - 2\lambda_i'M)Z_i$.
13. Identify the $m$ largest elements in matrix $Z$, placed in the $k_l$th positions, $l = 1, \ldots, m$.
14. Construct $H$ by zeroing out the $k_l$th row and the $k_l$th column, $l = 1, \ldots, m$, in $H$ and setting the diagonal elements equal to 1.
15. Construct $J$ by zeroing out the $k_l$th element of $J$, $l = 1, \ldots, m$.
16. Compute $W_i$ by solving the system $J_i = H W_i$.
17. Compute the off-diagonal elements of matrix $C$ by solving the next equation: $2Z^T (K' - \lambda M' - \lambda M')W_i + 3Z^T (K'' - \lambda M'' - \lambda_i''M - 2\lambda_i'M)W_i - 2\lambda_i'M V_i + Z C_i + Z^T (K'' - \lambda M'' - 3\lambda_i''M - 3\lambda_i'M)Z_i = 0$, for $j \neq i$.
18. Compute the elements of the diagonal of matrix $C$ as: $C_{i i} = -0.5Z^T M Z_i - V_i^T M Z_i$.
19. Finally, the sensitivity of the eigenvectors is $Z' = V + Z C$.

This situation giving an efficient method for eigenvector derivative calculation in this situation. Table 3 summarizes this algorithm. Notice that in this situation matrix $\Gamma$ for the computation of the adjacent eigenvectors is not unique, since there are equal eigenvalue derivatives in this case, so that a method has to be included in order to find the right set of adjacent eigenvectors for which derivatives can be calculated. For the technical details justifying this method we refer readers to [13].

Nelson, Dailey and Friswell’s are the three essential algorithms for computing eigenvector derivatives. There are however interesting refinements or modifications that could be of interest in certain situations, and, without doubt, worth mentioning. A key point in these algorithms is the solution of the degenerate linear system $GV = f$ (indeed in Friswell’s method this kind of calculation has to be performed twice), that it is solved by zeroing out certain rows and columns of the system matrix. With this procedure one gets a linear system admitting a unique solution, but its matrix might be poorly conditioned. To circumvent this issue, in [15] an iterative algorithm for solving this system is proposed, including an analysis of convergence of the method. In [14] an interesting method for computing solutions to this system is proposed by adding $m$ (recall that $m$ is the eigenvalue multiplicity) new equations to it. Their method
also works out when equal eigenvalue derivatives of any order of derivation occur, what is not considered by the Friswell’s method.

2.3. Remarks on differentiability

The previous methods provide the derivatives of eigenvalues and eigenvectors assuming those to exist. Since literature is somehow confusing regarding whether eigenvalues and eigenvectors are differentiable or not, in the sense that there are misleading comments or claims in several references, specially in those pertaining to the mechanical engineering literature, we would like, to the best of the authors knowledge, to clarify this issue by indicating proper references and known results.

The first point to take into account is that beyond eigenvalues and eigenvectors derivatives with respect to a single real variable, in practical optimization problems more than this is required, we actually need to have differentiability with respect to a vector of variables, meaning that jacobian matrix can be computed (all partial derivatives exist) and any directional derivative coincides with the jacobian matrix acting on the direction.

When an eigenvalue is simple, and stiffness and mass matrices, $K$ and $M$, are continuously differentiable with respect to a single real variable, then differentiability of both eigenvalue and associated eigenvectors holds. There are several references where this is shown, and we have already cited [2,21], and now we add [24]. When the eigenproblem depends on a vector of variables, with $K$ and $M$ continuously differentiable with respect to this vector, since partial derivatives with respect to each single variable are continuous, then differentiability with respect to the variables vector of both the single eigenvalue and its associated eigenvectors holds as well.

The situation for multiple eigenvalues is much more difficult. When considering eigenvalue differentiability with respect to a single variable, [24, Th. 6.8] establishes that if matrices $K$ and $M$ are continuously differentiable then there exist $n$ continuously differentiable functions that represent the (repeated) eigenvalues of the eigenproblem. Differentiability is broken when eigenvalues coalesce if we name then by the usual increasing order, as Fig. 1 graphically shows. Regarding eigenvector differentiability with respect to a single variable when multiple eigenvalue happens, the situation is more subtle. On the one hand, additional smoothness of the eigenproblem with respect to the variable is required. [24, Ex. 5.3] gives an example in which matrices $K$ and $M$ are not analytical with respect to variables and continuity of eigenvectors fails at the point where eigenvalues coalesce, so that analyticity of the eigenproblem with respect to the variable is the minimal hypothesis in order to guarantee eigenvectors differentiability. This is not an issue for structural, or topology, optimization problems where stiffness and mass matrices dependence of variables is usually polynomial, and therefore analytic. Under this analyticity hypothesis, it is shown in [24, Section 6.2], that there exists a $M$-orthonormal basis of $\mathbb{R}^n$ of eigenvectors continuously differentiable (and in fact analytic) with respect to the variable. This is congruent with the previous methods for computing eigenvectors derivatives, where the first step consists in finding the adjacent eigenvectors, which are the eigenvectors that the multiple eigenvalue contributes to the continuously differentiable $M$-orthonormal basis of eigenvectors.

Concerning differentiability with respect to a vector of variables, multiple eigenvalues are not differentiable (in the Fréchet sense) with respect to the full vector of variables, but they are Gâteaux differentiable, i.e. directional derivatives do exist at any direction. Proofs of this fact and formulas for the directional derivatives can be found in [2]. Regarding eigenvectors, the continuously differentiable $M$-orthonormal basis of eigenvectors mentioned in the previous paragraph depends in general on the single variable (in the notation of the previous subsection, the matrix $\Gamma$ depends on the variable with respect to which we are computing derivatives), so that, in general, there does not exist an $M$-orthonormal basis of eigenvectors differentiable with respect to the variables vector [24]. Ref. [13] reports about this issue, including examples in which the basis keeps constant for all the variables, and therefore differentiability is granted. However, for a fixed mode shape one could compute partial derivatives as linear combination of derivatives of the differentiable basis with respect to single variables, and finally to write the jacobian matrix of the fixed mode shape concluding its differentiability since partial derivatives are continuous.

Optimization problems in which cost or constraints depend on eigenvectors usually select them through a criterion permitting to follow a certain desired given mode shape. For instance, a problem in which we would like to find the structure optimizing a cost depending on the structure mode shape closed to a given reference one. Thinking in plate-type structures, that could be a prescribed mode shape of the whole plate (i.e., the whole domain) while optimizing the structure inside the domain. Of course, we would need derivatives for such eigenvectors, that eventually will correspond to double (or in general multiple) eigenvalues. Next section is devoted to this issue.
3. Practical computation of derivatives for selected eigenmodes

In the previous section, classical methods for the computation of derivatives with respect to single variables of eigenvectors associated to multiple eigenvalues have been introduced and commented. Those methods only provide derivatives of a certain eigenvector basis, the so-called adjacent eigenvector basis, associated to the multiple eigenvalues. On the other hand, when we have an optimization problem in which cost, constraints, or both, depend on eigenvectors, gradient-based optimization algorithms require the computation of the eigenvector derivatives, and those eigenvectors may not belong to the adjacent eigenvector basis in the case of repeated eigenfrequencies. In this section we go a step further of this by computing derivatives for fixed, or selected, eigenmodes associated to the multiple eigenvalues but not being an adjacent eigenvector in general. For this situation we propose a procedure that will be corroborated numerically with several examples. Our method provides the derivatives of any selected eigenvector taking advantage of the fact that any eigenvector associated to a multiple eigenvalue can be obtained as a linear expansion of the adjacent eigenvectors of such an eigenvalue. Our procedure admits multiple selection of eigenmodes, actually for the whole eigenvector basis. Our method is not an alternative to the previous methods, but a procedure for their practical implementation in optimization problems.

There are two issues to be solved. First, the derivative computation itself, and second the selection of the eigenmodes. The first one is easily solved. For computing the derivative of a fixed eigenmode we take advantage of the linear structure of the set of eigenmodes associated to a multiple eigenvalue. Let $\lambda$ be an eigenvalue of multiplicity $m$, and $v$ an eigenmode associated to $\lambda$, i.e. $(K - \lambda M)v = 0$, then the derivative of $V$ is computed in the following way:

1. Find the adjacent basis of eigenvectors, $Z^{(1)}, \ldots, Z^{(m)}$;
2. Compute the derivative of this vectors by any of the previous methods, Dailey’s method if possible, for instance;
3. Write the selected eigenmode as a linear expansion of the adjacent eigenvectors, $V = c_1 Z^{(1)} + \cdots + c_m Z^{(m)}$, or in matrix form $V = ZC$,
4. The derivative of the fixed mode is the linear expansion of the derivative of the adjacent eigenmodes with the same coefficients, $V' = c_1 (Z'^{(1)}) + \cdots + c_m (Z'^{(m)}) = Z'C$.

We would like to emphasize that, to the best of the authors knowledge, this algorithm, in spite of it is natural and easy, has not been neither proposed nor used in the literature before.

Regarding the second issue, the eigenvectors selection can be explicitly given by the model, but in general reference eigenvectors must be computed by solving an optimization problem in order to fulfill a proximity criterion in the spirit of the Modal Assurance Criterion (see [20] and the references therein). The idea is to select the closest eigenmodes to given vectors (which typically are eigenmodes of the structure filling the whole design domain). Therefore, if $\tilde{Y}$ is the matrix containing these target vectors (in the same number than the multiplicity of the eigenvalue, this is not strictly necessary but convenient as we point out below), $U$ is an initial eigenvector basis associated to the repeated eigenvalue (given by any numerical algorithm we may use), then the reference (or selected) eigenvectors, i.e. those taking part in our optimization algorithm, are obtained as

$$Y = U \Gamma_0,$$

where the orthogonal matrix $\Gamma_0$ is the unique solution of

$$\begin{align*}
\max_{\Gamma_0, \Gamma_0^T \Gamma_0 = I} \quad & \|\tilde{Y}^T M Y\| \\
\text{s.t.} & \quad Y = U \Gamma_0.
\end{align*}$$

(19)

(20)
Table 4
Derivative of a selected eigenmode.

1. Compute the reference set of eigenvectors, \( \mathbf{Y} \), through the optimization problem (20).
2. Compute the matrix \( \mathbf{D} \) as: \( \mathbf{D} = \mathbf{Y}^T (\mathbf{K}' - \lambda \mathbf{M}') \mathbf{Y} \), and solve the eigenproblem (16).
3. Adjacent eigenvectors are \( \mathbf{Z} = \mathbf{Y} \Gamma \), with \( \Gamma \) given by (16).
4. Compute \( \mathbf{Z}' \) by using Dailey’s (or Friswell’s, if necessary) method.
5. Compute the derivatives as \( \mathbf{Y}' = \mathbf{Z}' \Gamma - 1 = \mathbf{Z}' \Gamma^T \).

We select as many eigenvectors as the multiplicity of the eigenvalue, this is not strictly necessary but convenient taking into account that for computing the derivative of an eigenvector associated to a multiple eigenvalue we need to know both an adjacent basis of eigenvectors and their derivatives. This multiple selection procedure, even in the spirit of MAC, seems also to be a novelty of this paper, as far as the authors may claim.

Once the reference eigenvectors are computed, the adjacent eigenvectors are calculated from them, and derivatives are obtained from the derivatives computed for the adjacent eigenvectors using the linear transformation from this former basis to the reference one. This is summarized in Table 4.

The validity of this procedure is shown through two examples. First, a square plate clamped at its four edges with four elements and degrees of freedom, and second, a two-element cantilever. In both of them, we compute eigenvector derivatives using our method and by numerical differentiation, checking that the same numerical results are obtained.

We would like to remark that those examples are to a certain extend artificial since those have been built up to illustrate our method.

Four-edge clamped plate

In this example a square plate, shown in Fig. 2, is considered in order to ensure repeated natural frequencies. Four square finite elements are used to model the plate. The elemental stiffness and mass matrices are defined as

\[
\mathbf{K}_e = \frac{E_0 t}{8(1 - \nu^2)} \begin{bmatrix}
4 & 1 & -2 & -1 & -2 & -1 & 0 & 1 \\
1 & 4 & 1 & 0 & -1 & -2 & -1 & -2 \\
-2 & 1 & 4 & -1 & 0 & -1 & -2 & 1 \\
-3 & 0 & -1 & 4 & 1 & -2 & -1 & 1 \\
-2 & -1 & 0 & 1 & 4 & 1 & -2 & -1 \\
-1 & -2 & -1 & 2 & 1 & 4 & 1 & 0 \\
0 & -1 & -2 & 1 & 1 & 4 & 1 & -1 \\
1 & -2 & 1 & -2 & 1 & 0 & -1 & 4 \\
\end{bmatrix}
\]

\[
\mathbf{M}_e = \frac{\rho m t L^2}{36} \begin{bmatrix}
4 & 0 & 2 & 0 & 1 & 0 & 2 & 0 \\
0 & 4 & 0 & 2 & 0 & 1 & 0 & 2 \\
2 & 0 & 4 & 2 & 0 & 2 & 0 & 1 \\
0 & 2 & 0 & 4 & 0 & 2 & 0 & 1 \\
1 & 0 & 2 & 0 & 4 & 0 & 2 & 0 \\
0 & 1 & 0 & 2 & 0 & 4 & 0 & 2 \\
2 & 0 & 1 & 0 & 2 & 0 & 4 & 0 \\
0 & 2 & 0 & 1 & 0 & 2 & 0 & 4 \\
\end{bmatrix}
\]

where \( E_0 \) is the Young’s modulus of the material, \( t \) is the thickness and \( \nu \) is the Poisson’s ratio.

being \( \rho_m \) the material density and \( L \) the element length.
This structural problem involves one variable \(\rho_e\) per element, which is the structural density of the finite element. \(\rho_e = 0\) and \(\rho_e = 1\) mean void and solid, respectively [1].

Global stiffness and mass matrices (both depending linearly with the structural variable \(\rho_e\)) are constructed by assembling the element matrices \(K_e\) and \(M_e\) in the usual way,

\[
K = \sum_{e=1}^{4} \rho_e K_e, \quad M = \sum_{e=1}^{4} \rho_e M_e.
\]

(23)

The eigenproblem is expressed as follows,

\[
\begin{align*}
(K - \lambda_j M)U_j &= 0, & j = 1, 2 \\
U_i^T M U_j &= \delta_{ij}, & i, j = 1, 2
\end{align*}
\]

(24)

where the vector \(U\) stores the displacements of the free node in the \(x\)- and \(y\)-axes,

\[
U = \begin{bmatrix} u_x \\ u_y \end{bmatrix}.
\]

(25)

Set \(E_0 = 1\), \(v = 0.3\), \(L = 0.5\), \(t = 0.01\), \(\rho_m = 1\). We compute derivatives when variables vector \(x = \rho = [\rho_1, \rho_2, \rho_3, \rho_4] = [1, 1, 1, 1]\). For these values of \(\rho\), eigenvalues are double and equal to 17.8022, and obtained eigenvectors (by a standard algorithm for computing eigenvectors) are collected into the matrix \(U\),

\[
U = \begin{bmatrix} 21.2159 & 21.2159 \\
21.2159 & -21.2159 \end{bmatrix}.
\]

(26)

The target vectors are obtained from those eigenvectors under the action of a rotation, getting the ones parallel to the \(x\)- and \(y\)-axes, that is,

\[
\tilde{Y} = \begin{bmatrix} 30.0000 & 0.0000 \\
0.0000 & 30.0000 \end{bmatrix}.
\]

(27)

The eigenvectors \(U\) and \(\tilde{Y}\) are represented in Fig. 3 (both sets are scaled, since the \(M\)-normalized displacements are large).

In general, the reference set of eigenvectors \(Y\) should be obtained from \(\tilde{Y}\) through the procedure of the previous section, i.e. through the optimization problem (20), that is to say

\[
Y = \hat{U} \Gamma_0,
\]

with \(\Gamma_0\) the optimal solution of

\[
\begin{cases}
\max_{\Gamma_0: \Gamma_0^T \Gamma_0 = 1} \|\tilde{Y}^T M Y\| \\
\text{s.t.:} & Y = U \Gamma_0.
\end{cases}
\]

Eventually, it could happen \(Y = \tilde{Y}\) if the columns of \(\tilde{Y}\) are an \(M\)-orthonormal eigenvectors basis of the same eigenvalue, which is the case of this example since \(\tilde{Y}\) is obtained from \(U\) under the action of a rotation, so that \(Y = \tilde{Y}\).

Thanks to the symmetry of the problem, the derivatives with respect to any of the four variables are very similar, so that for the sake of brevity only the derivatives with respect to \(\rho_1\) will be shown. The procedure is the following:
1. Compute the adjacent eigenvectors from \( Y \), i.e. solve the eigenvalue problem

\[
D \Gamma = \Gamma \Lambda',
\]

with

\[
D = Y^T(K' - \lambda M')Y,
\]

and the adjacent eigenvectors are obtained as

\[
Z = Y\Gamma.
\]

Notice that \( \Lambda' \) is a diagonal matrix whose diagonal elements are the eigenvalue derivatives, which in this case are different permitting to use Dailey’s method. Solving (16) we arrive at

\[
\Gamma = \begin{bmatrix} 0.7071 & 0.7071 \\ 0.7071 & -0.7071 \end{bmatrix}
\]

and

\[
\Lambda' = \begin{bmatrix} 4.0797 & 0 \\ 0 & 13.7225 \end{bmatrix}.
\]

Now, we can compute the adjacent eigenvectors

\[
Z = \begin{bmatrix} 21.2132 & 21.2132 \\ 21.2132 & -21.2132 \end{bmatrix}.
\]

2. Compute the adjacent eigenvectors derivatives by Dailey’s method. We obtain

\[
\frac{\partial Z}{\partial \rho_1} = \begin{bmatrix} -2.6517 & -2.6517 \\ -2.6517 & 2.6517 \end{bmatrix}.
\]

3. Compute the derivative of \( Y \) as

\[
\frac{\partial Y}{\partial \rho_1} = \frac{\partial Z}{\partial \rho_1} \Gamma^T,
\]

obtaining

\[
\frac{\partial Y}{\partial \rho_1} = \begin{bmatrix} -3.7500 & 0.0000 \\ 0.0000 & -3.7500 \end{bmatrix}.
\]

Now, in order to validate our procedure, we compute these derivatives by numerical differentiation and check that the results coincide with those obtained with our analytical method. The numerical derivatives of the problem
are computed by using finite-difference methods. Both eigenvalues, \( \lambda_1, \lambda_2 \), are different and coalesce at the point \( \rho = [1, 1, 1, 1] \). After this point, the eigenvalues separate again, but their order has changed, as shown in Fig. 1. Backward formula is used instead of the centered one in order to prevent errors due to this fact. We remark that a centered scheme could be used but taking into account the eigenvalues change off. The first order backward formula used to compute numerically the derivatives with respect to \( \rho_1 \) is expressed as follows:

\[
\frac{\partial \mathbf{Y}}{\partial \rho_1} = \frac{\mathbf{Y}_{\rho_1} - \mathbf{Y}_{\rho_1 - h_N}}{h_N},
\]

where \( h_N \) is the step, \( \partial \mathbf{Y} / \partial \rho_1 \) is the numerically computed derivative and \( \mathbf{Y}_{\rho_1} \) are the set of eigenvectors obtained by solving Eq. (24) when the structural density of the first discrete element is perturbed as follows: \( \rho_1^- = \rho_1 - h_N \). Taking step \( h_N = 10^{-3} \) the numerical derivatives are

\[
\frac{\partial \mathbf{Y}}{\partial \rho_1} = \begin{bmatrix} -3.7507 & 0.0000 \\ 0.0000 & -3.7507 \end{bmatrix}.
\]

(28)

In this example the analytically and numerically computed derivatives differ no more that 0.02%, showing and corroborating the validity of our method.

**Two-element cantilever**

In this example we consider a beam-type finite element model for the cantilever shown in Fig. 4. For the sake of clarity only two finite elements are used, with 8 degrees of freedom per element, as usual for beam-type finite elements.

The element stiffness and mass matrices are given by the following expressions:

\[
\mathbf{K}_e = \frac{E_0}{L} \begin{bmatrix} 12I_e/L^2 & 0 & 0 & 6I_e/L & -12I_e/L^2 & 0 & 0 & 6I_e/L \\ 0 & 12I_e/L^2 & -6I_e/L & 0 & 0 & -12I_e/L^2 & 0 & -6I_e/L \\ 0 & -6I_e/L & 4I_e & 0 & 0 & 6I_e/L & 2I_e & 0 \\ 6I_e/L & 0 & 4I_e & -6I_e/L & 0 & 0 & 2I_e & 0 \\ -12I_e/L^2 & 0 & 0 & -6I_e/L & 12I_e/L^2 & 0 & 0 & -6I_e/L \\ 0 & -12I_e/L^2 & 6I_e/L & 0 & 0 & 12I_e/L^2 & 6I_e/L & 0 \\ 0 & -6I_e/L & 2I_e & 0 & 0 & 6I_e/L & 4I_e & 0 \\ 6I_e/L & 0 & 2I_e & -6I_e/L & 0 & 0 & 4I_e & 0 \end{bmatrix}
\]

(29)

\[
\mathbf{M}_e = \frac{\rho AL^3}{420} \begin{bmatrix} 156 & 0 & 0 & 22L & 54 & 0 & 0 & -13L \\ 0 & 156 & -22L & 0 & 0 & 54 & 13L & 0 \\ 0 & -22L & 4L^2 & 0 & 0 & -13L & -3L^2 & 0 \\ 22L & 0 & 0 & 4L^2 & 13L & 0 & 0 & -3L^2 \\ 54 & 0 & 0 & 13L & 156 & 0 & 0 & -22L \\ 0 & 54 & -13L & 0 & 0 & 156 & 22L & 0 \\ 0 & 13L & -3L^2 & 0 & 0 & 22L & 4L^2 & 0 \\ -13L & 0 & 0 & -3L^2 & -22L & 0 & 0 & 4L^2 \end{bmatrix}
\]

(30)

where \( E_0 \) is the stiffness, \( L \) is the length of the element, \( \rho \) is the mass density, \( A \) is the cross-sectional area and \( I_y \) and \( I_z \) are the area moments for the \( y \) and \( z \) axes, respectively. For the sake of clarity and to relate with the previous section, both variables form the vector of variables \( \mathbf{x} = [I_y, I_z] \). These last two parameters, the inertia moments, are the variables with respect to which the problem will be differentiated. In the case of a square section both area moments \( I_y \) and \( I_z \) are identical, causing the multiplicity of the eigenvalues.

The global stiffness and mass matrices are constructed by assembling the element matrices \( \mathbf{K}_e \) and \( \mathbf{M}_e \) of both elements:

\[
\mathbf{K} = \sum_{e=1}^{2} \mathbf{K}_e (I_y, I_z), \quad \mathbf{M} = \sum_{e=1}^{2} \mathbf{M}_e.
\]

(31)

The eigenproblem is

\[
\begin{bmatrix} (\mathbf{K} - \lambda_j \mathbf{M}) \mathbf{U}_j = 0, \quad & j = 1, \ldots, 8 \\
\mathbf{U}_i^T \mathbf{M} \mathbf{U}_j = \delta_{ij}, \quad & i, j = 1, \ldots, 8 \end{bmatrix}
\]

(32)
where $\lambda_j$ are the eigenvalues, i.e. the square of the natural frequencies, and $U_j$ the mode shapes, that store the rotations and displacements of the structure,

$$U = [v_1, w_1, \theta_{y1}, v_2, w_2, \theta_{y2}, \theta_{z2}]^T.$$  \hspace{1cm} (33)

Set $E = 1000$, $A = 1$ and $L = \rho = 1$. When $I_y = I_z = 1$ in Eq. (32), two first eigenvalues are equal, so that we have a double eigenvalue $\lambda_1 = \lambda_2 = 1.8414$, and a computed eigenvector basis collected in the matrix $U$,

$$U = 
\begin{bmatrix}
-0.0369 & 0.0584 \\
-0.0584 & -0.0369 \\
-0.0402 & -0.0254 \\
0.0254 & -0.0402 \\
-0.0125 & 0.0198 \\
-0.0198 & -0.0125 \\
-0.0340 & -0.0215 \\
0.0215 & -0.0340 \\
\end{bmatrix}.$$  \hspace{1cm} (34)

Target eigenvectors are collected in matrix $\tilde{Y}$, whose columns are linearly independent (and actually $M$-orthonormal) eigenvectors associated to the double eigenvalue (and therefore may be obtained as linear combinations of the vectors in $U$):

$$\tilde{Y} = 
\begin{bmatrix}
0.0488 & 0.0488 \\
0.0488 & -0.0488 \\
0.0336 & -0.0336 \\
-0.0336 & -0.0336 \\
0.0166 & 0.0166 \\
0.0166 & -0.0166 \\
0.0284 & -0.0284 \\
-0.0284 & -0.0284 \\
\end{bmatrix}.$$  \hspace{1cm} (35)

Now, the reference set of eigenvectors $Y$ is obtained from $\tilde{Y}$ as the solution of the maximization principle (20). Indeed, we get that the components of $Y$ differ less than $10^{-8}$ from the corresponding components of $\tilde{Y}$, so that with such a small difference in this situation we may assume $Y = \tilde{Y}$. The difference between both sets is related to how the reference set $\tilde{Y}$ is defined. In this example $\tilde{Y}$ is obtained by the inclusion of a small perturbation in the variables. Since this perturbation is very small (but big enough to separate the eigenvalues) both sets of eigenvector must be very similar and therefore $\tilde{Y}$ almost $M$-orthonormal, but they will never be exactly the same because matrices $K$ and $M$ change slightly with the change in the variables.
Dailey’s method is used to compute the eigenvector derivatives with respect to \( I_y \) and \( I_z \). Following the steps described in Table 2, we arrive at

\[
\Gamma_{I_y} = \begin{bmatrix} 0.7071 & 0.7071 \\ -0.7071 & 0.7071 \end{bmatrix}, \quad \Gamma_{I_z} = \begin{bmatrix} -0.7071 & 0.7071 \\ 0.7071 & 0.7071 \end{bmatrix}
\]

and

\[
\frac{\partial \mathbf{Z}_{I_y}}{\partial I_y} = \begin{bmatrix} 0.0000 & 0.0000 \\ -0.0015 & 0.0015 \\ -0.0036 & 0.0036 \\ 0.0000 & 0.0000 \\ 0.0008 & -0.0008 \\ 0.014 & -0.0014 \\ 0.0000 & 0.0000 \end{bmatrix}, \quad \frac{\partial \mathbf{Z}_{I_z}}{\partial I_z} = \begin{bmatrix} 0.0000 & -0.0021 \\ 0.0000 & 0.0000 \\ 0.0000 & 0.0000 \\ 0.0000 & 0.0000 \\ 0.0000 & 0.0000 \\ 0.0000 & -0.0020 \end{bmatrix}.
\]

Notice that the rotation matrices for obtaining the adjacent eigenvectors (\( \mathbf{Z}_{I_y}, \mathbf{Z}_{I_z} \)) are different for each variable, that is, \( \Gamma_{I_y} \neq \Gamma_{I_z} \). Finally, following procedure given in Table 4, we obtain that the partial derivatives of the eigenmodes given by \( \mathbf{Y} \) are

\[
\frac{\partial \mathbf{Y}}{\partial I_y} = \frac{\partial \mathbf{Z}_{I_y}}{\partial I_y} \Gamma_{I_y}^T, \quad \frac{\partial \mathbf{Y}}{\partial I_z} = \frac{\partial \mathbf{Z}_{I_z}}{\partial I_z} \Gamma_{I_z}^T.
\]

As we did in the above example, we compute these derivatives by numerical differentiation and check that the results coincide with those obtained with our analytical method. The first order backward formula used to compute numerically the derivatives with respect to both variables \( I_y \) and \( I_z \) is expressed as follows:

\[
\frac{\partial_N \mathbf{Y}}{\partial I_y} = \frac{\mathbf{Y}_{I_y} - \mathbf{Y}_{I_y^-}}{h_N}, \quad \frac{\partial_N \mathbf{Y}}{\partial I_z} = \frac{\mathbf{Y}_{I_z} - \mathbf{Y}_{I_z^-}}{h_N},
\]

where \( h_N \) is the step, \( \partial_N \mathbf{Y}/\partial I_y \) and \( \partial_N \mathbf{Y}/\partial I_z \) are the numerically computed eigenvector derivatives and \( \mathbf{Y}_{I_y^-} \) and \( \mathbf{Y}_{I_z^-} \) are the set of eigenvectors obtained by solving Eq. (32) when the variables \( I_y \) and \( I_z \) (for the upper element) are perturbed as follows: \( I_y^- = I_y - h_N \) and \( I_z^- = I_z - h_N \). The value of these numerically computed derivatives for \( h_N = 10^{-3} \) is

\[
\frac{\partial_N \mathbf{Y}}{\partial I_y} = \begin{bmatrix} 0.0000 & 0.0000 \\ -0.0015 & 0.0015 \\ -0.0036 & 0.0036 \\ 0.0000 & 0.0000 \\ 0.0008 & -0.0008 \\ 0.014 & -0.0014 \\ 0.0000 & 0.0000 \end{bmatrix}, \quad \frac{\partial_N \mathbf{Y}}{\partial I_z} = \begin{bmatrix} 0.0015 & -0.0015 \\ 0.0000 & 0.0000 \\ 0.0000 & 0.0000 \\ -0.0036 & 0.0036 \\ -0.0008 & 0.0008 \\ 0.014 & -0.0014 \\ 0.0000 & 0.0000 \end{bmatrix}.
\]

The difference between the analytically and the numerically computed eigenvectors derivatives is smaller than 0.01\%, showing the validity of the proposed method.

4. Application to optimization problems

The goal of our procedure is to compute sensitivities of functionals depending on eigenvectors appearing in optimization problems, either in the cost or the constraints, or both of them. In this section we analyze an example of an eigenmode optimization problem, showing how our procedure for computing eigenmode sensitivities succeeds
in applying a gradient-based optimization algorithm. The algorithm chosen to test our claim is MMA, the Method of Moving Asymptotes [25], of common use in topology optimization problems. We carry out this validation with a complex example in which both cost and constraint depend on eigenvectors. The example is the optimal design of modal piezoelectric transducers [26]. This problem has already been analyzed and studied in [9,10], and here it is used in order to illustrate how our method works out.

The problem is the following:

$$\max_{\rho_s, \rho_p, \alpha} : F_k - \alpha$$

subject to

$$\begin{align*}
(K - \mu_j M)\Phi_j &= 0, \quad j = 1, \ldots, J \\
\Phi_j^T M \Phi_l &= 0, \quad j, l = 1, \ldots, J, \quad j \neq l \\
\Phi_j^T \Phi_j &= 1, \quad j = 1, \ldots, J \\
|F_j| &\leq \alpha, \quad j = 1, \ldots, J, \quad j \neq k \\
\rho_s &\in [0, 1] \\
\rho_p &\in [0, 1] \\
\alpha &\geq 0.
\end{align*}$$

(4) The coefficient $F_j$ is

$$F_j = C^T \Phi_j,$$

(43) with

$$C = R(\rho_s)(2\rho_p - 1)B,$$

(44) where $R$ is an interpolation function that models the piezoelectric effect in empty areas, $\rho_s$ is the structural density variable, $\rho_p$ is the electrode profile and $B$ is the typical FE strain–displacement matrix. $\Phi_j$ is the $j$th mode shape, $\mu_j$ is the $j$th eigenvalue, $K$ and $M$ are the global stiffness and mass matrices, respectively, and $\alpha$ is an independent auxiliary variable used to model the problem. For details on the modeling the interested reader is referred to [9] and [10].

Coefficient $F_j$ models the charge collected by the electrodes for the $j$th mode shape, up to a piezoelectric constant. Then the objective of the problem is maximizing the charge collected for the mode shape of interest while the charge of the rest of mode shapes is suppressed as much as possible.

We consider the case of a square domain clamped at its four corners as shown in Fig. 5. The number of mode shapes used in the problem is $J = 4$. Young’s modulus is set to $E = 167$ GPa, Poisson’s ratio $\nu = 0.3$, length $L = 1$m and thickness $t = 0.01$m. First four mode shapes are represented in Fig. 6. Setting $k = 1$ the first mode shape is isolated from the rest. This means that the coefficient $F_1$ is maximized while coefficients $F_2, F_3$ and $F_4$ are minimized.

Now we would like to remark what ordering in the mode shapes means. We take as target mode shapes those of the whole square plate, when in a certain iteration we have a structure layout $\rho_s$, we compute the modes, and take those closest to the first, second, third and fourth target mode shapes. Since the algorithm is iterative, this has to be done at any iteration, being computationally costly, but preventing mode switching (change of order of eigenfrequencies in
Fig. 6. Mode shapes of the plate.

Fig. 7. Mode shapes computed at $i$th iteration step.

The way we select the closest modes is by the modal assurance criterion, MAC, or its variant for multiple selection introduced in the previous section through the maximization principle (20). At any iteration, a set of $\tilde{J}$ eigenvectors, where $\tilde{J} > J$ has to be computed and among those select the one whose inner product with the $k$th target mode shape is closest to one, and this will be the $k$-eigenvector for the actual iteration. Since, mode shapes and target mode shapes are normalized to be unitary (euclidean norm equal to 1), their inner product will be smaller than or equal to one, and it gets maximal when both vectors are equal. Fig. 6 shows the target mode shapes (the first four mode shapes of the whole plate), and in Fig. 7 we see the first eight mode shapes in a disordered way, and in Table 5 how through the MAC method they are assigned in the right way. We refer readers to [20] for an explanation of MAC, and to [9,10] for a complete exposition of the use of MAC in this situation. Mode tracking is imperative in this sort of problems, otherwise the mode shapes that are being optimized and suppressed would change in each iteration jeopardizing the convergence of the iterative method.
Table 5
MAC for mode shapes of reference.

<table>
<thead>
<tr>
<th></th>
<th>$\hat{Y}_1$</th>
<th>$\hat{Y}_2$</th>
<th>$\hat{Y}_3$</th>
<th>$\hat{Y}_4$</th>
<th>$\hat{Y}_5$</th>
<th>$\hat{Y}_6$</th>
<th>$\hat{Y}_7$</th>
<th>$\hat{Y}_8$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_1$</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$Y_2$</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$Y_3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$Y_4$</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Fig. 8. Mode shape of reference $Y$ (left) and computed at the first iteration step $U$ (right).

Coming back to our problem, note that the first and the second mode shapes present the same eigenfrequency in the first iteration, and then the derivatives of $F_1$ and $F_2$, that depend on the eigenvectors $\Phi_1$ and $\Phi_2$, must be computed by using the new method proposed in the previous section. During the iterative optimization procedure eigenvalues will split out and become simple, and on the other hand other eigenvalues will coalesce and become double during one iteration.

The reference mode shapes obtained in the first iteration are shown in Fig. 8.

In this case, it is easy to identify the mode shapes. The eigenfrequency for the first and second modes is repeated, and then the basis of eigenvectors is not unique. Before computing the derivatives, the finite dimensional optimization
problem proposed in the previous section must be solved. The derivatives of $F_3$ and $F_4$ are computed by using Nelson’s method, and the derivatives of $F_1$ and $F_2$ with the new method proposed.

Fig. 9 shows the results of the optimization problem. Fig. 9 left shows the structure density, where black and white mean solid and void areas, respectively. Fig. 9 right shows the polarization profile, where cyan and orange colors represent areas of the structure covered by positive and negative electrodes, respectively.

The evolution of the objective function is shown in Fig. 10. It is noticed that the cost converges in less than 100 iterations. The initial point of the iterative process is the optimized electrode of the mode shape of interest for the homogeneous plate.

A second example is presented in Fig. 11, where the second mode shape is isolated. It is shown in Fig. 8b that the plate moves in the direction which is a linear combination of the movements on the $x$- and $y$-axes. Nevertheless, at the end of the iterative process, the solution (optimized electrode profile) presents a polarity that coincides with the mode shape of reference $Y_2$, in other words, the jumps between positive and negative polarity coincides with the parts of the structure that are subject to traction and compression in the reference mode shape $Y_2$. A summary of the process is shown in Table 6.
Table 6

Steps of the optimization process.

1. Set the constants of the problem: stiffness, thickness, Poisson’s ratio and volume density.
2. Select the $J$ mode shapes of reference.
3. Initialize both design variables $\rho_s$ and $\rho_p$.
4. Compute the first $J$th eigenmodes of the plate and identify the $J$th closest modes to the reference by means of MAC.
5. Check the multiplicity of the eigenvalues.
   - If it is simple compute the coefficients $F_j$ and their sensitivities by using Nelson’s method.
   - If it is multiple proceed in the following way:
     – Find a new basis of orthonormal eigenvectors as solution of the finite dimensional optimization problem presented in Eq. (20).
     – Compute the coefficients $F_j$ and their sensitivities by using the new method proposed in this work.
6. Update the new set of eigenvectors of reference, replacing the previous one with the new set computed (if needed).
7. Update the design variables $\rho_s$ and $\rho_p$ by using MMA.
8. Until convergence, go back to step 4.

5. Conclusions

In this paper, we study eigenmode optimization problems from a general framework perspective whenever tracking specific mode shapes. The main questions we address are: mode shape selection, eigenmode differentiability and eigenmode derivative calculation. We present a novel method for mode shape selection very close, and in the spirit of the MAC (modal assurance criterion) but also considering the possibility of multiple selection, what is necessary in the case of mode shapes associated to multiple eigenvalues. For those selected mode shapes, which in general differ from the adjacent eigenmodes, we propose a simple algorithm, that takes advantage of the linear structure of eigenspaces, for computing their derivatives. Those derivatives are required by any gradient-based optimization method. A review on the existing mathematical results on eigenvector differentiability is given, including a result in [24] that gives theoretical validity to our method. Further, the success of our algorithm is numerically corroborated with three examples: two structural analysis examples in which the algorithm for derivatives computation is numerically validated comparing with the results obtained via numerical differentiation; and an optimization example in which both cost and constraints depend on eigenmodes associated to multiple eigenvalues, and for which the optimization MMA algorithm together with our selection mode shape and computation of derivatives procedures succeed in reaching optimal solutions.

Acknowledgments

We appreciate very much the comments and suggestions of an anonymous referee that considerably improved the quality of the manuscript. This work has been funded by MINECO (Spain) through grant MTM2013-47053-P and Junta de Comunidades de Castilla-La Mancha (Spain) through grant PEII-2014-010-P.

References