Theor Chim Acta (1989) 76: 331–351

Theoretica Chimica Acta



# Solution of the generalized eigenvalue equation perturbed by a generalized low rank perturbation*

Tomislav P. Živković**
Texas A & M University at Galveston, Galveston, TX 77553-1675, USA

(Received February 10, revised August 1/Accepted August 3, 1989)

**Summary.** The problem of finding eigenvalues and eigenstates of the generalized perturbed eigenvalue equation $(\mathbb{B}+\mathbb{V})\Psi=\varepsilon(\mathbb{C}+\mathbb{P})\Psi$ is considered. The eigenvalues and the eigenstates of the unperturbed eigenvalue equation $\mathbb{B}\Phi=\lambda\mathbb{C}\Phi$ are assumed to be known. Matrices $\mathbb{B}$, $\mathbb{V}$, $\mathbb{C}$ and $\mathbb{P}$ can be arbitrary, except for the requirement that $\mathbb{C}$ be nonsingular and that the eigenstates of the unperturbed equation be complete. It is shown that the eigenvalues and the eigenstates of the perturbed equation can be easily obtained if the rank of the generalized perturbation $\{\mathbb{V}, \mathbb{P}\}$ is small. A special case of low rank perturbations are piecewise local perturbations which are common in physics and chemistry. If the perturbation is piecewise local with fixed localizability, the operation count for the derivation of a single eigenvalue and/or a single eigenstate is $\mathcal{O}(n)$. If the perturbation has a fixed rank, the operation count for the derivation of all eigenvalues and/or all eigenstates is $\mathcal{O}(n^2)$.

**Key words:** Generalized perturbed eigenvalue equation — Low rank perturbation — Piecewise local perturbations — Non-Hermitian matrices

## 1. Introduction

The solution of the generalized eigenvalue equation

$$\mathbb{A}\Psi_k=\varepsilon_k\mathbb{S}\Psi_k, \tag{1}$$

* Research supported by the Welch Foundation of Houston, Texas, and by the Yugoslav Ministry for Development (Grant P-96)

** *Permanent address*: Rudjer Bošković Institute, 41000 Zagreb, Croatia, Yugoslavia

where $\mathbb{A}$ and $\mathbb{S}$ are $n$-order matrices, is one of the most important numerical problems in quantum chemistry. Typically, $\mathbb{A}$ represents the Hamiltonian, while $\mathbb{S}$ is the overlap matrix defined over some set of basis functions. In many quantum chemical models these basis functions are assumed to be orthonormalized, and instead of the generalized eigenvalue equation (1) one solves a simple eigenvalue equation with $\mathbb{S} = \mathbb{1}$. However, in the SCF *ab initio* and in the more sophisticated SCF semiempirical methods, one cannot neglect the overlap between the basis functions. For example, the overlap between the $2p_z$ atomic orbitals situated on adjacent carbon atoms is $\approx 0.25$, which is not negligible. The standard approach is an explicit or implicit orthonormalization of the initial orbitals. However, this is not always practical, especially when interacting molecules are considered. The reason for the neglect of the overlap is mainly numerical, it is much easier to solve the eigenvalue equation (1) with $\mathbb{S} = \mathbb{1}$. It is however more realistic to assume that the overlap matric $\mathbb{S}$ is not trivial.

There are some other quantum chemical problems which may lead to the generalized eigenvalue equation (1), and also to the eigenvalue equation involving non-Hermitian matrices. For example, the eigenvalue problems encountered in the study of molecular vibrational spectra involve two matrices, matrix $\mathbb{F}$ and matrix $\mathbb{G}$. The former matrix is associated with the force field, whereas the latter is associated with the kinetic energy. Depending on the implementation, the corresponding eigenvalue equation is either a generalized eigenvalue equation (1) with $\mathbb{S} \equiv \mathbb{G}^{-1} \neq \mathbb{1}$ or $\mathbb{S} \equiv \mathbb{F}^{-1} \neq \mathbb{1}$, or it is an eigenvalue equation for the (generally non-Hermitian) matrix $\mathbb{FG}$ [1].

Many problems involving the solution of the eigenvalue equation are related to each other. If the solution to some system described by the matrix $\mathbb{A}$ is required, one usually already knows the solution to some related "unperturbed" system described by the matrix $\mathbb{B}$. A natural idea is to exploit this knowledge in order to speed up the calculation and decrease the operation count. Such an approach may also provide a direct insight into the connection between related systems. Such systems are most naturally treated by the perturbation expansion method. This method is usually applied to the simple eigenvalue equation (1) with $\mathbb{S} = \mathbb{1}$, and it solves this equation exploiting the knowledge of the eigenvalues and the eigenstates of the related unperturbed system. The matrix $\mathbb{A}$ is represented as a sum $\mathbb{A} = \mathbb{B} + \mathbb{V}$, where matrix $\mathbb{B}$ describes the unperturbed system, whereas matrix $\mathbb{V}$ corresponds to the perturbation.

We present here an alternative method for the solution of the perturbed eigenvalue equation. In order to include all possibilities, we consider the generalized eigenvalue equation perturbed by the generalized perturbation

$$(\mathbb{B} + \mathbb{V})\Psi_k = \varepsilon_k(\mathbb{C} + \mathbb{P})\Psi_k. \tag{2}$$

With the choice $\mathbb{A} = \mathbb{B} + \mathbb{V}$ and $\mathbb{S} = \mathbb{C} + \mathbb{P}$ this equation is equivalent to the generalized eigenvalue equation (1). The eigenvalues $\lambda_i$ and the corresponding orthonormalized eigenstates $\Phi_i$ of the generalized "unperturbed" eigenvalue equation

$$\mathbb{B}\Phi_i = \lambda_i \mathbb{C}\Phi_i \tag{3}$$

are assumed to be known. For the sake of generality matrices $\mathbb{B}$, $\mathbb{C}$, $\mathbb{V}$, and $\mathbb{P}$ are not assumed to be Hermitian. This necessitates careful distinction between the left and right eigenstates.

Matrix $\mathbb{V}$ in (2) is the "perturbation" of the unperturbed matrix $\mathbb{B}$, while matrix $\mathbb{P}$ is the "perturbation" of the unperturbed matrix $\mathbb{C}$. Since matrices $\mathbb{B}$ and $\mathbb{C}$ usually correspond to the unperturbed Hamiltonian and to the unperturbed overlap matrix, respectively, matrices $\mathbb{V}$ and $\mathbb{P}$ can be interpreted as perturbations of the Hamiltonian and overlap matrix, respectively. The perturbation $\{\mathbb{V}, \mathbb{P}\}$ is "generalized", unless the perturbation of the "overlap" matrix $\mathbb{C}$ is trivial, i.e. unless $\mathbb{P} = \mathbb{0}$.

Many perturbations encountered in quantum chemistry are either "local" or "piecewise local". Such perturbations affect only a small portion of the total unperturbed system. For example, two molecules may differ in a single atom. One of these two molecules can be considered to describe the unperturbed system, and the "perturbation" corresponds to the replacement of one atom with another. This perturbation is nonzero only over a small portion of the unperturbed system, and hence it is local. Other examples of local perturbations are creation and destruction of a single bond, local interactions between two molecules, etc. A piecewise local perturbation is a more general perturbation which may be expressed in terms of a few local perturbations. Examples are the replacement of a few selected atoms, creation and destruction of a few selected bonds, etc. Unless the combined perturbation involves too many local perturbations, the region affected by the perturbation is still "small" with respect to the total region occupied by the unperturbed system. This property is essential for the perturbation to be considered "piecewise local".

Standard perturbation methods are not well suited to treat piecewise local perturbations. The essential condition for the numerical efficiency of the perturbation expansion: a small perturbation and thus fast convergence, is usually not satisfied. Though the substitution of an atom with another is a local perturbation, it is not necessarily small. Similarly, creation or destruction of a single bond is by no means a small perturbation.

The method to be presented is particularly suitable for the treatment of the generalized perturbed eigenvalue equation (2) with a generalized piecewise local perturbation $\{\mathbb{V}, \mathbb{P}\}$. More generally, this method can be efficiently applied to all perturbations $\{\mathbb{V}, \mathbb{P}\}$ such that matrices $\mathbb{V}$ and $\mathbb{P}$ have low rank. Piecewise local perturbations are a special case of a low rank perturbation (LRP).

The LRP method produces the eigenvalue(s) and the eigenstate(s) of the generalized eigenvalue equation (2) in a finite and predictable number of steps. The method produces correct results for an arbitrary perturbation $\{\mathbb{V}, \mathbb{P}\}$. The requirement that the ranks of $\mathbb{V}$ and $\mathbb{P}$ be small is only a numerical one. Unless this requirement is satisfied, the operation count of the LRP method can be large, and some other method may be more efficient in the numerical solution of (2).

If the perturbation is piecewise local (with arbitrary but fixed localizability), then the operation count needed to derive a single eigenvalue $\varepsilon_0$ of (2) is $\mathcal{O}(n)$. If the eigenvalue $\varepsilon_0$ is known, the operation count needed to derive the corresponding eigenstate(s) is also $\mathcal{O}(n)$. If the perturbation is not piecewise local but has a

low fixed rank, then the operation count needed to derive all eigenvalues of (2) is $\mathcal{O}(n^2)$. If these eigenvalues are known, the operation count needed to derive all the corresponding eigenstates is also $\mathcal{O}(n^2)$. For large $n$ these operation counts are significantly lower than the corresponding operation counts of any other method for the solution of a generalized eigenvalue equation (2).

The method presented here was originally applied to the Hermitian eigenvalue equation (2) with $\mathbb{C} = \mathbb{1}$ and $\mathbb{P} = \mathbb{0}$ [2]. In computer calculations performed on a PC computer in double precision [2], the eigenvalues and eigenstates obtained were found to be correct up to $\approx 14$ significant figures. The times needed to derive these eigenstates and eigenvalues were also in agreement with the predicted operation counts, and in particular with the predicted dependence on $n$ [2].

It should be noted that a formula equivalent to (11a), but restricted to the nongeneralized Hermitian eigenvalue equation ($\mathbb{C} = \mathbb{1}$, $\mathbb{P} = \mathbb{0}$, $\mathbb{B}^\dagger = \mathbb{B}$, $\mathbb{V}^\dagger = \mathbb{V}$ where $\dagger$ denotes complex conjugate transpose), with the additional restriction that $|u_s\rangle = |v_s\rangle$ are the eigenstates while $w_s$ are the corresponding nonvanishing eigenvalues of $\mathbb{V}$ (these quantities enter the definition (12a) of functions $S^a_{sp}(\varepsilon)$ and the relation (11a)), has been apparently stated for the first time by Beattie and Fox [3] and independently by Arbenz and Golub [4]. Subject to the same restrictions, a formula equivalent to (14a) has been derived by Arbenz and Golub [4]. The equations derived by these authors relate the number of the positive eigenvalues of the unperturbed Hermitian matrix $\mathbb{B}$ with the number of the positive eigenvalues of the perturbed Hermitian matrix $(\mathbb{B} + \mathbb{V})$ [4]. Simpson derived a similar formula involving the number of positive eigenvalues, but for an even more restricted eigenvalue problem [5]. His formula was used for the frequency analysis of mechanical structures [6, 7].

These earlier approaches are different from the approach presented here. They use either some generalization of the Weinstein–Aronszajn determinant known from the methods of intermediate problems [8, 9], or more recently from Sylvester's law of inertia [10, 11]. Arbenz and Golub [4], for example, obtained their results using Silvester's law. Since this law applies only to symmetric matrices [10, 11], it is unlikely that their approach can be generalized to arbitrary non-Hermitian matrices. The generalization to the generalized eigenvalue equation ($\mathbb{C} \neq \mathbb{1}$) and to the generalized perturbations ($\mathbb{P} \neq \mathbb{0}$) is also not obvious. It should be noted however that these authors did generalize their results to some special non-Hermitian matrices [4]. They considered the modified unitary eigenvalue problem $\mathbb{U}^\dagger \mathbb{B} \Psi = \varepsilon \Psi$, where $\mathbb{B}$ is a unitary matrix with known spectral decomposition, $\mathbb{U}$ is a unitary matrix such that $\mathbb{1} - \mathbb{U}$ has low rank. By means of the Cayley transform [12] they transformed this unitary eigenvalue problem to the Hermitian one [4].

The method presented here applies to a much wider class of eigenvalue problems. Firstly, it is complete since it treats explicitly eigenvalues as well as eigenstates. Secondly, it applies to the generalized eigenvalue equation with $\mathbb{C} \neq \mathbb{1}$. Thirdly it allows for generalized perturbations with $\mathbb{P} \neq \mathbb{0}$. Finally, it is generalized to non-Hermitian matrices, which are of interest in the study of molecular vibrational spectra and also in many applications outside the realm of quantum chemistry. The only restriction to complete generality is the requirement that matrix $\mathbb{C}$ be non-singular and matric $\mathbb{C}^{-1/2}\mathbb{B}\mathbb{C}^{-1/2}$ non-defective.

## 2. The method

We consider a generalized perturbed eigenvalue equation with arbitrary matrices $\mathbb{V}$ and $\mathbb{P}$, an arbitrary non-singular matrix $\mathbb{C}$, and a matrix $\mathbb{B}$ subject to a mild condition to be shortly stated. If needed, all the results thus obtained can be easily specified for Hermitian matrices.

In the case of arbitrary matrices one has be distinguish between left and right eigenstates. Accordingly, the unperturbed eigenvalue equation (3) is replaced with

$$\mathbb{B}|\Phi_i^R\rangle = \lambda_i \mathbb{C}|\Phi_i^R\rangle, \qquad \langle\Phi_i^L|\mathbb{B} = \lambda_i \langle\Phi_i^L|\mathbb{C}, \tag{4}$$

where $\lambda_i$ are the eigenvalues, while $|\Phi_i^R\rangle$ and $\langle\Phi_i^L|$ are the corresponding right and left eigenstates, respectively. Note that right and left eigenvalues are the same, whereas right and left eigenstates may differ [13]. We assume that the eigenstates $|\Phi_i^R\rangle$ and $\langle\Phi_i^L|$ are complete, and that they satisfy the generalized biorthonormalized relation

$$\langle\Phi_i^L|\mathbb{C}|\Phi_j^R\rangle = \delta_{ij}. \tag{5}$$

We also assume that these eigenstates and the corresponding eigenvalues $\lambda_i$ are known.

The assumption that left and right eigenstates of (4) are complete and that they can be chosen to satisfy (5) can be stated in different ways. This assumption is equivalent to the requirement that the matrix $\mathbb{C}$ be non-singular and the matrix $\mathbb{C}^{-1/2}\mathbb{B}\mathbb{C}^{-1/2}$ non-defective (see Appendix). In particular, if matrix $\mathbb{B}$ is Hermitian and if $\mathbb{C}$ is the overlap matrix defined over linearly independent basic functions, this assumption is satisfied. This assumption is the only condition imposed on the otherwise arbitrary matrices $\mathbb{B}$, $\mathbb{V}$, $\mathbb{C}$ and $\mathbb{P}$.

By analogy with (4), the generalized eigenvalue equation (2) is replaced with

$$(\mathbb{B}+\mathbb{V})|\Psi_k^R\rangle = \varepsilon_k(\mathbb{C}+\mathbb{P})|\Psi_k^R\rangle, \tag{6a}$$

$$\langle\Psi_k^L|(\mathbb{B}+\mathbb{V}) = \varepsilon_k\langle\Psi_k^L|(\mathbb{C}+\mathbb{P}), \tag{6b}$$

where $\varepsilon_k$ are the eigenvalues, while $|\Psi_k^R\rangle$ and $\langle\Psi_k^L|$ are the corresponding right and left eigenstates, respectively.

Let $\varrho_v$ and $\varrho_p$ be the ranks of matrices $\mathbb{V}$ and $\mathbb{P}$, respectively. Define the "rank" $\varrho$ of the generalized perturbation $\{\mathbb{V}, \mathbb{P}\}$ as the sum

$$\varrho = \varrho_v + \varrho_p. \tag{7}$$

We now show that a special case of low rank perturbations ($\varrho \ll n$) are piecewise local perturbations.

Let the region $\mathscr{R}_v$ affected by the perturbation $\{\mathbb{V}, \mathbb{P}\}$ be only a small fraction of the total region $\mathscr{R}$ occupied by the unperturbed system. Region $\mathscr{R}_v$ contains $l \ll n$ basic vectors, where $l$ is by definition the "localizability" of the perturbation $\{\mathbb{V}, \mathbb{P}\}$. With an appropriate rearrangement of the rows and columns of $\mathbb{V}$ and $\mathbb{P}$, these matrices have a single $l \times l$ block which contains all nonzero matrix elements. The diagonalization of $\mathbb{V}$ and $\mathbb{P}$ reduces to the diagonalization of these

$l \times l$ submatrices. Since the rank of the square matrix equals the number of nonzero eigenvalues, the ranks $\varrho_v$ and $\varrho_p$ of matrices $\mathbb{V}$ and $\mathbb{P}$ satisfy $\varrho_v \leqslant l$ and $\varrho_p \leqslant l$. Each piecewise local perturbation ($l \ll n$) is hence a low rank perturbation ($\varrho \ll n$).

Matrices $\mathbb{V}$ and $\mathbb{P}$ can be written in the form

$$\mathbb{V} = \sum_{s=1}^{m} \omega_s |u_s\rangle\langle v_s|, \qquad \mathbb{P} = \sum_{s=1}^{\mu} \tau_s |x_s\rangle\langle y_s|, \tag{8}$$

where $\omega_s \neq 0$ and $\tau_s \neq 0$ are arbitrary scalars, while $|u_s\rangle$, $\langle v_s|$, $|x_s\rangle$ and $\langle y_s|$ are arbitrary vectors. There are many possible representations (8) of the same perturbation $\{\mathbb{V}, \mathbb{P}\}$. For example, one can always rescale the vectors $|u_s\rangle$, $\langle v_s|$, $|x_s\rangle$ and $\langle y_s|$ in such a way that all $\omega_s$ and $\tau_s$ equal unity. If $\mathbb{V}$ is non-defective, the scalars $\omega_s$ can be chosen to be eigenvalues of $\mathbb{V}$ and the vectors $|u_s\rangle$ and $<v_s|$ can be chosen to be the corresponding right and left eigenstates, etc.

Another useful representation of the perturbation $\{\mathbb{V}, \mathbb{P}\}$ is the representation (8) satisfying $m = \mu$ and $\langle y_s| = \langle v_s|$ $(s = 1, \ldots, m)$

$$\mathbb{V} = \sum_{s=1}^{m} \omega_s |u_s\rangle\langle v_s|, \qquad \mathbb{P} = \sum_{s=1}^{m} \tau_s |x_s\rangle\langle v_s|, \tag{9}$$

where $|\omega_s| + |\tau_s| \neq 0$ $(s = 1, \ldots, m)$. There is a similar representation with $|u_s\rangle = |x_s\rangle$.

Each perturbation $\{\mathbb{V}, \mathbb{P}\}$ can be represented in the form (9). If the perturbation is piecewise local, representation (9) is particularly suitable for the application of the LRP method. In this case the vectors $\langle v_s|$ can be chosen to be unit row vectors and (9) reduces to

$$\mathbb{V} = \sum_{s=1}^{m} |u_s\rangle\langle s|, \qquad \mathbb{P} = \sum_{s=1}^{m} |x_s\rangle\langle s|, \tag{9a}$$

where $|u_s\rangle$ and $|x_s\rangle$ are columns of matrices $\mathbb{V}$ and $\mathbb{P}$, respectively, while $\langle s|$ are the corresponding unit row vectors. The summation is performed only over these unit vectors $\langle s|$ for which either the corresponding column vector $|u_s\rangle$ or the corresponding column vector $|x_s\rangle$ is nonzero. If $\mathbb{V}$ and $\mathbb{P}$ are Hermitian, then $m = l$ is the localizability of the perturbation $\{\mathbb{V}, \mathbb{P}\}$. We refer to the representation (9a) as a "column-wise" representation of the perturbation $\{\mathbb{V}, \mathbb{P}\}$. There is a similar row-wise representation of $\{\mathbb{V}, \mathbb{P}\}$.

If the vectors $|u_s\rangle$ as well as the vectors $\langle v_s|$ in (8) are linearly independent, then $m = \varrho_v$ is the rank of $\mathbb{V}$. Otherwise $m > \varrho_v$. Similarly, if the vectors $|x_s\rangle$ as well as the vectors $\langle y_s|$ are linearly independent, then $\mu = \varrho_p$ is the rank of $\mathbb{P}$. Otherwise $\mu > \varrho_p$.

Any linear dependence between the vectors $|u_s\rangle$, $|v_s\rangle$, $|x_s\rangle$ or $|y_s\rangle$ can easily be eliminated. For example, if the vectors $|u_s\rangle$ are linearly dependent, then there exist nontrivial coefficients $c_s$ such that $\sum_s c_s |u_s\rangle = 0$. Let the $r$th coefficient $c_r$ be nonzero. Then $\mathbb{V}$ can be written in the form

$$\mathbb{V} = \sum_{s \neq r} \omega_s |u_s\rangle\langle v'_s|, \tag{10a}$$

where new vectors $\langle v'_s|$ are

$$\langle v'_s| = \langle v_s| - (\omega_r/c_r)(c_s/\omega_s)\langle v_r|. \tag{10b}$$

This reduces $m$ by one. The elimination process can be continued until the remaining vectors $|u_s\rangle$ are linearly independent. The same applies to other sets of linearly dependent vectors in the representation (8). Hence one can without loss of generality assume $m = \varrho_v$ and $\mu = \varrho_p$, i.e. $\varrho = (m + \mu)$.

In a similar way $m$ in the representation (9) can be reduced. Let $\varrho_0$ be the smallest $m$ which can be obtained with such a reduction. One easily finds $\varrho_0 \leqslant \varrho$. We shall find the quantities $\varrho$ and $\varrho_0$ important in the estimation of the operation counts for the derivation of the eigenvalues and the eigenstates of (6).

In the LRP approach it is convenient to distinguish "cardinal" and "singular" eigenvalues and eigenstates. If the eigenvalue $\varepsilon_0$ of (6) differs from all the eigenvalues $\lambda_i$ of the unperturbed eigenvalue equation (4), it is "cardinal" [2]. Otherwise, i.e. if $\varepsilon_0 \in \{\lambda_i\}$ it is "singular". Each eigenstate $\Psi$ corresponding to the cardinal eigenvalue $\varepsilon_0$ is cardinal, and each eigenstate $\Psi$ corresponding to the singular eigenvalue $\varepsilon_0$ is singular [2].

In addition, singular eigenstates can be "normal" or "exotic". Let $\varepsilon_0 = \lambda_k$ be a singular eigenvalue of (6), and let $\Psi$ be the corresponding eigenstate. Further, let $\lambda_k$ be a $\nu$-degenerate eigenvalue of the unperturbed eigenvalue equation (4), and let $|\Phi^R_{k\varkappa}\rangle$ and $\langle\Phi^L_{k\varkappa}|$ $(\varkappa = 1, \ldots, \nu)$ be the corresponding biorthonormalized right and left eigenstates, respectively. If $\Psi$ is the right eigenstate of (6) and it is a linear combination of the right eigenstates $|\Phi^R_{k\varkappa}\rangle$ $(\varkappa = 1, \ldots, \nu)$ of (4), or if $\Psi$ is the left eigenstate of (6) and it is a linear combination of the left eigenstates $|\Phi^L_{k\varkappa}\rangle$ $(\varkappa = 1, \ldots, \nu)$ of (4), then it is "normal". Otherwise it is "exotic".

We now formulate a few theorems which provide an efficient algorithm for the solution of the generalized eigenvalue equation (6). In these theorems it is assumed that left and right eigenstates of the unperturbed eigenvalue equation (4) are complete and that they satisfy the generalized biorthonormalized relation (5). This restriction affects only matrices $\mathbb{B}$ and $\mathbb{C}$, and apart from this restriction these matrices are arbitrary. The first two theorems refer to the perturbation $\{\mathbb{V}, \mathbb{P}\}$ represented in the general form (8).

**Theorem 1** (cardinal eigenvalues and eigenstates). *Let the perturbation* $\{\mathbb{V}, \mathbb{P}\}$ *be represented in the form* (8).

*Then* (a) *each nonzero cardinal eigenvalue* $\varepsilon_k$ *of the generalized eigenvalue equation* (6) *is a solution of the equation*

$$\mathscr{D}(\varepsilon) \equiv \begin{vmatrix} S^a_{sp}(\varepsilon) - \delta_{sp}/\omega_s & S^b_{sp}(\varepsilon) \\ S^c_{sp}(\varepsilon) & S^d_{sp}(\varepsilon) + \delta_{sp}/(\varepsilon\tau_s) \end{vmatrix} = 0, \tag{11}$$

*where* $\mathscr{D}(\varepsilon)$ *is an* $(m + \mu)$*-order determinant and where the matrix elements of matrices* $S^a(\varepsilon)$, $S^b(\varepsilon)$, $S^c(\varepsilon)$ *and* $S^d(\varepsilon)$ *are*

$$S^a_{sp}(\varepsilon) = \sum_{i=1}^{n} \frac{\langle v_s|\Phi^R_i\rangle\langle\Phi^L_i|u_p\rangle}{\varepsilon - \lambda_i}, \qquad s, p = 1, \ldots, m, \tag{12a}$$

$$S^b_{sp}(\varepsilon) = \sum_{i=1}^{n} \frac{\langle v_s | \Phi_i^R \rangle \langle \Phi_i^L | x_p \rangle}{\varepsilon - \lambda_i}, \qquad s = 1, \ldots, m, \quad p = 1, \ldots, \mu, \tag{12b}$$

$$S^c_{sp}(\varepsilon) = \sum_{i=1}^{n} \frac{\langle y_s | \Phi_i^R \rangle \langle \Phi_i^L | u_p \rangle}{\varepsilon - \lambda_i}, \qquad s = 1, \ldots, \mu, \quad p = 1, \ldots, m, \tag{12c}$$

$$S^d_{sp}(\varepsilon) = \sum_{i=1}^{n} \frac{\langle y_s | \Phi_i^R \rangle \langle \Phi_i^L | x_p \rangle}{\varepsilon - \lambda_i}, \qquad s, p = 1, \ldots, \mu. \tag{12d}$$

*Conversely, each root* $\varepsilon_0 \notin \{\lambda_i\}$ *of* $\mathscr{D}(\varepsilon)$ *is a cardinal eigenvalue of the generalized eigenvalue equation* (6).

(b) *Let* $\varepsilon_0$ *be a nonzero cardinal eigenvalue of* (6). *Each right eigenstate* $\Psi^R$ *corresponding to this eigenvalue is of the form*

$$\Psi^R = \sum_{i=1}^{n} \left[ \left( \sum_{p=1}^{m} \langle \Phi_i^L | u_p \rangle \mathsf{C}_p + \sum_{p=1}^{\mu} \langle \Phi_i^L | x_p \rangle \mathsf{D}_p \right) \Big/ (\varepsilon_0 - \lambda_i) \right] | \Phi_i^R \rangle, \tag{13a}$$

*where*

$$\mathsf{C}_p = \omega_p \langle v_p | \Psi^R \rangle, \qquad \mathsf{D}_p = -\varepsilon_0 \tau_p \langle y_p | \Psi^R \rangle. \tag{13b}$$

*Moreover, the coefficients* $\mathsf{C}_p$ $(p = 1, \ldots, m)$ *and* $\mathsf{D}_p$ $(p = 1, \ldots, \mu)$ *satisfy*

$$\begin{aligned} &\sum_{p=1}^{m} [S^a_{sp}(\varepsilon_0) - \delta_{sp}/\omega_s] \mathsf{C}_p + \sum_{p=1}^{\mu} S^b_{sp}(\varepsilon_0) \mathsf{D}_p = 0, \qquad s = 1, \ldots, m, \\ &\sum_{p=1}^{m} S^c_{sp}(\varepsilon_0) \mathsf{C}_p + \sum_{p=1}^{\mu} [S^d_{sp}(\varepsilon_0) + \delta_{sp}/(\varepsilon_0 \tau_s)] \mathsf{D}_s = 0, \qquad s = 1, \ldots, \mu. \end{aligned} \tag{13c}$$

*Conversely, if* $\varepsilon_0$ *is a nonzero cardinal eigenvalue of* (6), *each state* $\Psi^R$ *of the form* (13*a*) *where the coefficients* $\mathsf{C}_p$ *and* $\mathsf{D}_p$ *are the (nontrivial) solution of the linear set* (13*c*) *is the corresponding right eigenstate. Moreover, these coefficients satisfy* (13*b*).

Besides cardinal eigenvalues and eigenstates, one has to find singular eigenvalues and eigenstates. This is provided by the following theorem.

**Theorem 2** (singular eigenvalues and eigenstates). *Let the perturbation* $\{\mathbb{V}, \mathbb{P}\}$ *be of the form* (8), *and let* $\varepsilon_0 = \lambda_k$ *be a nonzero singular eigenvalue of* (6). *Further, let* $\lambda_k$ *be a* $\nu$*-degenerate eigenvalue of the unperturbed equation* (4), *and let* $|\Phi^R_{k\varkappa}\rangle$ *and* $\langle \Phi^L_{k\varkappa}|$ $(\varkappa = 1, \ldots, \nu)$ *be the corresponding biorthonormalized right and left eigenstates, respectively. Then*

(a) *The eigenvalue* $\varepsilon_0$ *is a solution of the equation*

$$\mathscr{D}^0(\varepsilon) \equiv \left| \begin{array}{c|c|c} S^{a0}_{sp}(\varepsilon) - \delta_{sp}/\omega_s & S^{b0}_{sp}(\varepsilon) & \langle v_s | \Phi^R_{k\varkappa} \rangle \\ \hline S^{c0}_{sp}(\varepsilon) & S^{d0}_{sp}(\varepsilon) + \delta_{sp}/(\varepsilon\tau_s) & \langle y_s | \Phi^R_{k\varkappa} \rangle \\ \hline \langle \Phi^L_{k\varkappa} | u_p \rangle & \langle \Phi^L_{k\varkappa} | x_p \rangle & \mathbf{0} \end{array} \right| = 0 \tag{14}$$

*where* $\mathscr{D}^0(\varepsilon)$ *is an* $(m + \mu + \nu)$*-order determinant and where the matrix elements of matrices* $S^{a0}(\varepsilon)$, $S^{b0}(\varepsilon)$, $S^{c0}(\varepsilon)$ *and* $S^{d0}(\varepsilon)$ *are*

$$S_{sp}^{a0}(\varepsilon) = \sum_{i \neq k}^{n} \frac{\langle v_s | \Phi_i^R \rangle \langle \Phi_i^L | u_p \rangle}{\varepsilon - \lambda_i}, \qquad s, p = 1, \ldots, m, \tag{15a}$$

$$S_{sp}^{b0}(\varepsilon) = \sum_{i \neq k}^{n} \frac{\langle v_s | \Phi_i^R \rangle \langle \Phi_i^L | x_p \rangle}{\varepsilon - \lambda_i}, \qquad s = 1, \ldots, m, \quad p = 1, \ldots, \mu, \tag{15b}$$

$$S_{sp}^{c0}(\varepsilon) = \sum_{i \neq k}^{n} \frac{\langle y_s | \Phi_i^R \rangle \langle \Phi_i^L | u_p \rangle}{\varepsilon - \lambda_i}, \qquad s = 1, \ldots, \mu, \quad p = 1, \ldots, m, \tag{15c}$$

$$S_{sp}^{d0}(\varepsilon) = \sum_{i \neq k}^{n} \frac{\langle y_s | \Phi_i^R \rangle \langle \Phi_i^L | x_p \rangle}{\varepsilon - \lambda_i}, \qquad s, p = 1, \ldots, \mu. \tag{15d}$$

*Conversely, each root* $\varepsilon_0 \in \{\lambda_i\}$ *of* $\mathscr{D}^0(\varepsilon)$ *is a singular eigenvalue of the generalized eigenvalue equation* (6).

(b) *Each right eigenstate* $\Psi^R$ *corresponding to the singular eigenvalue* $\varepsilon_0 = \lambda_k$ *is of the form*

$$|\Psi^R\rangle = \sum_{i \neq k} \left[ \left( \sum_{p=1}^{m} \langle \Phi_i^L | u_p \rangle \mathsf{C}_p + \sum_{p=1}^{\mu} \langle \Phi_i^L | x_p \rangle \mathsf{D}_p \right) \Big/ (\varepsilon_0 - \lambda_i) \right] |\Phi_i^R\rangle + \sum_{\varkappa=1}^{\nu} \mathsf{E}_\varkappa |\Phi_{k\varkappa}^R\rangle, \tag{16a}$$

*where*

$$\mathsf{C}_p = \omega_p \langle v_p | \Psi^R \rangle, \qquad \mathsf{D}_p = -\varepsilon_0 \tau_p \langle y_p | \Psi^R \rangle, \qquad \mathsf{E}_\varkappa = \langle \Phi_{k\varkappa}^L | \mathbb{C} | \Psi^R \rangle. \tag{16b}$$

*Moreover, the coefficients* $\mathsf{C}_p$ $(p = 1, \ldots, m)$, $\mathsf{D}_p$ $(p = 1, \ldots, \mu)$ *and* $\mathsf{E}_\varkappa$ $(\varkappa = 1, \ldots, \nu)$ *satisfy*

$$\sum_{p=1}^{m} [S_{sp}^{a0}(\varepsilon_0) - \delta_{sp}/\omega_s] \mathsf{C}_p + \sum_{p=1}^{\mu} S_{sp}^{b0}(\varepsilon_0) \mathsf{D}_p + \sum_{\varkappa=1}^{\nu} \langle v_s | \Phi_{k\varkappa}^R \rangle \mathsf{E}_\varkappa = 0 \qquad s = 1, \ldots, m,$$

$$\sum_{p=1}^{m} S_{sp}^{c0}(\varepsilon_0) \mathsf{C}_p + \sum_{p=1}^{\mu} [S_{sp}^{d0}(\varepsilon_0) + \delta_{sp}/(\varepsilon_0 \tau_0)] \mathsf{D}_s + \sum_{\varkappa=1}^{\nu} \langle y_s | \Phi_{k\varkappa}^R \rangle \mathsf{E}_\varkappa = 0, \qquad s = 1, \ldots, \mu, \tag{16c}$$

$$\sum_{p=1}^{m} \langle \Phi_{k\varkappa} | u_p \rangle \mathsf{C}_p + \sum_{p=1}^{\mu} \langle \Phi_{k\varkappa} | x_p \rangle \mathsf{D}_p = 0, \qquad \varkappa = 1, \ldots, \nu.$$

*Conversely, if* $\varepsilon_0 = \lambda_k$ *is a nonzero singular eigenvalue of* (6), *each state* $\Psi^R$ *of the form* (16*a*) *where the coefficients* $\mathsf{C}_p$, $\mathsf{D}_p$ *and* $\mathsf{E}_\varkappa$ *are the (nontrivial) solution of the linear set* (16*c*) *is the corresponding singular eigenstate. Moreover, the coefficients* $\mathsf{C}_p$, $\mathsf{D}_p$ *and* $\mathsf{E}_\varkappa$ *satisfy* (16*b*).

The above theorems are proved in the Appendix. In the case of the simple Hermitian eigenvalue equation ($\mathbb{C} = \mathbb{1}$, $\mathbb{P} = \mathbb{0}$, $\mathbb{B}^\dagger = \mathbb{B}$, $\mathbb{V}^\dagger = \mathbb{V}$) the proof is given in [2].

For the sake of simplicity, in the above theorems only formulas for the derivation of the right eigenstates of (6) are given. Formulas for the derivation of the left eigenstates are analogous. These formulas can be obtained from the corresponding formulas for the derivation of the right eigenstates by the formal substitution $\langle v_s | \Leftrightarrow |u_s\rangle$, $\langle \Phi_i^L | \Leftrightarrow |\Phi_i^R\rangle$ and $\langle \Psi^L | \Leftrightarrow |\Psi^R\rangle$.

We have also explicitly excluded the point $\varepsilon = 0$ where many of the above expressions diverge. However, if one takes a proper limit $\varepsilon \to 0$, these formulas apply to $\varepsilon = 0$ as well. For example, in the limit $\varepsilon \to 0$ relation (11) reduces to $|S^a_{sp}(0) - \delta_{sp}/\omega_s| = 0$. Thus $\varepsilon = 0$ is a cardinal eigenvalue of (6) if and only if this relation is satisfied. Similarly all other relations can be specified for $\varepsilon = 0$.

One can reformulate the above expressions in such a way that the point $\varepsilon = 0$ is automatically included; however, this usually creates other problems. For example, (11) can be reformulated in the form

$$\mathscr{D}(\varepsilon) \equiv \begin{vmatrix} S^a_{sp}(\varepsilon) - \delta_{sp}/\omega_s & \sqrt{\varepsilon} S^b_{sp}(\varepsilon) \\ \sqrt{\varepsilon} S^c_{sp}(\varepsilon) & \varepsilon S^d_{sp}(\varepsilon) + \delta_{sp}/\tau_s \end{vmatrix} = 0, \tag{11'}$$

which includes $\varepsilon = 0$ as well. The drawback of this expression is that if $\varepsilon$ is known to be real (e.g., Hermitian matrices with positive definite $\mathbb{C}$ and $\mathbb{C} + \mathbb{P}$), then $\varepsilon^{1/2}$ is imaginary for negative $\varepsilon$, which complicates the numerical application of (11'). One can avoid this problem by yet another modification of (11); however, the resulting expression turns out to be non-symmetric. The drawback is again numerical. If all matrices are Hermitian and if matrices $\mathbb{C}$ and $\mathbb{C} + \mathbb{P}$ are in addition positive definite, the symmetric expression has a substantially lower operation count. On balance, the expressions given in the theorems seem to be the best choice, and the point $\varepsilon = 0$ can always be separately verified.

If the perturbation $\{\mathbb{V}, \mathbb{P}\}$ is not generalized ($\mathbb{P} = \mathbb{0}$), matrices $S^b(\varepsilon)$, $S^c(\varepsilon)$, $S^d(\varepsilon)$, $S^{b0}(\varepsilon)$, $S^{c0}(\varepsilon)$ and $S^{d0}(\varepsilon)$ vanish. Hence relations (11) and (14) simplify to

$$\mathscr{D}(\varepsilon) \equiv |S^a_{sp}(\varepsilon) - \delta_{sp}/\omega_s| = 0, \tag{11a}$$

$$\mathscr{D}^0(\varepsilon) \equiv \begin{vmatrix} S^{a0}_{sp}(\varepsilon) - \delta_{sp}/\omega_s & \langle v_s | \Phi^R_{k\varkappa} \rangle \\ \langle \Phi^L_{k\varkappa} | u_p \rangle & \mathbb{0} \end{vmatrix} = 0. \tag{14a}$$

The relations for the corresponding eigenstates simplify accordingly. All the relations so obtained are formally the same as the previously derived relations [2] for the simple Hermitian eigenvalue equation ($\mathbb{C} = \mathbb{1}$, $\mathbb{P} = \mathbb{0}$, $\mathbb{B}^\dagger = \mathbb{B}$, $\mathbb{V}^\dagger = \mathbb{V}$). The only difference is that for Hermitian matrices the left and right eigenstates of the unperturbed eigenvalue equation are the same. This simplifies expressions (12a) and (15a) for the LRP functions $S^a_{sp}(\varepsilon)$ and $S^{a0}_{sp}(\varepsilon)$. In spite of this formal similarity, formulas (11a) and (14a), as well as the corresponding formulas for the eigenstates, are much more general. These formulas apply to an arbitrary perturbed eigenvalue equation (6) with the only restriction that $\mathbb{P} = \mathbb{0}$.

In the case of the simple Hermitian eigenvalue equation ($\mathbb{C} = \mathbb{1}$, $\mathbb{P} = \mathbb{0}$, $\mathbb{B}^\dagger = \mathbb{B}$, $\mathbb{V}^\dagger = \mathbb{V}$), with the additional restriction that the perturbation $\mathbb{V}$ is represented symmetrically ($|u_s\rangle = |v_s\rangle$, $s = 1, \ldots, m$), formulas equivalent to (11a) and (14a) were independently derived by Arbenz and Golub [4]. Their derivation is based on the use of Silvester's law on inertia [11]. Their formulas are much less general than (11a) and (14a) except in one respect. Given some value $x$, those formulas relate the number of eigenvalues of the perturbed matrix that are greater than $x$ to the number of eigenvalues of the unperturbed matrix that are greater than $x$. From these formulas one can derive some useful inequalities for the bracketing of the eigenvalues [4]. The inequalities thus

obtained generalize well-known inequalities usually proven by means of the Courant–Weyl principle [9, 14]. Since the eigenvalues of (6) are generally complex, these inequalities cannot be generalized to apply to (6), at least not in their present form.

The application of Theorems 1 and 2 is straightforward. Theorem 1 can be used to find cardinal eigenvalues and eigenstates, whereas Theorem 2 can be used to find singular eigenvalues and eigenstates (if any).

In order to find cardinal eigenvalues one has to solve the characteristic equation $\mathscr{D}(\varepsilon) = 0$. The function $\mathscr{D}(\varepsilon)$ is a $(m + \mu)$-order determinant, and it is analytic in $\varepsilon$. Unless there is some accidental cancellation of terms [2], this determinant has a pole at each $\lambda_i$, and nowhere else. The characteristic equation $\mathscr{D}(\varepsilon)$ can be solved iteratively, and the corresponding operation count strongly depends on $(m + \mu)$. Since without loss of generality one can assume $m = \varrho_v$ and $\mu = \varrho_p$, this method of deriving cardinal eigenvalues is likely to be efficient if the rank $\varrho$ of the perturbation $\{\mathbb{V}, \mathbb{P}\}$ is small ($\varrho \ll n$).

Once the root $\varepsilon = \varepsilon_0$ of (6) is known, one finds the corresponding right eigenstate(s) using relations (13a) and (13c). Since the determinant $D(\varepsilon_0)$ vanishes, the homogeneous linear set (13c) of $(m + \mu)$ equations in $(m + \mu)$ unknowns, $\mathsf{C}_s$ and $\mathsf{D}_s$, has at least one nontrivial solution. After the coefficients $\mathsf{C}_s$ and $\mathsf{D}_s$ are determined, one obtains the corresponding right eigenstate $|\Psi^R\rangle$ by inserting these coefficients into (13a). Since, without loss of generality, one can assume $\varrho = m + \mu$, the derivation of $|\Psi^R\rangle$ should be efficient whenever $\varrho \ll n$.

Theorems 1 and 2 are convenient if the perturbation $\{\mathbb{V}, \mathbb{P}\}$ is represented in the general form (8). If this perturbation is represented in the form (9), one can instead consider Theorems 3 and 4.

**Theorem 3** (cardinal eigenvalues and eigenstates). *Let the perturbation* $\{\mathbb{V}, \mathbb{P}\}$ *be represented in the form* (9). *Then*

(a) *Each cardinal eigenvalue* $\varepsilon_k$ *of the generalized eigenvalue equation* (6) *is a solution of the equation*

$$\mathscr{D}(\varepsilon) \equiv |S_{sp}(\varepsilon) - \delta_{sp}| = 0, \tag{17a}$$

*where the function* $\mathscr{D}(\varepsilon)$ *is an* $m$*-order determinant and where the matrix elements of the matrix* $S(\varepsilon)$ *are*

$$S_{sp}(\varepsilon) = \sum_{i=1}^{n} \frac{\langle v_s | \Phi_i^R \rangle [\omega_p \langle \Phi_i^L | u_p \rangle - \varepsilon \tau_p \langle \Phi_i^L | x_p \rangle]}{\varepsilon - \lambda_i}, \qquad s, p = 1, \ldots, m. \tag{17b}$$

*Conversely, each solution* $\varepsilon_0 \notin \{\lambda_i\}$ *of the LRP equation* (17a) *is a cardinal eigenvalue of the generalized eigenvalue equation* (6).

(b) *Let* $\varepsilon_0$ *be a cardinal eigenvalue of* (6). *Each right eigenstate* $|\Psi^R\rangle$ *corresponding to this eigenvalue is of the form*

$$|\Psi^R\rangle = \sum_{i=1}^{n} \left[ \sum_{p=1}^{m} (\omega_p \langle \Phi_i^L | u_p \rangle - \varepsilon_0 \tau_p \langle \Phi_i^L | x_p \rangle) \mathsf{C}_p / (\varepsilon_0 - \lambda_i) \right] |\Phi_i^R\rangle, \tag{18a}$$

*where*

$$\mathsf{C}_p = \langle v_p | \Psi^R \rangle, \qquad p = 1, \ldots, m. \tag{18b}$$

*Moreover, the coefficients* $\mathsf{C}_p$ *satisfy*

$$\sum_{p=1}^{m} [S_{sp}(\varepsilon_0) - \delta_{sp}]\mathsf{C}_p = 0, \qquad s = 1, \ldots, m. \tag{18c}$$

*Conversely, if* $\varepsilon_0$ *is a cardinal eigenvalue of* (6), *each state* $\Psi^R$ *of the form* (18a), *where the coefficients* $\mathsf{C}_p$ *are the (nontrivial) solution of the linear set* (18c), *is the corresponding eigenstate. Moreover, these coefficients satisfy* (18b).

**Theorem 4** (singular eigenvalues and eigenstates). *Let the perturbation* $\{\mathbb{V}, \mathbb{P}\}$ *be of the form* (9), *and let* $\varepsilon_0 = \lambda_k$ *be a singular eigenvalue of* (6). *Let further* $\lambda_k$ *be a* $\nu$*-degenerate eigenvalue of the unperturbed eigenvalue equation* (4), *and let* $|\Phi^R_{k\varkappa}\rangle$ *and* $\langle\Phi^L_{k\varkappa}|$ $(\varkappa = 1, \ldots, \nu)$ *be the corresponding biorthonormalized right and left eigenstates, respectively. Then*

(a) *The eigenvalue* $\varepsilon_0$ *is a root of the LRP equation*

$$\mathscr{D}^0(\varepsilon) \equiv \begin{vmatrix} S^0_{sp}(\varepsilon) - \delta_{sp} & \langle v_s|\Phi^R_{k\varkappa}\rangle \\ \omega_p\langle\Phi^L_{k\varkappa}|u_p\rangle - \varepsilon\tau_p\langle\Phi^L_{k\varkappa}|x_p\rangle & \mathbb{0} \end{vmatrix} = 0 \tag{19}$$

*where* $\mathscr{D}^0(\varepsilon)$ *is an* $(m+\nu)$*-order determinant and where*

$$S^0_{sp}(\varepsilon) = \sum_{i\neq k} \frac{\langle v_s|\Phi^R_i\rangle[\omega_p\langle\Phi^L_i|u_p\rangle - \varepsilon\tau_p\langle\Phi^L_i|x_p\rangle]}{\varepsilon - \lambda_i}, \qquad s, p = 1, \ldots, m. \tag{20}$$

*Conversely, each root* $\varepsilon_0 \in \{\lambda_i\}$ *of the LRP equation* (19) *is a singular eigenvalue of the generalized eigenvalue equation* (6).

(b) *Each right eigenstate* $\Psi^R$ *of* (6) *which corresponds to the singular eigenvalue* $\varepsilon_0 = \lambda_k$ *is of the form*

$$|\Psi^R\rangle = \sum_{i\neq k}\left[\sum_{p=1}^{m} (\omega_p\langle\Phi^L_i|u_p\rangle - \varepsilon_0\tau_p\langle\Phi^L_i|x_p\rangle)\mathsf{C}_p/(\varepsilon_0 - \lambda_i)\right]|\Phi^R_i\rangle + \sum_{\varkappa=1}^{\nu} \mathsf{D}_\varkappa|\Phi^R_{k\varkappa}\rangle, \tag{21a}$$

*where*

$$\mathsf{C}_p = \langle v_p|\Psi^R\rangle, \qquad \mathsf{D}_\varkappa = \langle\Phi^L_{k\varkappa}|\mathbb{C}|\Psi^R\rangle. \tag{21b}$$

*Moreover, the coefficients* $\mathsf{C}_p$ $(p = 1, \ldots, m)$ *and* $\mathsf{D}_\varkappa$ $(\varkappa = 1, \ldots, \nu)$ *satisfy*

$$\begin{aligned} &\sum_{p=1}^{m} [S_{sp}(\varepsilon_0) - \delta_{sp}]\mathsf{C}_p + \sum_{\varkappa=1}^{\nu} \langle v_s|\Phi^R_{k\varkappa}\rangle\mathsf{D}_\varkappa = 0, \qquad s = 1, \ldots, m, \\ &\sum_{p=1}^{m} [\omega_p\langle\Phi^L_{k\varkappa}|u_p\rangle - \varepsilon_0\tau_p\langle\Phi^L_{k\varkappa}|x_p\rangle]\mathsf{C}_p = 0, \qquad \varkappa = 1, \ldots, \nu \end{aligned} \tag{21c}$$

*Conversely, if* $\varepsilon_0 = \lambda_k$ *is a singular eigenvalue, each state* $\Psi^R$ *of the form* (21a), *where the coefficients* $\mathsf{C}_p$ *and* $\mathsf{D}_p$ *are the (nontrivial) solution of* (21c), *is the corresponding eigenstate. Moreover, the coefficients* $\mathsf{C}_p$ *and* $\mathsf{D}_p$ *satisfy* (21b).

Theorems 3 and 4 are analogous to Theorems 1 and 2, respectively. Note that without loss of generality one can assume $m = \varrho_0 \leqslant \varrho$.

We now derive two lemmas which give some upper and lower bounds on the degeneracy of cardinal and singular eigenvalue, respectively. These lemmas can sometimes help to find eigenvalues and eigenstates of (6). Note first that if $\varepsilon_0$ is cardinal, the homogeneous linear set (18c) has at least one nontrivial solution. However, this set can generally have $\zeta \geqslant 1$ nontrivial solutions. Relation (18a) thus yields $\zeta$ eigenstates, one eigenstate for each nontrivial solution of (18c). All these eigenstates correspond to the same eigenvalue $\varepsilon_0$, and unless they are linearly dependent, $\varepsilon_0$ is $\zeta$-degenerate. Conversely, if $\varepsilon_0$ is cardinal and if it is $\zeta$-degenerate, the system (18c) should have $\zeta$ nontrivial solutions. However, this sytem can have at most $m$ nontrivial solutions. Since without loss of generality one can assume $m = \varrho_0$, this proves

**Lemma 1.** *Each cardinal eigenvalue is at most $\varrho_0$-degenerate.*

This lemma gives an upper bound to the possible degeneracy of the cardinal solutions. A similar lemma concerning the lower bound of singular solutions can be derived.

According to Theorem 4, a necessary and a sufficient condition for the eigenvalue $\lambda_k$ of the unperturbed system to coincide with some eigenvalue $\varepsilon_0$ of the perturbed system is the vanishing of the determinant $\mathscr{D}^0(\varepsilon_0)$. An equivalent requirement is that the set (21c) should have a nontrivial solution.

Let $m = \varrho_0$, and assume farther $\mathsf{C}_p = 0$ $(p = 1, \ldots, \varrho_v)$. This reduces relations (21c) to $\sum_\varkappa \langle v_s | \Phi^R_{k\varkappa} \rangle \mathsf{D}_\varkappa = 0$ $(s = 1, \ldots, \varrho_0)$. This is a homogeneous linear set of $\varrho_0$ equations in $\nu$ unknowns. If $\nu > \varrho_0$, this set has at least $(\nu - \varrho_0)$ nontrivial linearly independent solutions. Since all the $\mathsf{C}_s$ are zero, the corresponding eigenstates $\Psi^R$ are, according to (21a), normal singular. This proves

**Lemma 2.** *Let the eigenvalue $\lambda_k$ of the unperturbed system be $\nu$-degenerate. If $\varrho_0 < \nu$, then the perturbed system has at least $(\nu - \varrho_0)$ normal singular eigenstates corresponding to the singular eigenvalue $\varepsilon_0 = \lambda_k$.*

Note that $\varrho_0$ is by definition the smallest $m$ which can be obtained in the representation (9) of the perturbation $\{\mathbb{V}, \mathbb{P}\}$, and that $\varrho_0 \leqslant \varrho \equiv \varrho_v + \varrho_p$.

## 3. Numerical considerations

It is relatively easy to determine any eigenvalue and eigenstate of the generalized eigenvalue equation (6) by the LRP method. An efficient algorithm for the solution of the nongeneralized Hermitian eigenvalue equation $(\mathbb{C} = \mathbb{1}, \mathbb{P} = \mathbb{0}, \mathbb{B}^\dagger = \mathbb{B}, \mathbb{V}^\dagger = \mathbb{V})$ was described and implemented elsewhere [2]. Essentially the same algorithm can be applied to the generalized eigenvalue equation.

Consider for example the derivation of the cardinal eigenvalues of (6) using Theorem 1. Due to (10) one can without loss of generality assume $m = \varrho_v$ and $\mu = \varrho_p$. Cardinal eigenvalues are roots of (11), and this equation can be solved iteratively using as the initial value some approximate root of $\mathscr{D}(\varepsilon)$. The iteration can be performed by the modified [2] Brent method [13]. This method combines

the sureness of bisection with the efficiency of inverse quadratic interpolation [13]. One first calculates $2\varrho n$ scalar products $\langle v_s|\Phi_i^R\rangle$, $\langle y_s|\Phi_i^R\rangle$, $\langle\Phi_i^L|u_s\rangle$ and $\langle\Phi_i^L|x_s\rangle$. This requries $2\varrho ln$ multiplications. If the perturbation $\{\mathbb{V}, \mathbb{P}\}$ is piecewise local ($l \ll n$) this is $\mathcal{O}(n)$. If the perturbation is global ($l \approx n$) and low rank ($\varrho \ll n$), this is $\mathcal{O}(n^2)$. This operation count is calculated only once, irrespective of how many eigenvalues $\varepsilon_k$ are required.

In the simple iterative approach [2] each iteration step requires the calculation of $\varrho^2$ LRP functions $S_{sp}^a(\varepsilon)$, $S_{sp}^b(\varepsilon)$, $S_{sp}^c(\varepsilon)$ and $S_{sp}^d(\varepsilon)$, followed by the calculation of the determinant $\mathcal{D}(\varepsilon)$. First one forms $n$ inverses $(\varepsilon - \lambda_i)^{-1}$ ($n$ divisions), and then one forms $\varrho n$ products $\langle\Phi_i|u_p\rangle(\varepsilon - \lambda_i)^{-1}$ and $\langle\Phi_i|x_p\rangle(\varepsilon - \lambda_i)^{-1}$ ($\varrho n$ multiplications). Finally one forms LRP functions. This last step requires $n$ multiplications and $(n-1)$ additions per function. The determinant $\mathcal{D}(\varepsilon)$ can be calculated, for example, by the Gaussian elimination method which requires $\approx\varrho^3/3$ operations [13]. Excluding additions, this sums to approximately $[n(1+\varrho+\varrho^2)+\varrho^3/3]$ operations per iteration.

The iteration algorithm can be formulated in such a way that in the first iteration step one calculates the determinant $\mathcal{D}(\varepsilon)$ at two initial points, then each additional iteration requires the calculation of $\mathcal{D}(\varepsilon)$ at only one point [2]. Hence if $\eta$ roots of (11) are needed, and if on the average $\boldsymbol{I}$ iterations per root are performed, the total operation count is

$$\boldsymbol{M}_1(\eta) \approx 2n\varrho l + n(\boldsymbol{I}+1)\eta(1+\varrho+\varrho^2) + (\boldsymbol{I}+1)\eta\varrho^3/3. \tag{22a}$$

Considered as a function of $n$, and if all cardinal eigenvalues of (6) are required ($\eta \approx n$), this is $\mathcal{O}(n^2)$. If only few eigenvalues are required ($\eta \ll n$) and if in addition the perturbation $\{\mathbb{V}, \mathbb{P}\}$ is piecewise local ($l \ll n$), this is $\mathcal{O}(n)$.

The above iteration method can be improved [2]. The most time-consuming step is the calculation of $\varrho^2$ LRP functions (12). This step is repeated $(\boldsymbol{I}+1)$ times. The calculation of each of these functions is $\mathcal{O}(n)$. The main idea behind the improved algorithm is to replace the exact $\mathcal{O}(n)$ calculation of these functions with the approximate $\mathcal{O}(1)$ calculation of these functions. If the eigenstates are real, this can be done in the following way.

Let the eigenstate $\varepsilon_0$ be bracketed in the interval $\boldsymbol{P}$, or equivalently, let us be interested only in the eigenstates contained in $\boldsymbol{P}$. Assume that the dimension of $\boldsymbol{P}$ approximately equals the average distance between the adjacent eigenvalues $\lambda_i$ of the unperturbed system. Beside interval $\boldsymbol{P}$, consider an interval $\boldsymbol{R}$, which is a few times larger than $\boldsymbol{P}$, and which completely includes $\boldsymbol{P}$. Let $\boldsymbol{R}$ contain $n_{\boldsymbol{R}} \ll n$ eigenvalues $\lambda_i$. For big enough $n$, this can be usually satisfied. Express each LRP function (12) as a sum of two components: the "local" component which contains $n_{\boldsymbol{R}}$ terms with $\lambda_i \in \boldsymbol{R}$, and the "global" component which contains remaining $(n - n_{\boldsymbol{R}})$ terms with $\lambda_i \notin \boldsymbol{R}$. Then expand the global component in a Taylor series around some point which is currently a best approximation of $\varepsilon_0$, or around a midpoint of $\boldsymbol{P}$, and truncate this expansion at $d$ terms.

The resulting functions are approximate LRP functions. The terms which are most sensitive to small changes in $\varepsilon$ are contained in the local components which are calculated exactly, whereas all the remaining terms are contained in global components which are approximated with the first $d$ terms of the Taylor

expansion. Since $\boldsymbol{P}$ is by assumption a few times smaller than $\boldsymbol{R}$, the Taylor expansion converges quickly. If the quantities $n_{\boldsymbol{R}}$ and $d$ are well chosen, approximate LRP functions are reliable.

The calculation of the Taylor expansion coefficients is $\mathcal{O}(n)$ per function, but this should be done only once. Once these coefficients are known, the calculation of approximate LRP functions for each particular value of $\varepsilon$ is $\mathcal{O}(1)$. The solution of the LRP equation (11) is replaced with the solution of the approximate LRP equation, $\mathscr{D}'(\varepsilon) = 0$. Each iterative step in the solution of the approximate equation requires only $\mathcal{O}(1)$ operations per LRP function, which is negligible for big $n$. If the parameters $n_{\boldsymbol{R}}$ and $d$ are well chosen, the root $\varepsilon_0'$ of the approximate determinant $\mathscr{D}'(\varepsilon)$ is a good approximation of the corresponding root $\varepsilon_0$ of $\mathscr{D}(\varepsilon)$. In the case of simple Hermitian eigenvalue equations with random matrices $\mathbb{B}$ and $\mathbb{V}$ and with matrix elements of $\mathbb{V}$ approximately of the same order of magnitude as the matrix elements of $\mathbb{B}$, acceptable results were obtained with $n_{\boldsymbol{R}} = 4$ and $d = 3$ [2]. With this choice each root $\varepsilon_0'$ of $\mathscr{D}'(\varepsilon)$ usually approximates the corresponding eigenvalue $\varepsilon_0$ to five or more significant figures [2].

If the above process is repeated with the root $\varepsilon_0'$ being considered as a new initial approximation, a second set of approximate LRP functions is constructed, and a second approximate equation $\mathscr{D}''(\varepsilon) = 0$ is iteratively solved. Since $\varepsilon_0'$ is already very close to $\varepsilon_0$, the corresponding root $\varepsilon_0''$ of $\mathscr{D}''(\varepsilon)$ is practically exact. If the calculation is performed on a PC computer in double precision, root $\varepsilon_0''$ of $\mathscr{D}''(\varepsilon)$ approximates the eigenvalue $\varepsilon_0$ up to $\approx 15$ significant figures [2]. This is the maximum number of significant figures which one can hope for in double precision on a PC.

This improved algorithm "stabilizes" and lowers the operation count. Both advantages are due to the fact that the $\mathcal{O}(n)$ calculation of LRP functions is performed only twice, and not $(\boldsymbol{I}+1)$ times as in the simple iteration. With appropriate modification, this algorithm can be also applied to the calculation of complex eigenvalues.

The operation count for the derivation of $\eta$ roots of $\mathscr{D}(\varepsilon)$ using this improved algorithm can be estimated to be [2]

$$\boldsymbol{M}_1'(\eta) \approx 2n\varrho l + 2n\eta\varrho[1 + \varrho d] + \boldsymbol{I}\eta\varrho[n_{\boldsymbol{R}} + d + \varrho(n_{\boldsymbol{R}} - d) + \varrho^2/3], \qquad (22b)$$

where the quantities $d$ and $n_{\boldsymbol{R}}$ determine how well the LRP functions are approximated. The larger these quantities, the better the approximation. As stated above, if the matrices are random, good results are obtained with $d = 3$ and $n_{\boldsymbol{R}} = 4$. If the perturbation is unusually large, or if the eigenvalues $\lambda_i$ are distributed densely in the vicinity of $\varepsilon_0$, one should increase $n_{\boldsymbol{R}}$, but not $d$. If only $n_{\boldsymbol{R}}$ is increased, the two leading terms in (22b) are not affected.

Comparison of (22b) and (22a) demonstrates the advantage of the (22b) algorithm. Consider for example the calculation of a single eigenvalue ($\eta = 1$), and assume that the perturbation is piecewise local ($l \ll n$). In both expressions the second term is $\mathcal{O}(n)$, and this term substantially contributes to the operation count. In (22a) this term is very sensitive to the number of iterations $\boldsymbol{I}$, which is a highly variable and unpredictable quantity. In (22b) the dependence on $\boldsymbol{I}$ is confined to the third term which is $\mathcal{O}(1)$. In the limit of big $n$ the operation count

for (22b) is stable, whereas the operation count for (22a) is not. Moreover, assuming conservatively $I > 5$ and using $d = 3$ [2], one finds in the big $n$ limit $M_1'(\eta) < M_1(\eta)$. The relative advantage of the operation count $M_1'(\eta)$ further increases with the increase of $I$.

If the cardinal eigenvalue $\varepsilon_0$ of (6) is known, the derivation of the corresponding eigenstate(s) is straightforward. One first solves (13c). This can be done by Gaussian elimination followed by backsubstitution, which requires approxiamtely $\varrho^3/3 + \varrho^2/2 \approx \varrho^3/3$ operations [13]. If the eigenvalue $\varepsilon_0$ has been obtained as the root of $\mathscr{D}(\varepsilon)$, the LRP functions $S^a_{sp}(\varepsilon_0)$, $S^b_{sp}(\varepsilon_0)$, $S^c_{sp}(\varepsilon_0)$ and $S^d_{sp}(\varepsilon_0)$ are already known, and there is no overhead for their derivation. If this eigenvalue has been obtained in some other way, there is an initial overhead of $\approx 2\varrho nl$ operations and an additional overhead of $(1 + \varrho + \varrho^2)n$ operations per eigenstate.

Once (13c) is solved, one inserts the coefficients $\mathbf{C}_s$ and $\mathbf{D}_s$ in (13a). This requires $\approx(\varrho + 1)n$ operations. In order to derive $\eta$ eigenstates the total operation count is

$$M_2(\eta) \approx n\eta(\varrho + 1) + \eta\varrho^3/3, \tag{23}$$

if the corresponding eigenvalues were obtained by the LRP method and

$$M_3(\eta) \approx 2n\varrho l + n\eta(2 + 2\varrho + \varrho^2) + \eta\varrho^3/3, \tag{24}$$

if these eigenvalues have been obtained in some other way. Considered as a function of $n$, and if all the eigenstates are required ($\eta \approx n$), both operation counts are $\mathcal{O}(n^2)$. If only a few eigenstates are required ($\eta \ll n$), the operation count (23) is $\mathcal{O}(n)$, whereas the operation count (24), is either $\mathcal{O}(n)$ or $\mathcal{O}(n^2)$, depending on whether or not the perturbation is piecewise local.

Presently the most efficient method for the derivation of all eigenvalues or all eigenvalues and all eigenstates of a real symmetric matrix is a Householder tridiagonalization followed by the QL algorithm [13]. The operation count for the derivation of all eigenvalues of an $n$-order matrix by this method is [13]

$$H_1(n) \approx \tfrac{2}{3}n^3 + 30n^2, \tag{25}$$

whereas the operation count for the derivation of all eigenvalues and all eigenstates is [13]

$$H_2(n) \approx \tfrac{13}{3}n^3 + 30n^2. \tag{26}$$

Both operation counts are $\mathcal{O}(n^3)$, whereas the LRP operation counts are $\mathcal{O}(n^2)$ for $\eta \approx n$. For each fixed rank $\varrho$ of the perturbation $\{\mathbb{V}, \mathbb{P}\}$, there is some critical $n$ such that the LRP method is more efficient whenever this critical value is exceeded. The relative advantage of the LRP method further increases with the increase of $n$. If only a few eigenvalues or only a few eigenstates are required, the relative advantage of the LRP method is substantially greater, especially if the perturbation is peicewise local.

The estimated LRP operation counts may decrease in some special cases. For example, if matrices $\mathbb{B}$, $\mathbb{C}$, $\mathbb{V}$ and $\mathbb{P}$ are Hermitian then $|\Psi^R_k\rangle = |\Psi^L_k\rangle$ $(k = 1, \ldots, n)$. If in addition $\mathbb{C}$ and $\mathbb{C} + \mathbb{P}$ are positive definite, then the eigenvalues $\varepsilon_k$ and $\lambda_i$ are real. Moreover, since $\mathbb{V}^\dagger = \mathbb{V}$ and $\mathbb{P}^\dagger = \mathbb{P}$, one can

choose $|u_s\rangle = |v_s\rangle$ $(s = 1, \ldots, m)$ and $|x_s\rangle = |y_s\rangle$ $(s = 1, \ldots, \mu)$. Hence for real $\varepsilon$

$$S^a_{sp} = (S^a_{ps})^*, \qquad S^b_{sp} = (S^c_{ps})^*, \qquad S^d_{sp} = (S^d_{ps})^*. \tag{27}$$

Matrix $\|\mathscr{D}(\varepsilon)\|$ is thus Hermitian, and instead of $\varrho^2$ distinct LRP functions one has to calculate only $\varrho(\varrho + 1)/2$ such functions. In addition, instead of $2\varrho n$ scalar products, one has to calculate only $\varrho n$ scalar products $\langle u_s|\Phi_i\rangle$ and $\langle x_s|\Phi_i\rangle$. All the operation counts accordingly decrease.

The above estimates for the operation counts are based on Theorem 1, and they apply to the perturbation $\{\mathbb{V}, \mathbb{P}\}$ represented in the general form (8). If this perturbation is represented column-wise as in (9), one can use Theorem 3 instead of Theorem 1. Without loss of generality one can assume $m = \varrho_0$. The corresponding operation counts are analogous to the operation counts (22), (23) and (24). The most important difference is that rank $\varrho = (\varrho_v + \varrho_p)$ is replaced by $\varrho_0$. This replaces a $\varrho$-order LRP determinant with a $\varrho_0$-order LRP determinant. Since $\varrho_0 \leqslant \varrho$, the operation counts can be substantially decreased. Sometimes $\varrho_0$ can be as low as $\varrho_0 \approx \varrho/2$, which can decrease the operation counts up to fourfold. On the other hand, if relations (27) are satisfied, the relative performance of Theorem 1 improves, and the corresponding operation counts decrease approximately twofold. Which theorem produces a faster algorithm depends on a delicate trade off.

Similar conclusions apply to the derivation of singular eigenstates, which can be derived either with Theorem 2 or with Theorem 4.

## 4. Conclusion

We have considered the derivation of the eigenvalues and eigenstates of the generalized perturbed eigenvalue equation $(\mathbb{B} + \mathbb{V})\Psi = \varepsilon(\mathbb{C} + \mathbb{P})\Psi$, where the eigenvalues and the eigenstates of the related unperturbed eigenvalue equation $\mathbb{B}\Phi = \lambda\mathbb{C}\Phi$ are known. The only restriction imposed on matrices $\mathbb{B}$, $\mathbb{C}$, $\mathbb{V}$ and $\mathbb{P}$ is that matrix $\mathbb{C}$ be nonsingular, and that the eigenstates of the unperturbed eigenvalue equation form a complete set. This later condition can be replaced with the condition that matrix $\mathbb{C}^{-1/2}\mathbb{B}\mathbb{C}^{-1/2}$ be non-defective. These conditions are very mild, and the low rank perturbation (LRP) method applies to virtually arbitrary generalized eigenvalue equation with arbitrary generalized perturbation $\{\mathbb{V}, \mathbb{P}\}$.

For each fixed rank $\varrho \equiv \varrho_v + \varrho_p$ of the perturbation $\{\mathbb{V}, \mathbb{P}\}$, the operation count for the derivation of all eigenvalues and all eigenstates scales as $\mathcal{O}(n^2)$. If the perturbation $\{\mathbb{V}, \mathbb{P}\}$ is piecewise local, the operation count for the derivation of a single eigenvalue and a single eigenstate scales as $\mathcal{O}(n)$. The LRP method is hence potentially very useful for the solution of generalized eigenvalue problems involving low rank perturbations.

Besides the operation count, another important aspect of each numerical method is the stability of the method. In the case of Hermitian matrices the LRP method is stable and numerically reliable [2]. In the case of non-Hermitian

matrices this is not necessarily so. It is well known that the eigenvalue equations involving some non-Hermitian matrices are genuinely unstable, especially if the eigenstates are almost linearly dependent [13]. No numerical method can remedy a natural instability. At most one can hope for is that the LRP method does not introduce additional instabilities. Numerical experiments involving non-Hermitian matrices are in progress.

The LRP method and the standard perturbation expansion method complement each other. These methods address qualitatively different problems. In order for the perturbation expansion to be efficient, the perturbation should be small. In order for the LRP method to be efficient, the rank of the perturbation should be small. In addition, the perturbation expansion is a power series expansion which is not always guaranteed to converge, whereas the LRP method always produces correct eigenvalues and eigenstates, albeit sometimes (for large rank) with a large operation count.

One potential application of the LRP method is the treatment of large molecular systems perturbed by local or piecewise local perturbations. For example, if a solution to some hydrocarbon system is known, the LRP method could be used to derive solutions for related heterocompounds. Depending on the particular model (Hückel, PPP, *ab initio*, etc.), such an approach could be much faster than any presently known method. In a similar way the LRP method can be applied to numerical problems associated with breaking and forming of molecular bonds, local interactions between large molecules, local interactions between a single large molecule and small molecules or atoms, etc. In each particular case the feasibility of the LRP approach depends on the rank of the perturbation $\{\mathbb{V}, \mathbb{P}\}$. This rank should be "small" with respect to the dimension $n$ of the problem considered. How small depends on whether all or only few eigenstates and eigenvalues are requried, and it can be estimated from the quantitative expressions for operation counts.

Another potentially useful area of application is model testing. For example, in the Hückel approach one associates with each heteroatom X a Coulomb integral $\alpha_X$. One might ask how the change in the particular parameter $\alpha_X$ affects eigenvalues and eigenstates. This question can be answered using the standard perturbation approach. However, the perturbation method is inefficient if the change in $\alpha_X$ is substantial, and besides it gives only an approximate result. Direct methods, such as the Householder, Givens or Jacobi methods, have operation counts of $\mathcal{O}(n^3)$ [13]. The same is true for iterative methods, such as the power method or inverse iteration. The LRP method is fast (the rank of this perturbation is one), it is at most $\mathcal{O}(n^2)$, and it produces exact eigenstates and eigenvalue. Similarly, Hückel parameters $\beta_{xy}$ involving different pairs of atoms could be investigated, since the rank of the associated perturbation is again small, $\varrho = 2$. The same idea can be applied to more sophisticated semiempirical models, such as PPP. In this way one could use the LRP approach in order to find the "best" set of parameters for a given semiempirical model.

The same idea can be applied to SCF *ab initio* models. The "parameters" in an *ab initio* model are fixed; however, the basis set of functions is not. In order to assess how reliable a given *ab initio* model is, it is important to know how

much the predictions of the model change with changes in the basis set. The inclusion of additional functions amounts to the improvement of the model, and in view of the enormous numerical complexity of the *ab initio* treatment of not so large molecules, it is not easy to find out what is the effect of the addition of a particular function to a basis set. This problem is suitable for the LRP method. In the SCF approach the inclusion of a single basis function is equivalent to a perturbation with two distinct components, a dominant rank-one component which can be attributed to the inclusion of the new function, and another much smaller component which is due to the SCF readjustment of the system. The dominant rank-one component can be treated with the LRP method, and the remaining small component can be efficiently treated with a standard perturbation method. Such a combination of the LRP and the standard method could be very efficient in assessing the relative importance of various basis functions.

As a final example consider molecular vibrational spectra. The corresponding eigenvalue equation involves two matrices, a force field matrix $\mathbb{F}$ which is related to force constants (stretching, bending, etc), and a matrix $\mathbb{G}$ which is related to the molecular kinetic energy. Replacement of a single atom, change in the value of a particular force constant, change in the mass of the particular atom (isotope effect) etc. are all examples of local perturbations. Hence all these perturbations can be efficiently treated with the LRP method.

The relative efficiency of the LRP method is mainly due to the fact that this method replaces the solution of an $n$-order linear system with the solution of a $\varrho$- or $\varrho_0$-order nonlinear system. Though nonlinear systems are more difficult to solve, the trade-off is numerically favorable whenever $\varrho$ and $\varrho_0$ are sufficiently small with respect to $n$.

## 5. Appendix

We explicitly prove only Theorem 3. Except for much more involved notation, the proof of the other three theorems is analogous.

Let the perturbation $\{\mathbb{V}, \mathbb{P}\}$ be expressed in the form (9). Further let $\varepsilon_0 \notin \{\lambda_i\}$ be a cardinal eigenvalue of (6), and let $|\Psi^R\rangle$ be the corresponding right eigenstate. Substituting (9) in (6a), taking the scalar product of (6a) with $\langle \Phi_j^L|$ and using (4) one obtains

$$(\varepsilon_0 - \lambda_i)\langle \Phi_i^L|\mathbb{C}|\Psi^R\rangle = \sum_{p=1}^{m} [\omega_p \langle \Phi_i^L|u_p\rangle - \varepsilon_k \tau_p \langle \Phi_i^L|x_p\rangle]\langle v_p|\Psi^R\rangle. \quad \text{(A1)}$$

From the completeness of the unperturbed eigenstates $|\Phi_i^R\rangle$ and from (5) one derives the identity

$$\sum_{i=1}^{n} |\Phi_i^R\rangle\langle \Phi_i^L|\mathbb{C} = \mathbb{1}. \quad \text{(A2)}$$

Since $\varepsilon_0 \notin \{\lambda_i\}$, one can divide (A1) by $(\varepsilon_0 - \lambda_i)$. Multiplying the obtained

expression by $\langle v_s|\Phi_i^R\rangle$, summing over $i$ and using (A2) one obtains

$$\sum_{p=1}^{m}[S_{sp}(\varepsilon_0)-\delta_{sp}]\langle v_p|\Psi^R\rangle=0, \qquad s=1,\ldots,m, \tag{A3}$$

where the elements $S_{sp}(\varepsilon_0)$ are given by (17b) with $\varepsilon=\varepsilon_0$. At least one of the coefficients $\langle v_p|\Psi^R\rangle$ $(p=1,\ldots,m)$ in (A3) is nonzero. To see this, assume that all these coefficients are zero. Then (9) would imply $\mathbb{V}|\Psi^R\rangle=\mathbb{P}|\Psi^R\rangle=0$ and hence from (6a) $\varepsilon_0\in\{\lambda_i\}$, contrary to the assumption that $\varepsilon_0$ is cardinal. In conclusion, relations (A3), considered as a set of $m$ homogeneous linear equations in $m$ unknowns $\langle v_p|\Psi^R\rangle$ $(p=1,\ldots,m)$, have a nontrivial solutions. However, a homogeneous linear system can have a nontrivial solution if and only if the determinant of the system vanishes. Hence $\varepsilon_0$ satisfies (17a). Again dividing (A1) by $(\varepsilon_0-\lambda_i)$, multiplying by $|\Phi_i^R\rangle$ and using (A2) one derives (18a) where the coefficients $\mathsf{C}_p$ are given by (18b). This proves the first part of the above theorem namely that each cardinal eigenvalue $\varepsilon_0$ of (6) is a root of the LRP equation (17a) and each corresponding right eigenstate $\Psi^R$ is of the form (18a) where the coefficients $\mathsf{C}_p$ are given by (18b).

Now let $\varepsilon_0\notin\{\lambda_i\}$ be a root of the LRP equation (17a). This implies that a homogeneous linear system (18c) has at least one nontrivial solution $\{\mathsf{C}_p\}$. Consider the state $\Psi^R$ given by (18a), where the $\mathsf{C}_p$ are the solutions of (18c). We will show that this state is nonzero and that it is a right eigenstate of (6). Multiplying (18a) by $\langle v_s|$ and rearranging one obtains

$$\sum_{p=1}^{m}S_{sp}(\varepsilon_0)\mathsf{C}_p=\langle v_s|\Psi^R\rangle. \tag{A4}$$

Since the $\mathsf{C}_p$ also satisfy (18c), this implies $\mathsf{C}_p=\langle v_p|\Psi^R\rangle$. Since the $\mathsf{C}_p$ are nontrivial, the state $\Psi^R$ is nonzero.

One may now verify that $\Psi^R$ satisfies (6a) with the eigenvalue $\varepsilon_0$. Multiplying (18a) by $(\varepsilon_0-\lambda_j)\langle\Phi_j^L|\mathbb{C}$ using (5) and rearranging, one finds $\langle\Phi_j^L|\mathbb{B}+\mathbb{V}|\Psi^R\rangle=\varepsilon_0\langle\Phi_j^L|\mathbb{C}+\mathbb{P}|\Psi^R\rangle$. Since the $\langle\Psi_j^L|$ are complete, this proves that $\Psi^R$ is the right eigenstate of (6) with the eigenvalue $\varepsilon_0$.

*Note.* The condition that the unperturbed eigenstates $|\Phi_i^R\rangle$ be complete, and that (5) be satisfied can be expressed in many different ways. Thus if $\mathbb{C}$ is non-singular, (4) is equivalent to $\mathbb{B}'|\chi_i^R\rangle=\lambda_i|\chi_i^R\rangle$ and $\langle\chi_i^L|\mathbb{B}'=\lambda_i\langle\chi_i^L|$, where $\mathbb{B}'=\mathbb{C}^{-1/2}\mathbb{B}\mathbb{C}^{-1/2}$, $|\chi_i^R\rangle=\mathbb{C}^{1/2}|\Phi_i^R\rangle$ and $\langle\chi_i^L|=\langle\Phi_i^L|\mathbb{C}^{1/2}$. The completeness of $|\chi_i^R\rangle$ implies the completeness of $|\Phi_i^R\rangle$, and vice versa. By definition, $\mathbb{B}'$ is non-defective if the $|\chi_i^R\rangle$ are complete [13]. However. left and right eigenstates of a non-defective matrix can be always biorthonormalized, $\langle\chi_i^L|\chi_j^R\rangle=\delta_{ij}$ [13]. This implies (5). Hence the condition of the above theorem can be satisfied whenever $\mathbb{C}$ is non-singular and $|\Phi_i^R\rangle$ are complete, or equivalently, whenever $\mathbb{C}$ is non-singular and $\mathbb{B}'$ is non-defective.

The proofs of Theorems 1, 2 and 4 are analogous. In the case of Theorems 2 and 4, the identity (A2) should be expressed in the form

$$\sum_{i\neq k}|\Phi_i^R\rangle\langle\Phi_i^L|\mathbb{C}+\sum_{\varkappa=1}^{\nu}|\Phi_{k\varkappa}^R\rangle\langle\Phi_{k\varkappa}^L|\mathbb{C}=\mathbb{1}. \tag{A5}$$

## References


1. Wilson EB, Decius JC, Cross PC (1955) Molecular vibrations. The theory of infrared and Raman vibrational spectra. McGraw-Hill, New York
2. Živković TP; submitted to SIAM
3. Beattie C, Fox D (February 1987) Technical Report UMSI 87/11, University of Minnesota Supercomputer Institute
4. Arbenz P, Golub GH (1988) SIAM J Matrix Anal Appl 9:40–58
5. Simpson A (1974) Quart J Mech Appl Math 27:27–43
6. Sehmi NS (1985) J Sound Vibration 100:409–421
7. Simpson A (1980) Aeronaut J 84:417–433
8. Arbenz P, Gander W, Golub GH (January 1987) Research Report No. 87-01, Seminar für Angewandte Mathematik, ETH Zurich
9. Weinstein A, Stenger W (1972) Methods of intermediate problems for eigenvalues. Academic Press, New York
10. Golub GH, Van Loan CF (1983) Matrix computations. The Johns Hopkins University Press, Baltimore
11. Parlett BN (1980) The symmetric eigenvalue problem. Prentice Hall-Englewood Cliffs, NJ
12. Gantmacher FR (1986) Matrizentheorie, 2nd edn. Springer, Berlin Heidelberg New York
13. Press WH, Flannery BP, Teukolsky SA, Vetterling WT (1986) Numerical recipes. Cambridge University Press, Cambridge
14. Weinberger HF (1974) Variational methods for eigenvalue approximation. Regional Conference Series in Applied Mathematics, vol 15. SIAM, Philadelphia