EIGENVALUES OF THE REAL GENERALIZED EIGENVALUE EQUATION PERTURBED BY A LOW-RANK PERTURBATION

Tomislav P. ŽIVKOVIĆ
Rudjer Bošković Institute, 41001 Zagreb, The Republic of Croatia

Received 6 May 1991

Abstract

The low-rank perturbation (LRP) method solves the perturbed eigenvalue equation $(\mathbf{B} + \mathbf{V})\Psi_k = \varepsilon_k(\mathbf{C} + \mathbf{P})\Psi_k$, where the eigenvalues and the eigenstates of the related unperturbed eigenvalue equation $\mathbf{B}\Phi_i = \lambda_i\mathbf{C}\Phi_i$ are known. The method is designed for arbitrary n-by-n matrices **B**, **V**, **C**, and **P**, with the only restriction that the eigenstates Φ_i of the unperturbed equation should form a complete set. We consider here a real LRP problem where all matrices are Hermitian, and where in addition matrices **C** and **(C + P)** are positive definite. These conditions guarantee reality of the eigenvalues ε_k and λ_i. In the original formulation of the LRP method, each eigenvalue ε_k is obtained iteratively, starting from some approximate eigenvalue ε_k'. If this approximate eigenvalue is not well chosen, the iteration may sometimes diverge. It is shown that in the case of a real LRP problem, this danger can be completely eliminated. If the rank ρ of the generalized perturbation {**V**, **P**} is "small" with respect to n, then one can easily bracket and hence locate to any desirable accuracy the eigenvalues ε_k $(k = 1, \ldots, n)$ of the perturbed equation. The calculation of all n eigenvalues requires $O(\rho^2 n^2)$ operations. In addition, if the perturbation {**V**, **P**} is local with the localizability $l \approx \rho$, then only $O(\rho^2 n)$ operations are required for a derivation of a single eigenvalue.

1. Introduction

The Low-Rank Perturbation (LRP) method [1] solves the generalized perturbed eigenvalue equation

$$(\mathbf{B} + \mathbf{V})\Psi_k = \varepsilon_k(\mathbf{C} + \mathbf{P})\Psi_k, \tag{1}$$

where **B**, **V**, **C**, and **P** are n-by-n matrices, and where the solution of the corresponding generalized unperturbed eigenvalue equation

$$\mathbf{B}\Phi_i = \lambda_i\mathbf{C}\Phi_i \tag{2}$$

is known. Matrices **B**, **V**, and **P** are arbitrary, with the only restriction that matrix **C** should be nonsingular and matrix $\mathbf{C}^{-1/2}\mathbf{B}\mathbf{C}^{-1/2}$ nondefective [1]. These conditions

are equivalent to the requirement that relation (2) should have a complete set of eigenfunctions.

The LRP method works for any "perturbation" {**V**, **P**}. However, the method is most efficient if the rank ρ of this perturbation is "small" with respect to n [1,2]. The operation count for the derivation of all eigenvalues and all eigenstates of (1) by the LRP method is $O(\rho^2 n^2)$. In comparison, the Householder-QR method, which is presently the best method for the derivation of all eigenvalues and all eigenstates of a real symmetric matrix [3], requires $O(n^3)$ operations [3]. If the solution of the unperturbed system is known, and if the rank ρ of the perturbation is sufficiently small with respect to n, the LRP method is by an order $O(n)$ faster.

A special case of low-rank perturbations are local and piecewise local perturbations. Such perturbations are confined to a small fraction of the initial unperturbed system described by matrices **B** and **C**. If the generalized perturbation {**V**, **P**} is local or piecewise local with the localizability l, then only $O(l^2 n)$ operations are needed for the derivation of a single eigenvalue and/or eigenstate [1]. If the localizability l is sufficiently "small", the LRP approach is almost two orders of magnitude faster than any presently known method.

The LRP method is potentially very useful for various quantum chemical problems. Physically, most perturbations frequently encountered are either local or piecewise local. For example, a formal replacement of an atom with another atom and creation or annihilation of a chemical bond are local perturbations. Also, the interaction of two molecules A and B is usually confined to a small fraction of these molecules, and hence it is a local perturbation. A piecewise local perturbation is a more general perturbation which consists of few local perturbations, such as creation and/or annihilation of a few bonds and/or substitutions of a few atoms.

In the case of simple molecular models, where no SCF readjustment of molecular orbitals is required, the LRP approach can be applied directly to all the above perturbations. In the case of various SCF-type models, such perturbations can be represented as a sum of a dominant local matrix and a small nonlocal matrix. The nonlocal contribution to the perturbation matrix is due to the SCF readjustment of molecular orbitals. A major part due to the readjustment can also be included in the dominant local matrix. The remaining nonlocal component is usually very small. Hence, one can apply the LRP method to the dominant low-rank component in an iterative SCF manner, and the remaining small component can be efficiently treated either with a standard perturbation method or in some other way.

Similar combined methods can be applied to the calculation of isoenergy surfaces. If the underlying model does not include SCF readjustment, the LRP method produces each isoenergy surface in a closed analytical form [4]. In the case of various SCF models, one can again treat the residual contribution due to the nonlocal readjustment of molecular orbitals as a small perturbation. Note finally that in its present formulation, the LRP method can not be directly applied to the CI-type calculations. Although the perturbation can be physically local, the rank of the corresponding perturbation matrix is very large in a CI formalism.

In its general formulation [1], the LRP method applies to arbitrary matrices which are not necessarily Hermitian. In particular, the eigenvalues ε_k of the eigenvalue eq. (1) can be complex. However, in quantum mechanics one almost invariably encounters the eigenvalue eq. (1), where matrices **B** and **V** are Hermitian and where matrices **C** and (**C** + **P**) are Hermitian and positive definite. For example, in a variety of molecular models matrices **B** and (**B** + **V**) represent Hamiltonians, while matrices **C** and (**C** + **P**) represent overlap matrices defined over some set of linearly independent atomic orbitals. In such a case, all matrices are Hermitian, while matrices **C** and (**C** + **P**) are in addition positive definite. This guarantees that the eigenvalues ε_k of (1) and the eigenvalues λ_i of (2) are real. We will refer to the LRP problem involving matrices with the above-stated properties as a "real" LRP problem.

A real LRP problem is mathematically a rather special case of the general LRP problem. It is hence reasonable to expect that in the case of a real LRP problem one could utilize special properties of the matrices involved, thus improving the original LRP method. This is indeed the case. In particular, the LRP calculation of the eigenvalues is relatively unstable in the case of arbitrary matrices, since the calculation of each eigenvalue ε_k requires a good choice of the initial approximate eigenvalue $\varepsilon_k' \approx \varepsilon_k$, which is then iteratively improved [1,2]. We will show that in the case of a real LRP problem, the calculation of the eigenvalues can be successfully stabilized. In particular, we will show that the number $\mathcal{N}(\mathbf{B} + \mathbf{V},\ \mathbf{C} + \mathbf{P}, \varepsilon)$ of eigenvalues of the perturbed eigenvalue eq. (1) that are $> \varepsilon$ is related to the number $\mathcal{N}(\mathbf{B}, \mathbf{C}, \varepsilon)$ of eigenvalues of the unperturbed eigenvalue eq. (2) that are $> \varepsilon$ through the relation

$$\mathcal{N}(\mathbf{B} + \mathbf{V},\ \mathbf{C} + \mathbf{P}, \varepsilon) = \mathcal{N}(\mathbf{B}, \mathbf{C}, \varepsilon) + \pi(S(\varepsilon) - N^{-1}(\varepsilon)) - \pi(-N(\varepsilon)),$$

where $\pi(\mathbf{A})$ is the number of positive eigenvalues of a matrix **A**, and where $S(\varepsilon) - N^{-1}(\varepsilon)$ and $N(\varepsilon)$ are ρ-by-ρ matrices which can be easily constructed from the known solution of the unperturbed eq. (2) and from the perturbation matrices **V** and **P**. This relation enables successful bracketing of the eigenvalues ε_k of the perturbed eigenvalue eq. (1). In combination with the fast iterative evaluation of these eigenvalues [1,2], this relation provides a fast and completely stable and reliable algorithm for the calculation of these eigenvalues.

Two points should be emphasized. First, the efficiency of the LRP method is not affected by the magnitude of the perturbation. In the standard perturbation expansion method, the magnitude of the perturbation limits the usefulness of the method. Unless the perturbation is small, the perturbation expansion is numerically not efficient, and in some cases the perturbation diverges. In the LRP method, the dominant factor is the rank of the perturbation and the efficiency of the method is not affected by its magnitude.

Another point concerns degeneracy and quasi-degeneracy of unperturbed systems. In the standard perturbation methods, degenerate and quasi-degenerate systems

require a special treatment. In the LRP method all systems, nondegenerate as well as degenerate and quasi-degenerate, are treated in a uniform way. Numerically, the LRP treatment of degenerate and quasi-degenerate systems is of the same complexity as the LRP treatment of nondegenerate systems. We consider both points in more detail in section 4.

2. The LRP method

In the LRP method, the eigenvalues λ_i and the corresponding eigenstates Φ_i of the generalized unperturbed eq (2) are assumed to be known. Without loss of generality, one can assume that the eigenstates Φ_i satisfy the generalized biorthonormalized relation [1]

$$\langle \Phi_i \,|\, \mathbf{C} \,|\, \Phi_j \rangle = \delta_{ij}. \tag{3}$$

The matrix **V** in (1) is the "perturbation" of the matrix **B**, while the matrix **P** is the "perturbation" of the matrix **C**. We refer to a pair {**V**, **P**} as a "generalized" perturbation.

Let ρ_v and ρ_p be the ranks of matrices **V** and **P**, respectively. Define the "rank" ρ of the generalized perturbation {**V**, **P**} as the sum

$$\rho = \rho_v + \rho_p. \tag{4}$$

This quantity determines the numerical efficiency of the LRP method. In order for the LRP method to be efficient, the rank ρ should be "small" with respect to n.

Perturbation matrices **V** and **P** can be written in the form

$$\mathbf{V} = \sum_{s=1}^{m} \omega_s \boldsymbol{u}_s \boldsymbol{v}_s^+, \quad \mathbf{P} = \sum_{s=1}^{\mu} \tau_s \boldsymbol{x}_s \boldsymbol{y}_s^+, \tag{5a}$$

where $\omega_s \neq 0$ and $\tau_s \neq 0$ are scalars, $\boldsymbol{u}_s$ and $\boldsymbol{x}_s$ are column vectors, while $\boldsymbol{v}_s^+$ and $\boldsymbol{y}_s^+$ are row vectors. By definition, components of the row vector $\boldsymbol{u}^+$ are the complex conjugate of the corresponding components of the column vector $\boldsymbol{u}$.

We will use the notation $\mathbf{U} = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m]$ in order to denote an n-by-m matrix whose ith column is a column vector $\boldsymbol{u}_i$. We will also use the notation $\mathbf{A} = \mathrm{diag}(a_1, \ldots, a_m)$ in order to denote an m-by-m diagonal matrix with diagonal elements a_i. In addition, we will sometimes use a bra–ket notation $\langle \boldsymbol{u} |$ and $| \boldsymbol{v} \rangle$ in order to emphasize the independence of the particular choice of the basis. For example, we will use the notation $\langle \boldsymbol{u} \,|\, \boldsymbol{v} \rangle$ in order to denote a scalar product $\boldsymbol{u}^+\boldsymbol{v}$ between vectors $\boldsymbol{u}$ and $\boldsymbol{v}$.

With the above conventions, relations (5a) can be written in the matrix notation

$$\mathbf{V} = \mathbf{U}\,\mathbf{\Omega}\,\mathbf{V}^{+}, \qquad \mathbf{P} = \mathbf{X}\,\mathbf{T}\,\mathbf{Y}^{+}, \tag{5b}$$

where $\mathbf{\Omega} = \mathrm{diag}(\omega_1, \ldots, \omega_m)$ and $\mathbf{T} = \mathrm{diag}(\tau_1, \ldots, \tau_\mu)$ are diagonal matrices, and where $\mathbf{U} = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m]$, $\mathbf{V} = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m]$, $\mathbf{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_\mu]$ and $\mathbf{Y} = [\boldsymbol{y}, \ldots, \boldsymbol{y}_\mu]$ are matrices constructed from vectors $\boldsymbol{u}_i$, $\boldsymbol{v}_i$, $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$.

In general, all vectors entering the above expressions can be linearly dependent. If the vectors $|u_s\rangle$ as well as the vectors $\langle v_s|$ are linearly independent, then $m = \rho_v$ is the rank of $\mathbf{V}$. Otherwise, $m > \rho_v$. Similarly, if the vectors $|x_s\rangle$ as well as the vectors $\langle y_s|$ are linearly independent, then $\mu = \rho_p$ is the rank of $\mathbf{P}$. Otherwise, $\mu > \rho_p$. However, any linear dependence between the vectors $|u_s\rangle$, $|v_s\rangle$, $|x_s\rangle$ or $|y_s\rangle$ can be easily eliminated [1]. Hence, one can, without loss of generality, assume $m = \rho_v$ and $\mu = \rho_p$, i.e. $\rho = m + \mu$.

The LRP method can be formulated in various ways [1]. Here, we will use the formulation which is appropriate for the intended improvement of the derivation of the eigenvalues.

In order to simplify the notation, we write all vectors which determine representation (5) of the perturbation $\{\mathbf{V}, \mathbf{P}\}$ in a compact form

$$|z_s\rangle = \begin{cases} |u_s\rangle & s = 1, \ldots, m, \\ |x_{s-m}\rangle & s = m+1, \ldots, m+\mu; \end{cases}$$

$$\langle w_s| = \begin{cases} \langle v_s| & s = 1, \ldots, m, \\ \langle y_{s-m}| & s = m+1, \ldots, m+\mu. \end{cases} \tag{6a}$$

We also consider a diagonal matrix $\mathbf{N}(\varepsilon) = \mathrm{diag}(\eta_1, \ldots, \eta_{m+\mu})$, where diagonal elements $\eta_s \equiv \eta_s(\varepsilon)$ are defined in terms of the quantities ω_s and τ_s

$$\eta_s(\varepsilon) = \begin{cases} \omega_s & s = 1, \ldots, m, \\ -\varepsilon\tau_{s-m} & s = m+1, \ldots, m+\mu. \end{cases} \tag{6b}$$

In matrix form

$$\mathbf{N}(\varepsilon) = \begin{pmatrix} \mathbf{\Omega} & \mathbf{O} \\ \mathbf{O} & -\varepsilon\mathbf{T} \end{pmatrix}. \tag{7}$$

We can now define $(m+\mu)$-by-$(m+\mu)$ matrix $\mathbf{S}(\varepsilon)$, which depends on a parameter ε and which has matrix elements

$$\mathbf{S}_{sp}(\varepsilon) = \sum_{i=1}^{n} \frac{\langle w_s|\Phi_i\rangle\langle\Phi_i|z_p\rangle}{\varepsilon - \lambda_i} \qquad (s, p = 1, \ldots, m+\mu). \tag{8a}$$

Each nonzero eigenvalue $\varepsilon_k \notin \{\lambda_i\}$ of the eigenvalue eq. (1) is now a solution of the equation

$$\mathcal{D}(\varepsilon) \equiv |\mathbf{S}(\varepsilon) - \mathbf{N}^{-1}(\varepsilon)| = 0, \tag{9}$$

where $\mathcal{D}(\varepsilon)$ is the determinant of a matrix $\mathbf{R}(\varepsilon) \equiv \mathbf{S}(\varepsilon) - \mathbf{N}^{-1}(\varepsilon)$. Inversely, each root $\varepsilon_0 \notin \{\lambda_i\}$ of $\mathcal{D}(\varepsilon)$ is an eigenvalue of eq. (1) [1].

Further, if $\varepsilon_0 \notin \{\lambda_i\}$ is a nonzero root of (9) and hence an eigenvalue of (1), each eigenstate Ψ corresponding to this eigenvalue is of the form

$$|\Psi\rangle = \sum_{i=1}^{n} \left[\left(\sum_{p=1}^{m+\mu} \langle \Phi_i | z_p \rangle C_p \right) / (\varepsilon_0 - \lambda_i) \right] |\Phi_i\rangle, \tag{10a}$$

where

$$C_p = \eta_p(\varepsilon_0) \langle w_p | \Psi \rangle. \tag{10b}$$

Moreover, the coefficients C_p $(p = 1, \ldots, m + \mu)$ satisfy

$$\sum_{p=1}^{m+\mu} \left[\mathbf{S}_{sp}(\varepsilon_0) - \delta_{sp} / \eta_s(\varepsilon_0) \right] C_p = 0, \qquad s = 1, \ldots, m + \mu. \tag{10c}$$

Inversely, if $\varepsilon_0 \notin \{\lambda_i\}$ is a nonzero eigenvalue of (1), each state Ψ of the form (10a) where the coefficients C_p are the (nontrivial) solution of the linear set (10c) is the corresponding eigenstate. Moreover, these coefficients satisfy (10b) [1].

In the general LRP method where matrices **B**, **V**, **C**, and **P** can be almost arbitrary, one has to distinguish between right and left eigenstates. In particular, relations (10) apply only to the right-hand side eigenstates $|\Psi\rangle \equiv |\Psi^{R}\rangle$ of the perturbed eq. (1), while bra vectors $\langle \Phi_i | \equiv \langle \Phi_i^{L} |$ in (10a) refer to the left-hand side eigenstates of the unperturbed eq. (2). There are analogous relations for the left-hand side eigenstates of $\langle \Psi^{L}|$ of (1) [1]. However, in this paper we will explicitly consider only a real LRP problem. One can easily show that in this case, left- and right-hand side eigenstates coincide.

Relations (9) and (10) solve the perturbed eigenvalue eq. (1) for all "cardinal" eigenvalues $\varepsilon_k \notin \{\lambda_i\}$ except, possibly, for the eigenvalue $\varepsilon_k = 0$. This eigenvalue can be easily treated separately. There are analogous relations for the "singular" eigenvalues $\varepsilon_k \in \{\lambda_i\}$ [1]. For the sake of simplicity, here we will consider only cardinal eigenvalues.

In order to find cardinal eigenvalues, one has to solve the characteristic equation $\mathcal{D}(\varepsilon) = 0$. The function $\mathcal{D}(\varepsilon)$ is an $(m + \mu)$-order determinant, and it is analytic in ε. Unless there is some accidental cancellation of terms [1,2], this determinant has a pole at each λ_i, and nowhere else. The characteristic equation $\mathcal{D}(\varepsilon) = 0$ can be solved iteratively, and the corresponding operation count strongly

depends on $(m + \mu)$. Since one can assume $m + \mu = \rho$, this method of deriving cardinal eigenvalues is efficient whenever the rank ρ of the perturbation $\{\mathbf{V}, \mathbf{P}\}$ is small ($\rho \ll n$).

Once the root $\varepsilon = \varepsilon_0$ of (9) is known, one finds the corresponding eigenstate(s) using relations (10a) and (10c). Since the determinant $\mathcal{D}(\varepsilon_0)$ vanishes, the homogeneous linear set (10c) of $(m + \mu)$ equations in $(m + \mu)$ unknowns C_p has at least one nontrivial solution. After the coefficients C_p are derived, one obtains the corresponding eigenstate Ψ by inserting these coefficients into (10a). Assuming again $\rho = m + \mu$, the solution of the homogeneous linear system (10c) requires $O(\rho^3)$ operations. The subsequent insertion of the obtained coefficients C_p into (10a) requires $\approx \rho n$ operations. The derivation of Ψ is hence also efficient whenever $\rho \ll n$.

The derivation of the eigenstate(s) corresponding to a given eigenvalue $\varepsilon = \varepsilon_0$ presents no problem numerically. There are very stable and reliable methods for the solution of the homogeneous linear set of equations [3,8]. The solution of relations (10a) and (10c) is hence simple and straightforward. This is not the case for the derivation of the eigenvalue ε_0. Each eigenvalue $\varepsilon_0 \notin \{\lambda_i\}$ of (1) is a root of the LRP eq. (9). In general, one obtains this root by some iterative method starting from an approximate solution $\varepsilon_0' \approx \varepsilon_0$ [1,2]. As predicted theoretically [1] and as shown by numerical experiments [2], the calculation of a single eigenvalue requires $O(\rho^2 n)$ operations. However, if the initial approximate root ε_0' is not well chosen, the iteration may sometimes diverge. There is no simple and easy way to guarantee the convergence in all cases. This problem is further complicated in the case of degenerate and nearly degenerate roots. It is hence quite important to formulate a stable and reliable algorithm for the calculation of the eigenvalues of (1). We will show how this can be done in the case of the real LRP problem, i.e. in the case where all matrices are Hermitian and where in addition matrices $\mathbf{C}$ and $(\mathbf{C} + \mathbf{P})$ are positive definite.

3. Eigenvalues in the case of the real LRP problem

Consider a real LRP problem. Since matrix $(\mathbf{C} + \mathbf{P})$ is by assumption Hermitian and positive definite, matrix $(\mathbf{C} + \mathbf{P})^{-1/2}$ is a well-defined Hermitian matrix. Relation (1) can hence be transformed to a simple eigenvalue equation with matrix $\mathbf{A} = (\mathbf{C} + \mathbf{P})^{-1/2}(\mathbf{B} + \mathbf{V})(\mathbf{C} + \mathbf{P})^{-1/2}$. Since matrices $\mathbf{B}$, $\mathbf{V}$ and $(\mathbf{C} + \mathbf{P})^{-1/2}$ are Hermitian, matrix $\mathbf{A}$ is also Hermitian. It follows that the eigenvalues ε_k of the eigenvalue eq. (1) are real and that the corresponding eigenstates Ψ_k form a complete set.

Instead of the perturbed eigenvalue eq. (1), we will consider the auxiliary eigenvalue equation

$$(\mathbf{B} + \mathbf{V} - \xi \mathbf{P})\, \chi_j(\xi) = \gamma_j(\xi)\, \mathbf{C}\, \chi_j(\xi), \tag{11}$$

where ξ is a parameter. Each eigenvalue $\gamma_j(\xi)$ and the corresponding eigenstate(s) $\chi_j(\xi)$ of the eigenvalue eq. (11) is a function of a parameter ξ. One easily finds that

in the point ξ such that $\gamma_j(\xi) = \xi$, one has $\xi = \varepsilon_k$ for some eigenvalue ε_k of (1). Inversely, each eigenvalue ε_k of (1) satisfies $\gamma_j(\varepsilon_k) = \varepsilon_k$ for some j. Further, since all eigenvalues ε_k are real, the relation $\gamma_j(\xi) = \xi$ can have only real solutions. Geometrically, each eigenvalue ε_k of relation (1) lies on the intersection of some function $\gamma_j(\xi)$ with the line $\gamma = \xi$ (see fig. 1).

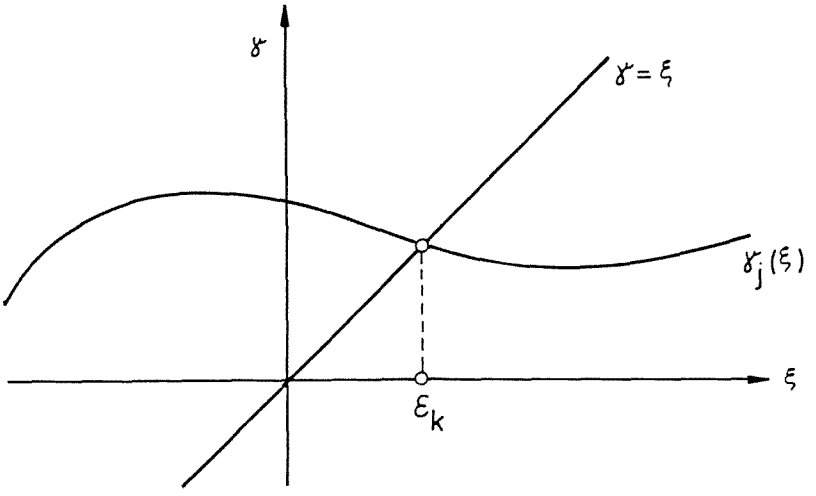

Fig. 1. Each eigenvalue $\gamma_j(\xi)$ of the auxiliary eq. (9) crosses the line $\gamma = \xi$ exactly once. These crossing points are eigenvalues ε_k of the perturbed eq. (1).

The auxiliary eigenvalue eq. (11) and the unperturbed eigenvalue eq. (2) are equivalent to

$$(\mathbf{B}' + \mathbf{V}' - \xi \mathbf{P}')\, \chi_j'(\xi) = \gamma_j(\xi)\, \chi_j'(\xi), \tag{12a}$$

$$\mathbf{B}' \Phi_i' = \lambda_i \Phi_i', \tag{12b}$$

respectively, where

$$\mathbf{B}' = \mathbf{C}^{-1/2}\mathbf{B}\mathbf{C}^{-1/2}, \quad \mathbf{V}' = \mathbf{C}^{-1/2}\mathbf{V}\mathbf{C}^{-1/2}, \quad \mathbf{P}' = \mathbf{C}^{-1/2}\mathbf{P}\mathbf{C}^{-1/2},$$

$$\chi_j' = \mathbf{C}^{1/2}\chi_j, \quad \Phi_i' = \mathbf{C}^{1/2}\Phi_i. \tag{12c}$$

Since $\mathbf{C}$ is by assumption Hermitian and positive definite, $\mathbf{C}^{-1/2}$ is well defined and Hermitian. Since matrices $\mathbf{B}$, $\mathbf{V}$, and $\mathbf{P}$ are also Hermitian, it follows that matrices $\mathbf{B}'$, $\mathbf{V}'$, and $\mathbf{P}'$ are Hermitian as well. Each eigenvalue $\gamma_j(\xi)$ is hence a real function of ξ. In addition, since all matrices are finite dimensional, each eigenvalue $\gamma_j(\xi)$ is

an analytic function on the real ξ axis [5]. In particular, $\gamma_j(\xi)$ has a well-defined derivative for each real ξ.

Consider now the derivative $\partial\gamma_j/\partial\xi$. Using the identity

$$\partial\langle\chi_j'(\xi)\,|\,\chi_j'(\xi)\rangle/\partial\xi \equiv \langle\partial\chi_j'(\xi)/\partial\xi\,|\,\chi_j'(\xi)\rangle + \langle\chi_j'(\xi)\,|\,\partial\chi_j'(\xi)/\partial\xi\rangle = 0,$$

where $\chi_j'(\xi)$ is a normalized eigenstate of (12a), one finds

$$\partial\gamma_j/\partial\xi = -\langle\chi_j'(\xi)\,|\,\mathbf{P}'\,|\,\chi_j'(\xi)\rangle. \tag{13}$$

Using this expression, we prove in the appendix the following lemma:

LEMMA 1

Let matrices **B**, **V**, **C**, and **P** be Hermitian, and let matrices **C** and **C** + **P** be positive definite. Further, let $\gamma_j(\xi)$ be the eigenvalue of (11). Then, for real ξ,

$$\partial\gamma_j(\xi)/\partial\xi < 1. \tag{14}$$

Geometrically, the slope of the function $\gamma_j(\xi)$ is smaller than 1. Hence, each function $\gamma_j(\xi)$ crosses the line $\gamma = \xi$ at most once. In the appendix, we show that this function crosses the line $\gamma = \xi$ exactly once, and hence the equation $\gamma_j(\xi) = \xi$ has exactly one solution (see fig. 1).

Let now $\mathcal{N}(\mathbf{B}, \mathbf{C}, \varepsilon)$ denote the number of eigenvalues of the generalized eigenvalue eq. (2) which are $> \varepsilon$, and let $\mathcal{N}(\mathbf{A}, \varepsilon) \equiv \mathcal{N}(\mathbf{A}, \mathbf{I}, \varepsilon)$ denote the number of eigenvalues of the Hermitian matrix **A** that are $> \varepsilon$. In the appendix, we prove:

THEOREM 1

Let matrices **B**, **V**, **C**, and **P** be Hermitian, and let matrices **C** and (**C** + **P**) be positive definite. Further, let ξ be arbitrary but fixed. Then

$$\mathcal{N}(\mathbf{B} + \mathbf{V} - \xi\mathbf{P}, \mathbf{C}, \xi) = \mathcal{N}(\mathbf{B} + \mathbf{V}, \mathbf{C} + \mathbf{P}, \xi). \tag{15}$$

The number of eigenvalues $\gamma_j(\xi)$ of relation (11) that are $> \xi$ equals the number of eigenvalues ε_k of relation (1) that are $> \xi$. This theorem thus relates the eigenvalues ε_k of the perturbed eigenvalue eq. (1) with the eigenvalues $\gamma_j(\xi)$ of the auxiliary eigenvalue eq. (11).

Since we are considering a real LRP problem, matrices **V** and **P** are Hermitian, and hence they can be represented in the symmetric form

$$\mathbf{V} = \sum_{s=1}^{m} \omega_s \boldsymbol{u}_s \boldsymbol{u}_s^+ = \mathbf{U}\boldsymbol{\Omega}\mathbf{U}^+, \quad \mathbf{P} = \sum_{s=1}^{\mu} \tau_s \boldsymbol{x}_s \boldsymbol{x}_s^+ = \mathbf{X}\mathbf{T}\mathbf{X}^+, \tag{5c}$$

where $\boldsymbol{\Omega} = \mathrm{diag}(\omega_1, \ldots, \omega_m)$ and $\mathbf{T} = \mathrm{diag}(\tau_1, \ldots, \tau_\mu)$ are now real diagonal matrices, while $\mathbf{U} = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m]$ and $\mathbf{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_\mu]$ are n-by-m and n-by-μ matrices, respectively. In particular, vectors $|\boldsymbol{u}_s\rangle$ and $|\boldsymbol{x}_s\rangle$ can be chosen to be normalized eigenstates of operators $\mathbf{V}$ and $\mathbf{P}$, respectively, in which case ω_s and τ_s are the corresponding eigenvalues.

Since matrices $\mathbf{V}$ and $\mathbf{P}$ are now represented in a symmetric form (5c), relation (8a) reduces to

$$\mathbf{S}_{sp}(\varepsilon) = \sum_{i=1}^{n} \frac{\langle z_s|\Phi_i\rangle\langle\Phi_i|z_p\rangle}{\varepsilon - \lambda_i} \qquad (s,p = 1, \ldots, m+\mu), \tag{8b}$$

where vectors z_s are given with (6a). Matrix $\mathbf{S}(\varepsilon)$ is hence Hermitian whenever ε is real. Since $\boldsymbol{\Omega}$ and $\mathbf{T}$ are real, matrix $\mathbf{N}(\varepsilon)$ is also Hermitian for real ε.

Consider now matrices $\mathbf{K}(\varepsilon, \xi)$ and $\mathbf{L}(\varepsilon, \xi)$

$$\mathbf{K} = \begin{pmatrix} \mathbf{B}' + \mathbf{V}' - \xi\mathbf{P}' - \varepsilon & \mathbf{O} \\ \mathbf{O} & -\mathbf{N}^{-1}(\xi) \end{pmatrix}, \quad \mathbf{L} = \begin{pmatrix} \mathbf{B}' - \varepsilon & \mathbf{O} \\ \mathbf{O} & \mathbf{S}(\varepsilon) - \mathbf{N}^{-1}(\xi) \end{pmatrix}. \tag{16}$$

These matrices are defined for each $\varepsilon \notin \{\lambda_i\}$ and each $\xi \neq 0$. In the points $\varepsilon \in \{\lambda_i\}$, matrix $\mathbf{S}(\varepsilon)$ is not defined, while in the point $\xi = 0$, matrix $\mathbf{N}(\xi)$ has no inverse. Hence, in these points matrices $\mathbf{K}$ and $\mathbf{L}$ are not defined. Note also that matrices $\mathbf{K}(\varepsilon, \xi)$ and $\mathbf{L}(\varepsilon, \xi)$ are Hermitian whenever the parameters ε and ξ are real. This follows from the Hermiticity of matrices $\mathbf{B}'$, $\mathbf{V}'$, $\mathbf{P}'$, $\mathbf{S}(\varepsilon)$, and $\mathbf{N}(\xi)$.

We now recall the definition of the congruency of matrices $\mathbf{A}$ and $\mathbf{B}$, and the definition of the inertia of a Hermitian matrix $\mathbf{A}$ [6–8]. By definition, matrices $\mathbf{A}$ and $\mathbf{B}$ are "congruent" if there exists a nonsingular matrix $\mathbf{M}$ such that $\mathbf{B} = \mathbf{M}\mathbf{A}\mathbf{M}^+$ [6]. Further, the inertia of a Hermitian matrix $\mathbf{A}$ is a triad $(\nu(\mathbf{A}), \zeta(\mathbf{A}), \pi(\mathbf{A}))$, where $\nu(\mathbf{A})$, $\zeta(\mathbf{A})$, and $\pi(\mathbf{A})$ are, respectively, the number of negative, zero, and positive eigenvalues of $\mathbf{A}$ [6–8].

In the appendix, we prove the following:

LEMMA 2

Matrices $\mathbf{K}(\varepsilon, \xi)$ and $\mathbf{L}(\varepsilon, \xi)$ are congruent.

By Sylvester's law [6–8], if matrices $\mathbf{A}$ and $\mathbf{B}$ are Hermitian and if they are congruent, then they have the same inertia. Lemma 2 hence implies that for real ξ and ε, matrices $\mathbf{K}$ and $\mathbf{L}$ have the same inertia. In particular, these matrices have

the same number of positive eiganvalues, i.e. $\pi(\mathbf{K}) = \pi(\mathbf{L})$. However, according to (16),

$$\begin{aligned} \pi(\mathbf{K}) &= \pi(\mathbf{B}' + \mathbf{V}' - \xi\mathbf{P}' - \varepsilon) + \pi(-\mathbf{N}^{-1}(\xi)), \\ \pi(\mathbf{L}) &= \pi(\mathbf{B}' - \varepsilon) + \pi(\mathbf{S}(\varepsilon) - \mathbf{N}^{-1}(\xi)). \end{aligned} \tag{17a}$$

One also finds $\pi(\mathbf{A}) = \pi(\mathbf{A}^{-1})$ for a nonsingular matrix $\mathbf{A}$. Further, one can write

$$\pi(\mathbf{B}' + \mathbf{V}' - \xi\mathbf{P}' - \varepsilon) = \mathcal{N}(\mathbf{B}' + \mathbf{V}' - \xi\mathbf{P}', \varepsilon), \ \pi(\mathbf{B}' - \varepsilon) = \mathcal{N}(\mathbf{B}', \varepsilon), \tag{17b}$$

where $\mathcal{N}(\mathbf{A}, \varepsilon) \equiv \mathcal{N}(\mathbf{A}, \mathbf{I}, \varepsilon)$ is the number of eigenvalues of a Hermitian matrix $\mathbf{A}$ that are $> \varepsilon$.

Collecting the above relations, one finds

$$\mathcal{N}(\mathbf{B}' + \mathbf{V}' - \xi\mathbf{P}', \varepsilon) = \mathcal{N}(\mathbf{B}', \varepsilon) + \pi(\mathbf{S}(\varepsilon) - \mathbf{N}^{-1}(\xi)) - \pi(-\mathbf{N}(\xi)).$$

This expression relates the number of eigenvalues γ_k of the matrix $(\mathbf{B}' + \mathbf{V}' - \xi\mathbf{P}')$ that are $> \varepsilon$ to the number of eigenvalues λ_i of the matrix $\mathbf{B}'$ that are $> \varepsilon$. This relation is true for each real ε, and in particular for the choice of $\varepsilon = \xi$. However, by theorem 1 the number of eigenvalues of the matrix $(\mathbf{B}' + \mathbf{V}' - \xi\mathbf{P}')$ that are $> \xi$ equals the number of eigenvalues of relation (1) that are $> \xi$. Moreover, matrix $\mathbf{B}'$ and the eigenvalue eq. (2) have the same eigenvalues. One thus finally obtains:

THEOREM 2

Let ε be arbitrary but fixed, and let matrices $\mathbf{B}$, $\mathbf{V}$, $\mathbf{C}$, and $\mathbf{P}$ be Hermitian. Further, let matrices $\mathbf{C}$ and $(\mathbf{C} + \mathbf{P})$ be positive definite. Then

$$\mathcal{N}(\mathbf{B} + \mathbf{V}, \mathbf{C} + \mathbf{P}, \varepsilon) = \mathcal{N}(\mathbf{B}, \mathbf{C}, \varepsilon) + \pi(\mathbf{S}(\varepsilon) - \mathbf{N}^{-1}(\varepsilon)) - \pi(-\mathbf{N}(\varepsilon)). \tag{18}$$

Expression (18) relates the number $\mathcal{N}(\mathbf{B} + \mathbf{V}, \mathbf{C} + \mathbf{P}, \varepsilon)$ of eigenvalues ε_k of the perturbed eigenvalue eq. (1) that are $> \varepsilon$ to the number $\mathcal{N}(\mathbf{B}, \mathbf{C}, \varepsilon)$ of eigenvalues λ_i of the unperturbed eigenvalue eq. (2) that are $> \varepsilon$. This expression generalizes a result obtained previously by Arbenz and Golub [9]. These authors have obtained a similar equation for a special case $\mathbf{C} = \mathbf{I}$ and $\mathbf{P} = \mathbf{O}$. An expression analogous to the one obtained by Arbenz and Golub was independently derived by Beatie and Fox [10]. A corresponding formula for a substantially more restricted eigenvalue problem was given by Simpson [11].

The proof of the above theorem depends in an essential way on the fact that matrices $\mathbf{C}$ and $\mathbf{C} + \mathbf{P}$ are Hermitian and positive definite. In particular, it is not sufficient for these matrices to be only Hermitian. For example, if $\mathbf{C}$ is Hermitian

but not positive definite, then $\mathbf{C}^{-1/2}$ is either not defined (in the case when $\mathbf{C}$ is singular) or if defined, it is not Hermitian. Namely, if $\mathbf{C}$ has some negative eigenvalue, then the corresponding eigenvalue of $\mathbf{C}^{-1/2}$ is imaginary and hence $\mathbf{C}^{-1/2}$ can not be Hermitian. Hence, neither the eigenvalues λ_i of the unperturbed eigenvalue eq. (2), nor the eigenvalues $\gamma_j(\xi)$ of the auxiliary eigenvalue eq. (11) are guaranteed to be real. Matrices $\mathbf{B}'$, $\mathbf{V}'$, and $\mathbf{P}'$ are also not Hermitian, except possibly by accident. On the other hand, if $\mathbf{C} + \mathbf{P}$ is Hermitian but not positive definite, then one finds, following an analogous argument as above, that the eigenvalues ε_k of the perturbed eigenvalue eq. (1) are not guaranteed to be real. The same conclusion can be obtained in yet another way. The proof of lemma 1 depends in an essential way on the fact that $\mathbf{C} + \mathbf{P}$ is positive definite. If this matrix is not positive definite, eigenvalues $\gamma_j(\xi)$ of the auxiliary eigenvalue eq. (11) are not guaranteed to have slope less than one. Hence, the function $\gamma_j(\xi)$ is not guaranteed to cross the line $\gamma = \xi$, i.e. $\gamma_j(\xi) = \xi$ is not guaranteed to have a solution for real ξ.

Relation (18) enables an efficient bracketing of the eigenvalues ε_k of the perturbed eigenvalue eq. (1). This relation is defined for each real ε, except for the points $\varepsilon \in \{\lambda_i\}$ which are eigenvalues of the unperturbed eigenvalue eq. (2) and which are poles of the matrix $\mathbf{S}(\varepsilon)$, and for the point $\varepsilon = 0$ which is a pole of a matrix $\mathbf{N}^{-1}(\varepsilon)$. If one evaluates $\mathcal{N}(\mathbf{B} + \mathbf{V}, \mathbf{C} + \mathbf{P}, \varepsilon)$ in two points $\varepsilon = \varepsilon'$ and $\varepsilon = \varepsilon''$, the difference of the two numbers obtained is the number of eigenvalues ε_k that are inside the interval $(\varepsilon', \varepsilon'']$. Choosing different points $\varepsilon = \varepsilon'$ for the evaluation of the expression $\mathcal{N}(\mathbf{B} + \mathbf{V}, \mathbf{C} + \mathbf{P}, \varepsilon)$, one can successfully bracket each eigenvalue ε_k.

Relation (18) can be modified in such a way as to include points $\varepsilon \in \{\lambda_i\}$. However, this more general relation involves matrices with larger dimensions, thus increasing the operation count. From a theoretical point of view, relation (18) is sufficient since, using this relation, one can bracket in an arbitrarily small interval each eigenvalue ε_k, including eigenvalues $\varepsilon_k \in \{\lambda_i\}$. Some numerical difficulties could arise only if $\mathbf{S}(\varepsilon)$ is evaluated at some point too close to an eigenvalue λ_i, since $\mathbf{S}(\varepsilon)$ diverges at $\varepsilon \in \{\lambda_i\}$.

4. Numerical considerations

Relation (18) is our basic equation for the derivation of the eigenvalues ε_k of the perturbed eigenvalue eq. (1) in the case of the real LRP problem. Usually, one has to recalculate this expression repeatedly for various values of the parameter ε. This is necessary in order to bracket and finally isolate the eigenvalues ε_k. It is hence important that repeated recalculations of this expression be fast and efficient.

All the quantities on the right-hand side of (18) can be obtained relatively easily. Calculation of the quantities $\mathcal{N}(\mathbf{B}, \mathbf{C}, \varepsilon)$ and $\pi(-\mathbf{N}(\varepsilon))$ presents no problem. The number $\mathcal{N}(\mathbf{B}, \mathbf{C}, \varepsilon)$ of eigenvalues λ_i that are $> \varepsilon$ is by assumption known. In addition, this number is constant if ε is required to be contained inside the interval between two consecutive eigenvalues λ_i.

Concerning the number $\pi(-\mathbf{N}(\varepsilon)) = \nu(\mathbf{N}(\varepsilon))$, one has from (7)

$$\pi(-\mathbf{N}(\varepsilon)) = \begin{cases} \nu(\Omega) + \nu(\mathbf{T}) & \text{if } \varepsilon < 0, \\ \nu(\Omega) + \pi(\mathbf{T}) & \text{if } \varepsilon > 0. \end{cases}$$

Since $\mathbf{\Omega}$ and $\mathbf{T}$ are diagonal, this is trivial to calculate. This number does not change unless ε changes sign. In addition, if the vectors $\boldsymbol{u}_s$ as well as the vectors $\boldsymbol{x}_s$ in the decomposition (5a) are linearly independent, then $\nu(\mathbf{\Omega}) = \nu(\mathbf{V})$, $\pi(\mathbf{T}) = \pi(\mathbf{P})$ and $\nu(\mathbf{T}) = \nu(\mathbf{P})$.

Only the number $\pi(\mathbf{S}(\varepsilon) - \mathbf{N}^{-1}(\varepsilon))$ of the positive eigenvalues of a matrix $\mathbf{R}(\varepsilon) = \mathbf{S}(\varepsilon) - \mathbf{N}^{-1}(\varepsilon)$ requires some nontrivial calculation. As a function of ε this number is constant, except for the points $\varepsilon = \varepsilon_k \notin \{\lambda_i\}$ which are eigenvalues of (1). In these points, matrix $\mathbf{R}(\varepsilon) = \mathbf{S}(\varepsilon) - \mathbf{N}^{-1}(\varepsilon)$ is singular, and hence the corresponding determinant vanishes. This recovers the original LRP eq. (9), with the restriction $|w_s\rangle = |z_s\rangle$ in (8a). Considered as a function of ε, matrix $\mathbf{R}(\varepsilon)$ changes the number of positive eigenvalues at each $\varepsilon = \varepsilon_k$.

The number $\pi(\mathbf{R}(\varepsilon))$ can be calculated in three steps. First, one has to form scalar products $\langle z_s | \Phi_i \rangle$, which are required in order to form matrix elements $\mathbf{S}_{sp}(\varepsilon)$, and hence matrix $\mathbf{R}(\varepsilon)$. These scalar products should be calculated only once, irrespective of how many eigenvalues ε_k are to be calculated. Next, one has to form matrix $\mathbf{R}(\varepsilon)$. For each eigenvalue ε_k, one has to recalculate this matrix many times in order to obtain the required convergence. Finally, each time matrix $\mathbf{R}(\varepsilon)$ is constructed, one has to find the number $\pi(\mathbf{R}(\varepsilon))$ of positive eigenvalues of this matrix.

We will now consider in more detail these three steps. Assume for simplicity $m + \mu = \rho_v + \rho_p = \rho$.

The first step, calculation of the scalar products $\langle z_s | \Phi_i \rangle$ is straightforward. There are ρn such scalar products, and each requires n multiplications. The operation count is hence $\approx \rho n^2$. As stated above, this calculation should be done only once, irrespective of how many eigenvalues ε_k are required. In addition, if the perturbation $\{\mathbf{V}, \mathbf{P}\}$ is local with localizability l, then at most l components of vectors z_s are nonzero. Hence, the calculation of each scalar product $\langle z_s | \Phi_i \rangle$ requires at most l multiplications and the operation count reduces to $\approx \rho l n$. Since the rank ρ_v and ρ_p can not exceed the localizability l of the corresponding matrices $\mathbf{V}$ and $\mathbf{P}$, one finds $\rho \leq 2l$. In addition, if the nonzero columns of the matrix $\mathbf{V}$ as well as the nonzero columns of the matrix $\mathbf{P}$ are linearly independent (which is usually the case), then $\rho \geq l$. The operation count for the calculation of the scalar products $\langle z_s | \Phi_i \rangle$ in the case of a local perturbation $\{\mathbf{V}, \mathbf{P}\}$ is hence of the order $O(\rho^2 n)$. Finally, in some quantum chemical applications scalar products $\langle z_s | \Phi_i \rangle$ can be simply postulated, in which case the associated operation count is $O(1)$.

Consider now the third step, which involves the calculation of the number $\pi(\mathbf{R}(\varepsilon))$ once the matrix $\mathbf{R}(\varepsilon)$ is constructed. One way to calculate this number is

to diagonalize $\mathbf{R}(\varepsilon)$. Since $\mathbf{R}(\varepsilon)$ is a Hermitian ρ-by-ρ matrix, this requires approximately $(2/3)\rho^3 + 30\rho^2$ operations if the very efficient Householder-QR method is applied [3]. However, a more efficient way to calculate this number is to perform the so-called $\mathbf{LDL}^+$ decomposition of the matrix $\mathbf{R}(\varepsilon)$:

$$\mathbf{L}(\varepsilon)\mathbf{D}(\varepsilon)\mathbf{L}(\varepsilon)^+ = \mathbf{R}(\varepsilon),$$

where $\mathbf{L}(\varepsilon)$ is a unit lower triangular matrix, $\mathbf{L}(\varepsilon)^+$ is a unit upper triangular matrix (complex conjugate of $\mathbf{L}(\varepsilon)$), while $\mathbf{D}(\varepsilon) = \mathrm{diag}(d_1(\varepsilon), d_2(\varepsilon), \ldots, d_\rho(\varepsilon))$ is a diagonal matrix. By Sylvester's law [6] matrices $\mathbf{D}(\varepsilon)$ and $\mathbf{R}(\varepsilon)$ have the same inertia. Hence, $\pi(\mathbf{R}(\varepsilon)) = \pi(\mathbf{D}(\varepsilon))$. Further, the $\mathbf{LDL}^+$ decomposition of a ρ-by-ρ Hermitian matrix requires approximately $\rho^3/6$ operations [8], which is much less than in the case of a diagonalization.

Although this $\mathbf{LDL}^+$ decomposition has a low operation count, it can not be recommended. Such a direct decomposition is not guaranteed to be stable and it can even fail [3,8]. In order to stabilize this decomposition, some previous pivoting strategy should be used [8]. Since the inertia of $\mathbf{D}(\varepsilon)$ should be the same as the inertia of $\mathbf{R}(\varepsilon)$, only a symmetric pivoting strategy can be used. A natural choice is to use the so-called diagonal pivoting method [8,12–14]:

$$\mathbf{L}(\varepsilon)\,\mathbf{B}(\varepsilon)\,\mathbf{L}(\varepsilon)^+ = P\mathbf{R}(\varepsilon)P^{\mathrm{T}}, \tag{19}$$

where $\mathbf{B}(\varepsilon)$ is a block diagonal matrix with 1-by-1 and 2-by-2 blocks, $\mathbf{L}(\varepsilon)$ is a unit lower triangular matrix, and P is a permutation determined in such a way that $|\mathbf{L}_{sp}| \leq 1$. Decomposition (19) is always possible, it is stable, and it requires approximately $\rho^3/6$ operations [12–14]. By Sylvester's law, the inertia of the matrix $\mathbf{R}(\varepsilon)$ equals the inertia of the matrix $\mathbf{B}(\varepsilon)$. Since $\mathbf{L}$ is a unit triangular matrix, the determinant of $\mathbf{R}(\varepsilon)$ also equals the determinant of $\mathbf{B}(\varepsilon)$. Further, since $\mathbf{B}(\varepsilon)$ is a block diagonal matrix with at most 2-by-2 blocks, it can be easily diagonalized to a diagonal matrix $\mathbf{D}(\varepsilon) = \mathrm{diag}(\boldsymbol{d}_1(\varepsilon), \ldots, \boldsymbol{d}_\rho(\varepsilon))$ with a known inertia. This step requires $\approx O(\rho)$ operations. The total operation count for the derivation of a matrix $\mathbf{D}(\varepsilon)$ is hence $\approx \rho^3/6$.

The above algorithm determines the inertia and the determinant of $\mathbf{R}(\varepsilon)$ in a stable and efficient way. Diagonal elements $\boldsymbol{d}_i(\varepsilon)$ of $\mathbf{D}(\varepsilon)$ and the determinant $|\mathbf{R}| = |\mathbf{D}| = \prod_i \boldsymbol{d}_i(\varepsilon)$ can be used in order to accelerate the search for the eigenvalues ε_k of the eigenvalue eq. (1). Since the determinant $|\mathbf{R}(\varepsilon_k)|$ vanishes, at least one diagonal element $\boldsymbol{d}_i(\varepsilon_k)$ of the matrix $\mathbf{D}(\varepsilon_k)$ vanishes. One can further show that diagonal elements $\boldsymbol{d}_i(\varepsilon)$ are smooth functions of ε, except possibly in a few isolated points. Hence, one can use the analytical properties of the diagonal elements $\boldsymbol{d}_i(\varepsilon)$ in order to improve the convergence to the eigenvalue ε_k. In practical applications, one either knows some initial approximate eigenvalue $\varepsilon_k' \approx \varepsilon_k$, or one looks for the

root(s) ε_k in some in advance given small interval $(\varepsilon', \varepsilon'')$. In either case, after some I iterations one obtains the eigenvalue ε_k. The total operation count associated with the decomposition (19) is hence estimated to be $\approx I\rho^3/6$ per eigenvalue.

The decomposition (19) provides an efficient way to calculate the inertia of a matrix $\mathbf{R}(\varepsilon)$, once this matrix is known. The calculation of this matrix represents our second step, and it requires the calculation of $\rho(\rho+1)/2$ matrix elements $\mathbf{S}_{sp}(\varepsilon)$ of a Hermitian matrix $\mathbf{S}(\varepsilon)$. According to (8b), the derivation of these matrix elements requires $O(\rho^2 n)$ operations, $O(n)$ operations per matrix element. In addition, one has to recalculate matrix $\mathbf{S}(\varepsilon)$ repeatedly for different values of ε. The total operation count associated with the repeated recalculation of these matrix elements is hence of the order $O(In\rho^2)$ per eigenvalue. This is substantially larger than the operation count associated with the decomposition (19). It is hence quite important to decrease this operation count as much as possible.

One way in which this can be done is to replace the exact $O(\rho^2 n)$ calculations of $\rho(\rho+1)/2$ functions $\mathbf{S}_{sp}(\varepsilon)$ with the approximate $O(\rho^2)$ calculation of these functions [1,2]. Assume that the exact eigenvalue ε_k is contained in the interval $(\varepsilon', \varepsilon'')$. We calculate $n_R \ll n$ terms of the sum (8b) that correspond to the eigenvalues λ_i close to this interval exactly. We approximate the remaining $(n - n_R) \approx n$ terms with the first few terms of the Taylor expansion at some point $\varepsilon_0 \in (\varepsilon', \varepsilon'')$ which is the best current estimate of ε_k. In this way, matrix elements $\mathbf{S}_{sp}(\varepsilon)$ are replaced with approximate matrix elements $\mathbf{S}'_{sp}(\varepsilon)$, and the matrix $\mathbf{R}(\varepsilon)$ with the approximate matrix $\mathbf{R}'(\varepsilon)$. The initial construction of the Taylor coefficients requires $O(\rho^2 n)$ operations, while each subsequent iteration requires only $O(\rho^2)$ operations. After I iterations, one obtains an approximate root $\varepsilon'_0 \approx \varepsilon_k$. Matrix elements $\mathbf{S}_{sp}(\varepsilon)$ are then approximated once more with matrix elements $\mathbf{S}''_{sp}(\varepsilon)$, this time using the Taylor expansion at the point $\varepsilon = \varepsilon'_0$. Since the approximate root ε'_0 is already quite close to the exact root ε_k, the new approximate root ε''_0 is practically exact [1,2]. As predicted theoretically [1] and as shown with numerical experiments [2], such an approach reduced the $O(I\rho^2 n)$ calculation of the matrix elements $\mathbf{S}_{sp}(\varepsilon)$ to a $O(\rho^2 n)$ calculation of these matrix elements. The main numerical load is associated with the Taylor expansion of the functions $\mathbf{S}_{sp}(\varepsilon)$. This expansion is done twice, and it is of the order $O(\rho^2 n)$ per eigenvalue. The overhead which is associated with the repeated recalculation of these matrix elements for different values of ε is of the order $O(I\rho^2)$ per eigenvalue. To this one should add the overhead $\approx I\rho^3/6$ associated with the $\mathbf{LBL}^+$ decomposition of matrix $\mathbf{R}(\varepsilon)$. Both overheads are negligible with respect to $O(\rho^2 n)$, especially for large n. The suggested approach thus substantially reduces the operation count for the repeated recalculation of the matrix elements $\mathbf{S}_{sp}(\varepsilon)$. In addition, this approach is also quite robust, since the operation count is essentially independent of the number of iterations I required to derive a single eigenvalue.

In the above numerical analysis of the real LRP method, there is no reference to the magnitude of the perturbation and no reference to the eventual degeneracy or quasi-degeneracy of the unperturbed system. In particular, the estimated operation

counts are not affected by the magnitude of the perturbation and/or by the degeneracy of the unperturbed system. In comparison, the standard perturbation method is not efficient and it can even fail, unless the perturbation is small. Also, a standard perturbation approach requires a special treatment of degenerate and quasi-degenerate systems.

In summary, the suggested real LRP approach requires approximately $O(\rho^2 n^2)$ operations for the calculation of all n eigenvalues ε_k of the perturbed system. If the perturbation $\{\mathbf{V}, \mathbf{P}\}$ is in addition local with localizability $l \approx \rho \ll n$, then the initial construction of scalar products $\langle z_s | \Phi_i \rangle$ requires only $O(\rho^2 n)$ operations. In this case, only $O(\rho^2 n)$ operations are required for the derivation of a single perturbed eigenvalue. If l is sufficiently small, this is almost two orders of magnitude less than any presently known method.

5. Conclusions

The LRP method is designed to find eigenvalues and eigenstates of the generalized eigenvalue equation $(\mathbf{B} + \mathbf{V})\Psi_k = \varepsilon_k(\mathbf{C} + \mathbf{P})\Psi_k$, where the eigenvalues and the eigenstates of the unperturbed eigenvalue equation $\mathbf{B}\Phi_i = \lambda_i \mathbf{C}\Phi_i$ are known, and where $\mathbf{B}$, $\mathbf{V}$, $\mathbf{C}$, and $\mathbf{P}$ are n-by-n matrices. The method is numerically efficient if the rank ρ of the perturbation $\{\mathbf{V}, \mathbf{P}\}$ is "small" with respect to n. In its original formulation, the LRP method applied to arbitrary matrices, with the only restriction that the unperturbed system should have a complete set of eigenstates. However, physically most important is a real LRP problem, where matrices $\mathbf{B}$, $\mathbf{V}$, $\mathbf{C}$, and $\mathbf{P}$ are Hermitian and where in addition matrices $\mathbf{C}$ and $(\mathbf{C} + \mathbf{P})$ are positive definite. These requirements guarantee the reality of the eigenvalues ε_k and λ_i.

The LRP method can be substantially improved in the case of a real LRP problem. In particular, the main numerical problem in the general LRP method is the calculation of the eigenvalues ε_k. In the general LRP method, each eigenvalue ε_k is usually derived in an iterative way, starting from some initial approximate eigenvalue ε_k'. However, if the initial eigenvalue is not well chosen, the iteration may sometimes diverge [1,2]. In the case of the real LRP problem, this danger can be completely avoided. It is shown that in this case the eigenvalues of the perturbed eigenvalue equation can be efficiently bracketed. The method is fast and stable, and it requires $O(\rho^2 n^2)$ operations for the calculation of all n eigenvalues. If the perturbation $\{\mathbf{V}, \mathbf{P}\}$ is in addition local with localizability $l \approx \rho$, only $O(\rho^2 n)$ operations are needed for the derivation of a single eigenvalue. Unlike the standard perturbation method whose efficiency strongly depends on the magnitude of the perturbation and which requires a special treatment of degenerate and quasi-degenerate systems, the efficiency of the LRP method is not affected by the magnitude of the perturbation and/or by the degeneracy or quasi-degeneracy of the unperturbed system.

Appendix

Proof of lemma 1

Since matrix **C** is Hermitian and positive definite, matrix $\mathbf{C}^{-1/2}$ is also Hermitian. This trivially follows from the unitary transformation of matrix **C** in a diagonal form, in which form matrix $\mathbf{C}^{-1/2}$ is obviously Hermitian. Matrices $\mathbf{C}+\mathbf{P}$ and $\mathbf{C}^{-1/2}(\mathbf{C}+\mathbf{P})\mathbf{C}^{-1/2}=\mathbf{I}+\mathbf{P}'$ are hence congruent [6,7]. By Sylvester's law [6,7], these matrices have the same inertia. Since further $\mathbf{C}+\mathbf{P}$ is positive definite, matrix $\mathbf{I}+\mathbf{P}'$ is also positive definite. Hence, $\langle\phi|\mathbf{I}+\mathbf{P}'|\phi\rangle>0$ for each vector ϕ. In particular, this holds for the eigenvectors $\chi_j'(\xi)$ of the eigenvalue eq. (12a). Hence, and from (13), it follows that the derivative $\partial\gamma_j(\xi)/\partial\xi$ is less than one. This proves lemma 1. □

Note that the fact that the slope of the function $\gamma_j(\xi)$ is less than one does not yet imply that the equation $\gamma_j(\xi)=\xi$ has a solution for real ξ. It only implies that this equation has at most one solution. For example, it is quite conceivable that the function $\gamma_j(\xi)$ has a slope less than one, and that it approaches the line $\gamma=\xi$ asymptotically from above, never crossing this line. However, since each eigenvalue ε_k of (1) is a solution of $\gamma_j(\xi)=\xi$ for some j, this equation should have *exactly* one solution. This follows from the fact that there are n eigenvalues ε_k and n functions $\gamma_j(\xi)$.

Proof of theorem 1

For an arbitrary but fixed $\xi=\xi'$, eigenvalues $\gamma_j(\xi')$ of relation (12a) lie on the intersections of the vertical line $\xi=\xi'$ with functions $\gamma_j(\xi)$, while eigenvalues of ε_k of the eigenvalue eq. (1) lie on the intersections of the line $\gamma=\xi$ with functions $\gamma_j(\xi)$ (see fig. 2). Eigenvalues $\gamma_j(\xi')$ that are $>\xi'$ are all these intersections of the vertical line $\xi=\xi'$ with functions $\gamma_j(\xi)$ which are situated above the line $\gamma=\xi$, while the eigenvalues ε_k that are $>\xi'$ are all these intersections of the line $\gamma=\xi$ with functions $\gamma_j(\xi)$ which lie on the right-hand side of the line $\xi=\xi'$. Since the slope of each $\gamma_j(\xi)$ is less than one, each function $\gamma_j(\xi)$ that crosses the vertical line $\xi=\xi'$ above the line $\gamma=\xi$ crosses this line on the right-hand side of the line $\xi=\xi'$ exactly once (see fig. 2). There is hence a one-to-one correspondence between the eigenvalues $\gamma_j(\xi')$ of the eigenvalue eq. (12a) that are $>\xi'$ with the eigenvalues ε_k of the eigenvalue eq. (1) that are $>\xi'$. This proves theorem 1. □

Proof of lemma 2

Matrix elements $\mathbf{S}_{sp}(\varepsilon)$ of a matrix $\mathbf{S}(\varepsilon)$ can be written in the form

$$\mathbf{S}_{sp}(\varepsilon)=\sum_{i=1}^{n}\frac{\langle z_s'|\Phi_i'\rangle\langle\Phi_i'|z_p'\rangle}{\varepsilon-\lambda_i}, \tag{A1}$$

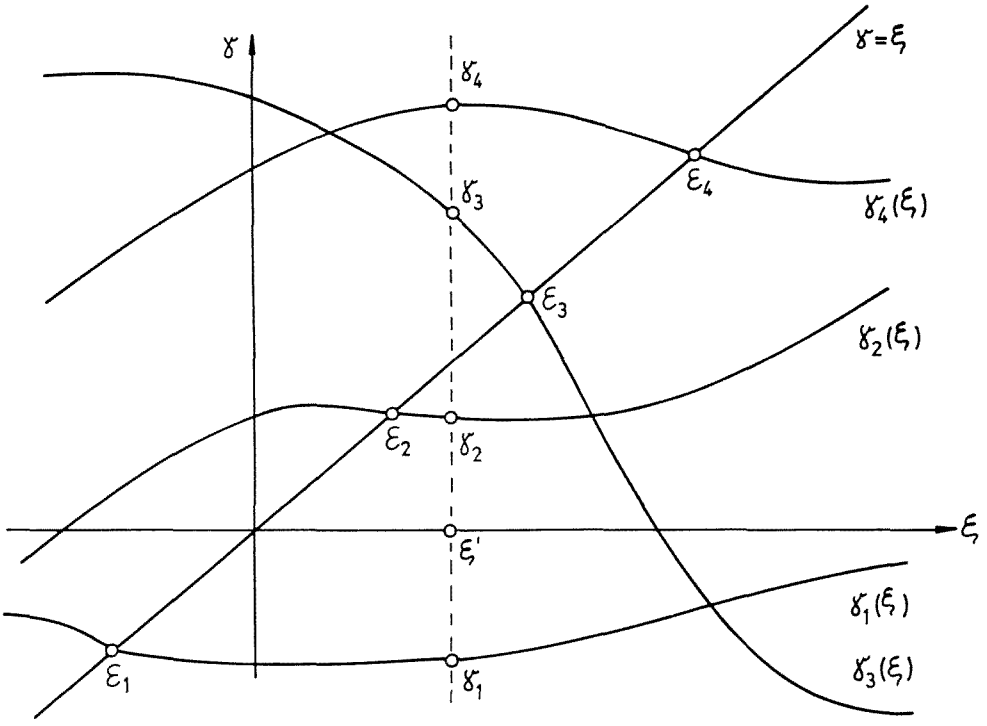

Fig. 2. Eigenvalues ε_k of the perturbed eq. (1) lie on the intersections of the line $\gamma = \xi$ with the functions $\gamma_j(\xi)$, while for an arbitrary but fixed $\xi = \xi'$, eigenvalues $\gamma_j(\xi')$ of the auxiliary eq. (9) lie on the intersections of the vertical line $\xi = \xi'$ with the functions $\gamma_j(\xi)$. Each function $\gamma_j(\xi)$ that crosses the vertical line $\xi = \xi'$ above the line $\gamma = \xi$, crosses the line $\gamma = \xi$ on the right-hand side of the vertical line $\xi = \xi'$. This proves theorem 1.

where $|z'_s\rangle = \mathbf{C}^{-1/2}|z_s\rangle$. Hence,

$$\mathbf{S}(\varepsilon) = \mathbf{Z}'^{+}(\varepsilon - \mathbf{B}')^{-1}\mathbf{Z}', \tag{A2}$$

where $\mathbf{Z}' = [z'_1, \ldots, z'_{m+\mu}]$ is an n-by-$(m+\mu)$ matrix with matrix elements $\mathbf{Z}'_{sp} = \langle s|z'_p\rangle$, where we have used the identity

$$(\varepsilon - \mathbf{B}')^{-1} = \sum_{i=1}^{n} \frac{\Phi'_i \Phi'^{+}_i}{\varepsilon - \lambda_i}, \tag{A3}$$

which follows from (12b). One also finds $\mathbf{V}' - \xi\mathbf{P}' = \sum_s \eta_s(\xi)\, z'_s z'^{+}_s$, or in a matrix form

$$\mathbf{V}' - \xi\mathbf{P}' = \mathbf{Z}'\mathbf{N}(\xi)\mathbf{Z}'^{+}. \tag{A4}$$

Consider now matrix $\mathbf{M}(\varepsilon, \xi)$:

$$\mathbf{M} = \begin{pmatrix} \mathbf{I} & (\mathbf{B}' - \varepsilon)^{-1}\mathbf{Z}' \\ \mathbf{N}(\xi)\mathbf{Z}'^{+} & \mathbf{I} - \mathbf{N}(\xi)\mathbf{S}(\varepsilon) \end{pmatrix}. \tag{A5}$$

Inserting (A2) and (A4) into expression (16) for matrices **K** and **I**, one finds after some algebra $\mathbf{L} = \mathbf{M}^{+}\mathbf{K}\mathbf{M}$. This proves lemma 2. □

Remark

If one analyzes the derivation of the relation $\mathbf{L} = \mathbf{M}^{+}\mathbf{K}\mathbf{M}$, one finds that it is not necessary for matrix $\mathbf{N}(\xi)$ to be diagonal. It is sufficient for this matrix to be Hermitian and non-singular. In particular, matrices $\mathbf{\Omega}$ and **T** which define $\mathbf{N}(\xi)$ can be arbitrary Hermitian and non-singular matrices. Since these matrices determine the decomposition (5) of perturbation matrices **V** and **P**, one can accordingly generalize all the obtained results. However, this generalization is probably not necessary, since the decomposition (5) is already general enough for all practical applications.

Acknowledgement

The author would like to thank Darko Babić for useful discussions and critical reading of the manuscript.

References

[1] T.P. Živković, Theor. Chim. Acta 76(1989)331.
[2] T.P. Živković, J. Math. Chem. 4(1990)1143.
[3] W.H. Press, B.P. Flannery, S.A. Teukolsky and W.T. Vetterling, *Numerical Recipes* (Cambridge University Press, Cambridge, 1986).
[4] T.P. Živković, Comput. Chem. 14(1990)335.
[5] T. Kato, *Perturbation Theory for Linear Operators* (Springer, Berlin/Heidelberg/New York, 1966).
[6] D.K. Faddeev and V.N. Faddeev, *Computational Methods of Linear Algebra* (Freeman, San Francisco/London, 1963).
[7] B.N. Parlett, *The Symmetric Eigenvalue Problem* (Prentice–Hall, Englewood Cliffs, NJ, 1980).
[8] G.H. Golub and C.F. Van Loan, *Matrix Computations* (The Johns Hopkins University Press, Baltimore, MD, 1983).
[9] P. Arbenz and G.H. Golub, SIAM J. Matrix. Anal. Appl. 9(1988)40.
[10] C. Beattie and D. Fox, Tech. Report UMSI 87/11, University of Minnesota Supercomputer Institute (1987).
[11] A. Simpson, Quart. J. Mech. Appl. Math. 27(1974)27.
[12] J.R. Bunch and B.N. Parlett, SIAM J. Num. Anal. 8(1971)639.
[13] J.R. Bunch and L. Kaufman, Math. Comp. 31(1977)162.
[14] J.R. Bunch, L. Kaufman and B.N. Parlett, Numer. Math. 27(1976)95.