Brief Comments on Perturbation Theory of a Nonsymmetric Matrix: The GF Matrix†

R. A. Marcus

Noyes Laboratory of Chemical Physics, California Institute of Technology, Pasadena, California 91125-7200

Received: November 10, 2000; In Final Form: December 20, 2000

Perturbation theory for nonsymmetric matrices is discussed for the **GF** matrix for molecular vibrations. As a simple extension of early results, two approaches are given: one a direct diagonalization of the nonsymmetric matrix, the other a presymmetrization. Presymmetrization of the **GF** matrix, well known for Δ**F**, is also described here for Δ**G**. It permits the use of standard perturbation theory for symmetric matrices. Application of the second-order expression for the Δ**G** case to the determination of the frequencies of many ozone isotopomers is given elsewhere.

1. Introduction

Perturbation theory for the nonsymmetric **GF** matrix for molecular vibrations is well known in the literature for the **F**-matrix.[1−3] In an almost forgotten article, a rigorous first- and second-order treatment for the **G**-matrix was given by Edgell.[4] An approximate derivation for the Δ**G**-perturbation is given in standard texts for the first-order case.[5,6] While it does contain a tacit assumption that the formalism for nonsymmetric matrices is the same as that for symmetric matrices, the final result for the first-order term is the same as that obtained by Edgell.[4] Our interest in the topic arose in a calculation of many unknown frequencies of ozone isotopomers[7] for use in a kinetic study of the "mass-independent" isotope effect in ozone formation.[8]

Using the simple compact formalism of more recent perturbation theory,[9−13] the derivation in ref 4 is readily extended. The treatment for both Δ**G** and Δ**F** is described below in two ways, one of which involves a direct diagonalization of the nonsymmetric matrices[9] (sections 1, 3, and 5). The other is based on a presymmetrization of **GF** (sections 4 and 5), both for Δ**F** and Δ**G**. Presymmetrization for the Δ**F** case, well known in the literature,[1−3] is given for comparison, while that for Δ**G** appears to be new. The present brief comments are intended to reexamine and extend the early results.[4]

2. Theory

A perturbation theory for nonsymmetric matrices or, in general, for operators which are not self-adjoint, is straightforward. For comparison it is useful to consider first the well known elementary perturbation theory for the symmetric or self-adjoint case, and then the very minor modification needed to extend it to the nonsymmetric case.We let **A** denote a symmetric matrix or self-adjoint operator with eigenvectors x_i whose columns form a matrix **X** and with eigenvalues λ_i, the elements of a diagonal matrix Λ.

$$\mathbf{AX} = \mathbf{X}\Lambda \tag{2.1}$$

Since the present application is to vibrations and involves real matrices **A**, we shall use the transpose $\mathbf{A}^T$ of **A** rather than referring to the adjoint. For symmetric matrix **A**, $\mathbf{A} = \mathbf{A}^T$. We let $\mathbf{A}^o$ and **V** denote the unperturbed and perturbation operators or matrices, respectively, and consider the nondegenerate case. Making the customary expansion in a small parameter ϵ, which is later replaced by unity, we have

$$\mathbf{A} = \mathbf{A}^o + \epsilon\mathbf{V} \tag{2.2}$$

$$\lambda_i = \lambda_i^o + \epsilon\lambda_i^{(1)} + \epsilon^2\lambda_i^{(2)} + \cdots \tag{2.3}$$

$$x_i = x_i^o + \epsilon x_i^{(1)} + \epsilon^2 x_i^{(2)} + \cdots \tag{2.4}$$

The usual equating of equal powers of ϵ yields

$$\mathbf{A}^o x_i^o = \lambda_i^o x_i^o \tag{2.5}$$

$$(\mathbf{A}^o - \lambda_i^{\ o})x_i^{(1)} = (\lambda_i^{(1)} - \mathbf{V})x_i^o \tag{2.6}$$

$$(\mathbf{A}^o - \lambda_i^o)x_i^{(2)} = \lambda_i^{(2)}x_i^o + \lambda_i^{(1)}x_i^{(1)} - \mathbf{V}x_i^{(1)} \tag{2.7}$$

The x_i^o form an orthonormal set of eigenvectors, as in eq 2.8 below, and for x_i we use the convenient linear normalization[10,14,15] present in the second part of eq 2.8:

$$x_i^{oT}x_j^o = \delta_{ij} \qquad x_i^{oT}x_i = 1 \tag{2.8}$$

where x_i^{oT} is a row vector corresponding to the column matrix x_i^o. It follows that

$$x_i^{oT}x_i^{(1)} = x_i^{oT}x_i^{(2)} = ... = 0 \tag{2.9}$$

Application of x_i^{oT} to the left of eqs 2.6 and 2.7, taking into account the symmetric nature of $\mathbf{A}^o$, yields

$$\lambda_i^{(1)} = x_i^{oT}x_i^o \tag{2.10}$$

$$x_i^{(2)} = x_i^{oT}\mathbf{V}x_i^{(1)} \tag{2.11}$$

Application of x_j^{oT}, $j \neq i$, to the left of eq 2.6 yields the component of $x_i^{(1)}$ along x_j^o

$$(\lambda_j^o - \lambda_i^o)x_j^{oT}x_i^{(1)} = -x_j^{oT}\mathbf{V}x_i^o \tag{2.12}$$

The vector $x_i^{(1)}$ has no component along x_i^o, as seen in eq 2.9,

† Part of the special issue "William H. Miller Festschrift".

10.1021/jp004164d CCC: $20.00
Published on Web 02/10/2001

and so from eq 2.12 we have

$$x_i^{(1)} = -\sum_{j\neq i} \frac{x_j^{oT}\mathbf{V}x_i^o}{\lambda_j^o - \lambda_i^o} x_j^o \tag{2.13}$$

and eq 2.11 then yields the standard result for $\lambda_i^{(2)}$:

$$\lambda_i^{(2)} = \sum_{j\neq i} \frac{(x_i^{oT}\mathbf{V}x_j^o)(x_j^{oT}\mathbf{V}x_i^o)}{\lambda_i^o - \lambda_j^o} \tag{2.14}$$

The special property of symmetric nature of $\mathbf{A}^o$ in proceeding from eqs 2.6 and 2.7 to eqs 2.10 and 2.14 is reflected in the orthonormality of the x^o's. To be sure, the use of that property could have been delayed in the derivation, as later in section 5, by using a resolvent formalism, but at some point the symmetric nature of $\mathbf{A}$ was used to obtain eqs 2.10 and 2.14 for $\lambda_i^{(1)}$ and $\lambda_i^{(2)}$.

We consider next the case of a nonsymmetric matrix $\mathbf{A}$, which in our case will be the **GF** matrix. For a nonsymmetric $\mathbf{A}$ there is a very simple expedient which permits $\lambda_i^{(1)}$, $\lambda_i^{(2)}$, and higher $\lambda_i^{(n)}$ to be determined immediately from the series of equations such as eqs 2.5−2.7 and higher orders: A reciprocal basis set y_i^o, which forms a biorthogonal set of eigenvectors,[16−18] is introduced, and we again use a linear normalization for x_i

$$y_j^{oT}x_i^o = \delta_{ij},\ y_i^{oT}x_i = 1, \text{ so that } y_i^{oT}x_i^{(1)} = y_i^{oT}x_i^{(2)} = \cdots = 0 \tag{2.15}$$

where each y_i^{oT} is a left eigenvector of $\mathbf{A}^o$ (which in our case is the unperturbed **GF** matrix, $\mathbf{G}^o\mathbf{F}^o$), the x_i^o remain the right eigenvectors. As such, the y_i^o become the right eigenvectors of the transpose, $\mathbf{A}^{oT}$, and have the same eigenvalues λ_i^o as $\mathbf{A}^o$:[16,17]

$$\mathbf{A}^T y_i^o = \lambda_i^o y_i^o \tag{2.16}$$

Operating on the left of eqs 2.6 and 2.7 by y_i^{oT}, instead of x_i^{oT}, leads to the counterpart of eqs 2.10 and 2.11:

$$\lambda_i^{(1)} = y_i^{oT}\mathbf{V}x_i^o = (\mathbf{Y}^{oT}\mathbf{V}\mathbf{X})_{ii} \qquad \lambda_i^{(2)} = y_j^{oT}\mathbf{V}x_i^{(1)} \tag{2.17}$$

Operating on the left of eq 2.6 by y_j^{oT} instead of by x_j^{oT}, $j \neq i$, together with the last half of eq 2.17, leads to the counterpart of eq 2.14

$$\lambda_i^{(2)} = \sum_{j\neq i} \frac{(y_i^{oT}\mathbf{V}x_j^o)(y_j^{oT}\mathbf{V}x_i^o)}{\lambda_i^o - \lambda_j^o} \equiv \sum_{j\neq i} \frac{(\mathbf{Y}^{oT}\mathbf{V}\mathbf{X})_{ij}(\mathbf{Y}^{oT}\mathbf{V}\mathbf{X})_{ji}}{\lambda_i^o - \lambda_j^o} \tag{2.18}$$

Diagonal and off-diagonal elements of the matrix $\mathbf{Y}^{oT}\mathbf{V}\mathbf{X}^o$ are seen to occur in eqs 2.17 and 2.18.

Two sets of eigenvectors x_i^o and y_i^{oT} are clearly needed,[18] whereas in the symmetric case only x_i^o and its transpose x_i^{oT} were required. In general, x_i^o and y_i^o are different vectors. For example, while x_i^o and y_j^{oT} are biorthogonal, x_i^o and x_j^{oT} are seen later not to be orthonormal. There is clearly, nevertheless, a strong parallelism between eqs 2.10 and 2.14 and eqs 2.17 and 2.18. This parallelism extends (as in section 5) equally to higher-order perturbation theory and to degenerate perturbation theory, and so the methods derived for the symmetric operators are again immediately adaptable to a nonsymmetric matrix or operator $\mathbf{A}$. Further, because of the special properties[19] of $\mathbf{F}$ and $\mathbf{G}$, eqs 2.17 and 2.18 can be simplified so that only one type of eigenvector, x^o or y^o, is needed per equation, instead of x^o and y^o.

3. Application to the GF Matrix

The above results are immediately applicable to the eigenvalue equation for the **GF** matrix

$$\mathbf{GFX} = \mathbf{X}\Lambda \text{ or } \mathbf{GFL} = \mathbf{L}\Lambda \tag{3.1}$$

where we have written the $\mathbf{X}$ as $\mathbf{L}$ in the second equation, using the notation that is standard in the molecular vibration literature.[1] In eq 3.1, Λ is a diagonal matrix consisting of the desired eigenvalues λ_i, and $\mathbf{X}$ is a matrix whose columns are the eigenvectors x_i. The λ_i are $4\pi^2\upsilon_i^2$, where the υ_i are the desired vibration frequencies of the molecule. Since $\mathbf{V}$ denotes $\Delta\mathbf{F}$ or $\Delta\mathbf{G}$ in the perturbation expressions in the preceding equations, eqs 2.17 and 2.18 are immediately applicable, and so one has the appropriate first or higher order or degenerate (section 5) perturbation theory.

These expressions can be simplified to reduce the two sets of eigenvectors, x_i^o and y_i^{oT}, to one set, using the specific property that the $\mathbf{G}$ and $\mathbf{F}$ matrices can be simultaneously diagonalized, following the ideas in ref 19. This latter diagonalization is related to current methods for obtaining the eigenvalues of the **GF** matrix based on a factorization of $\mathbf{G}$ to form the product of a matrix and its transpose.[20] In the following, we use the familiar notation of ref 1, $\mathbf{L}$, instead of $\mathbf{X}$. Following Wilson et al.,[19] the matrix $\mathbf{L}$ is introduced to diagonalize simultaneously the symmetric matrices $\mathbf{G}$ and $\mathbf{F}$, giving rise to the normal coordinates for the vibrations:

$$\mathbf{L}^T\mathbf{F}\mathbf{L} = \Lambda,\ \mathbf{L}^{oT}\mathbf{F}^o\mathbf{L}^o = \Lambda^o \tag{3.2}$$

$$\mathbf{L}^T\mathbf{G}^{-1}\mathbf{L} = \mathbf{I},\ \mathbf{L}^{oT}\mathbf{G}^{o-1}\mathbf{L}^o = \mathbf{I} \tag{3.3}$$

where $\mathbf{I}$ is the unit matrix. Equations 3.2 and 3.3 lead to

$$\mathbf{GFL} = \mathbf{L}\Lambda,\ \mathbf{L}\mathbf{L}^T = \mathbf{G} \tag{3.4}$$

and

$$\mathbf{G}^o\mathbf{F}^o\mathbf{L}^o = \mathbf{L}^o\Lambda^o,\ \mathbf{L}^o\mathbf{L}^{oT} = \mathbf{G}^o \tag{3.5}$$

Thereby, the $\mathbf{L}$ which diagonalizes simultaneously the symmetric matrices $\mathbf{F}$ and $\mathbf{G}^{-1}$ is clearly from eq 3.4 not an orthogonal matrix, i.e., $\mathbf{L}\mathbf{L}^T \neq \mathbf{I}$, and a similar remark applies from eq 3.5 to $\mathbf{L}^o$. We also note that the biorthogonality of $\mathbf{X}^o$ and $\mathbf{Y}^{oT}$, in section 2, i.e., $\mathbf{Y}^{oT}\mathbf{X}^o = \mathbf{I}$, now corresponds to the relation $(\mathbf{L}^o)^{-1}\mathbf{L}^o = \mathbf{I}$.

For the perturbation $\Delta\mathbf{G}$ of the $\mathbf{G}$ matrix, we have

$$\mathbf{V} = (\Delta\mathbf{G})\mathbf{F}^o \tag{3.6}$$

The terms in eqs 2.17 and 2.18 contain, in the notation of the present section, diagonal and off-diagonal elements of a matrix $(\mathbf{L}^o)^{-1}(\Delta\mathbf{G})\mathbf{F}^o\mathbf{L}^o$. If we rewrite the second half of eq 3.2, using eq 3.6, as $\mathbf{F}^o\mathbf{L}^o = (\mathbf{L}^{oT})^{-1}\Lambda^o$, then the $(\Delta\mathbf{G})\mathbf{F}^o\mathbf{L}^o$ in eqs 2.17 and 2.18 becomes $(\Delta\mathbf{G})(\mathbf{L}^{oT})^{-1}\Lambda^o$ and we obtain

$$(\mathbf{L}^o)^{-1}(\Delta\mathbf{G})\mathbf{F}^o\mathbf{L}^o = (\mathbf{L}^o)^{-1}(\Delta\mathbf{G})(\mathbf{L}^{oT})^{-1}\Lambda^o \tag{3.7}$$

Thereby, instead of the two sets of eigenvectors only the $(\mathbf{L}^o)^{-1}$ and its transpose $(\mathbf{L}^{oT})^{-1}$ appear. The $(\mathbf{L}^{oT})^{-1}$ symbol in eq 3.7 is less compact than its equivalent $\mathbf{Y}^o$ in section 2 (the right-hand side of eq 3.7 would now read $\mathbf{Y}^{oT}(\Delta\mathbf{G})\mathbf{Y}^o\Lambda^o$), though is more familiar in the vibrations literature.

Similarly, for the perturbation of the **F** matrix, **V** now denotes $\mathbf{G}^{o}\Delta\mathbf{F}$, so the elements of the matrix $(\mathbf{L}^{o})^{-1}\mathbf{G}^{o}(\Delta\mathbf{F})\mathbf{L}^{o}$ are needed for eqs 2.17 and 2.18. Using the expressions in eqs 3.5, $(\mathbf{L}^{o})^{-1}\mathbf{G}^{o} = \mathbf{L}^{oT}$. Thus,

$$(\mathbf{L}^{o})^{-1}\mathbf{G}^{o}(\Delta\mathbf{F})\mathbf{L}^{o} = \mathbf{L}^{oT}(\Delta\mathbf{F})\mathbf{L}^{o} \tag{3.8}$$

We may conclude from eqs 2.17 and 2.18 and eqs 3.7 and 3.8 that for the perturbation of the **G** matrix we have

$$\lambda_i = \lambda_i^{o} + [(\mathbf{L}^{o})^{-1}(\Delta\mathbf{G})(\mathbf{L}^{oT})^{-1}\Lambda^{o}]_{ii} + \sum_{j\neq i}\frac{[(\mathbf{L}^{o})^{-1}(\Delta\mathbf{G})(\mathbf{L}^{oT})^{-1}\Lambda^{o}]_{ij}[(\mathbf{L}^{o})^{-1}(\Delta\mathbf{G})(\mathbf{L}^{oT})^{-1}\Lambda^{o}]_{ji}}{\lambda_i^{o} - \lambda_j^{o}} \tag{3.9}$$

For the perturbation of the **F** matrix we have

$$\lambda_i = \lambda_i^{o} + [\mathbf{L}^{oT}(\Delta\mathbf{F})\mathbf{L}^{o}]_{ii} + \sum_{j\neq i}\frac{[\mathbf{L}^{oT}(\Delta\mathbf{F})\mathbf{L}^{o}]_{ij}[\mathbf{L}^{oT}(\Delta\mathbf{F})\mathbf{L}^{o}]_{ji}}{\lambda_i^{o} - \lambda_j^{o}} \tag{3.10}$$

Equation 3.10 for the $\Delta\mathbf{G}$ case has been applied to the calculation of the frequencies of many isotopomers of ozone in a treatment of mass-independent isotope effects.[7] (For that purpose eqs 2.17 and 2.18 sufficed, with $\mathbf{V} = \Delta\mathbf{G}$.)[7] It reduced the errors of calculating some twenty-six unknown frequencies of the isotopomers to about 1 cm^{-1} (30 GHz). The second-order term was needed only for the asymmetric isotopomers XYZ, as the first-order expression vanished in the unusual formalism used. This accuracy sufficed for the kinetic purposes needed, namely for the densities of states and zero-point energies of the ozone isotopomers.

Some comment on why the first-order perturbation vanished is perhaps of interest because of its novelty, though this point is not immediately relevant for the present paper: The unperturbed **G** matrix, $\mathbf{G}^{o}$, used in ref 7 was not that of an actual symmetric bent molecule XYX, but rather was that of a fictitious one XYZ containing the masses of X, Y, and Z but for which one **G**-matrix element was deleted, G_{13} ($= G_{31}$) in the notation of ref 3, p 243. This deletion permitted the same factorization of the 3×3 **GF** matrix for the asymmetric XYZ into the two blocks, 2×2 and 1×1, as that found for a symmetric molecule, but with the actual masses. This omitted G_{13} served as the perturbation **V**. Using symmetry arguments, the first-order perturbation was shown to vanish. An off-diagonal matrix element containing G_{13} did not vanish, again by symmetry, and so the second-order perturbation term was proportional to $(G_{13})^2$. In the work it was assumed that isotopic *shifts* are insensitive to anharmonicities.

It is useful, for comparison with the results derived in the next section, where a presymmetrization is used, to rewrite eq 3.10 as a trace, the trace meaning here that the *ii*th diagonal element is selected.[21] We then have

$$\lambda_i = \lambda_i^{o} + \mathrm{tr}_{ii}(\mathbf{L}^{o})^{-1}(\Delta\mathbf{G})(\mathbf{L}^{oT})^{-1}\Lambda^{o} + \mathrm{tr}_{ii}[(\mathbf{L}^{o})^{-1}(\Delta\mathbf{G})(\mathbf{L}^{oT})^{-1}\Lambda^{o}]^{T}\mathbf{S}(\mathbf{L}^{o})^{-1}(\Delta\mathbf{G})(\mathbf{L}^{oT})^{-1}\Lambda^{o} \tag{3.11}$$

where $\mathbf{S} = \mathbf{Q}^{o}(\mathbf{G}^{o}\mathbf{F}^{o} - \lambda_i)^{-1}\mathbf{Q}^{o}$, $\mathbf{Q}^{o}$ being an operator which projects onto the subspace that is the orthogonal complement to x_i^{o}, i.e., to $(\mathbf{L}^{o})_i$.

Because Λ^{o} and **S** commute (both are functions of $\mathbf{G}^{o}\mathbf{F}^{o}$) and because the trace is invariant to a cyclic permutation, eq 3.11 can be rewritten in a more symmetric way as

$$\lambda_i = \lambda_i^{o} + \mathrm{tr}_{ii}\Lambda^{o1/2}(\mathbf{L}^{o})^{-1}(\Delta\mathbf{G})(\mathbf{L}^{oT})^{-1}\Lambda^{o1/2} + \mathrm{tr}_{ii}[\Lambda^{o1/2}(\mathbf{L}^{o})^{-1}(\Delta\mathbf{G})(\mathbf{L}^{oT})^{-1}\Lambda^{o1/2}]^{T}\mathbf{S} \times [\Lambda^{o1/2}(\mathbf{L}^{o})^{-1}(\Delta\mathbf{G})(\mathbf{L}^{oT})^{-1}\Lambda^{o1/2}] \tag{3.12}$$

to which we return in the next section.

4. An Alternative Approach, Presymmetrization

We next consider a different approach, a presymmetrization, to the perturbation theory for **GF**, one which has been used in the literature[1−3] for $\Delta\mathbf{F}$, though not to our knowledge, for $\Delta\mathbf{G}$. This presymmetrization for $\Delta\mathbf{F}$ is a consequence of introducing equations such as eqs 3.2 and 3.3, in contrast with $\Delta\mathbf{G}$. This alternative method for $\Delta\mathbf{F}$ is based on converting the problem to that of diagonalization of a related symmetric matrix.[1−3] In that case the standard perturbation formalism for symmetric matrices immediately applies. This method has proved to be eminently practical for the **F** matrix, and a related approach is used in the literature for diagonalizing the **GF** matrix itself, without the perturbation aspects.[20] To this end, a matrix **C** is defined[2] relating **L** to $\mathbf{L}^{o}$

$$\mathbf{L} = \mathbf{L}^{o}\mathbf{C} \tag{4.1}$$

where **C** is an orthogonal matrix because it transforms one set of normal coordinates to another.[2] The eigenvalue equation, eq 3.1, after multiplying on the left by $(\mathbf{L}^{o})^{-1}$ can then be written as

$$[(\mathbf{L}^{o})^{-1}\mathbf{G}\mathbf{F}\mathbf{L}^{o}]\mathbf{C} = \mathbf{C}\Lambda \tag{4.2}$$

This equation is converted to a symmetrized operator equation using

$$(\mathbf{L}^{o})^{-1}\mathbf{G}\mathbf{F}\,\mathbf{L}^{o} = (\mathbf{L}^{o})^{-1}\mathbf{G}^{o}\mathbf{F}^{o}\mathbf{L}^{o} + (\mathbf{L}^{o})^{-1}\mathbf{G}^{o}(\Delta\mathbf{F})\mathbf{L}^{o} = \Lambda^{o} + \mathbf{L}^{oT}(\Delta\mathbf{F})\mathbf{L}^{o} \tag{4.3}$$

where eqs 3.5 and 3.8 were introduced. Thus, the diagonalization of $(\mathbf{L}^{o})^{-1}\mathbf{G}\mathbf{F}\mathbf{L}^{o}$, obtained by solving eq 4.2, is seen in eq 4.3 to be equivalent to the diagonalization of a symmetric matrix, namely the matrix on the extreme right-hand side of eq 4.3:

$$[\Lambda^{o} + \mathbf{L}^{oT}(\Delta\mathbf{F})\mathbf{L}^{o}]\mathbf{C} = \mathbf{C}\Lambda \tag{4.4}$$

The standard perturbation theory for symmetric matrices is then used to obtain a perturbation expression for Λ for this case.[1−3]

For this same approach, but applied to perturbation of the **G**-matrix, $\Delta\mathbf{G}$, we consider

$$(\mathbf{L}^{o})^{-1}\mathbf{G}\mathbf{F}^{o}\mathbf{L}^{o} = (\mathbf{L}^{o})^{-1}\mathbf{G}^{o}\mathbf{F}^{o}\mathbf{L}^{o} + (\mathbf{L}^{o})^{-1}(\Delta\mathbf{G})\mathbf{F}^{o}\mathbf{L}^{o} = \Lambda^{o} + (\mathbf{L}^{o})^{-1}(\Delta\mathbf{G})(\mathbf{L}^{oT})^{-1}\Lambda^{o} \tag{4.5}$$

where eq 3.5 was again introduced. Thus, from eq 4.2 we have

$$[\Lambda^{o} + (\mathbf{L}^{o})^{-1}(\Delta\mathbf{G})(\mathbf{L}^{oT})^{-1}\Lambda^{o}]\mathbf{C} = \mathbf{C}\Lambda \tag{4.6}$$

However, the matrix on the left of eq 4.6 is not symmetric, in contrast with the $\Delta\mathbf{F}$ case, eq 4.4. (Perhaps for this reason, a treatment parallel to that for eq 4.4 does seem to have appeared in the literature.) Nevertheless, eq 4.6 can readily be written in a symmetric form by minor manipulation: Multiplying on the left by $\Lambda^{o1/2}$ we have

$$[\Lambda^{o} + \{(\mathbf{L}^{oT})^{-1}\Lambda^{o1/2}\}^{T}(\Delta\mathbf{G})(\mathbf{L}^{oT})^{-1}\Lambda^{o1/2}]\mathbf{D} = \mathbf{D}\Lambda \tag{4.7}$$

where $\mathbf{D} = \Lambda^{o1/2}\mathbf{C}$. The matrix on the left-hand side of eq 4.7 is now symmetric, and so conventional perturbation theory can be applied. Using standard perturbation theory, the perturbation result obtained from eq 4.7 is readily seen to be the same as that given in eq 3.12.

We next comment briefly on a perturbation theory for $\Delta\mathbf{G}$ in refs 5 and 6. The transpose matrix **FG** was considered there, as an alternative to considering left eigenvectors of **GF**. The matrix equation and that for the unperturbed problem are now, in present notation,

$$\mathbf{FG}(\mathbf{L}^{T})^{-1} = (\mathbf{L}^{T})^{-1}\Lambda,\ \mathbf{F}^{o}\mathbf{G}^{o}(\mathbf{L}^{oT})^{-1} = (\mathbf{L}^{oT})^{-1}\Lambda^{o} \tag{4.8}$$

They noted that

$$\mathbf{L}^{oT}\mathbf{F}^{o}\mathbf{G}^{o}(\mathbf{L}^{oT})^{-1} = \Lambda^{o} \tag{4.9}$$

and so

$$\mathbf{L}^{oT}\mathbf{F}^{o}\mathbf{G}(\mathbf{L}^{oT})^{-1} = \Lambda^{o} + \mathbf{L}^{oT}\mathbf{F}^{o}(\Delta\mathbf{G})(\mathbf{L}^{oT})^{-1} \tag{4.10}$$

Multiplying the first equation in eq 4.8 by $\mathbf{L}^{oT}$ on the left, and introducing eq 4.10, we have, because $\mathbf{F} = \mathbf{F}^{o}$,

$$[\Lambda^{o} + \mathbf{L}^{oT}\mathbf{F}^{o}(\Delta\mathbf{G})(\mathbf{L}^{oT})^{-1}]\mathbf{L}^{oT}(\mathbf{L}^{T})^{-1} = \mathbf{L}^{oT}(\mathbf{L}^{T})^{-1}\Lambda \tag{4.11}$$

It was then tacitly assumed,[5,6] perhaps based on the known result in ref 4, that the standard first-order perturbation theory for symmetric matrices can be used for the nonsymmetric matrix $\Lambda^{o} + \mathbf{L}^{oT}\mathbf{F}^{o}(\Delta\mathbf{G})(\mathbf{L}^{oT})^{-1}$ in eq 4.11. The final answer is indeed correct. The more rigorous derivation is, nevertheless, given in ref 4 or in either of the derivations given above.

5. Higher Order and Degenerate Perturbation Theory

The results obtained in sections 2 and 3 are immediately extended in several respects using a resolvent formalism. We comment briefly on the extension here, denoting by $\mathbf{P}^{o}$ and $\mathbf{P}$ the projection operators of a vector onto the unperturbed space Ω^{o} of any given eigenvalue λ^{o} and onto the perturbed space Ω, respectively. The projection operator complementary to $\mathbf{P}^{o}$ is denoted by the customary $\mathbf{Q}^{o}$. We consider first the nondegenerate case. The space Ω^{o} then consists of vectors proportional to x_i^{o} and Ω consists of vectors proportional to x_i. In the interests of brevity we revert to the $\mathbf{A}^{o}$, $\mathbf{X}$ and $\mathbf{Y}^{T}$ notation. The desired eigenvalue λ_i is obtained from $\mathbf{P}$ from

$$(\mathbf{A}^{o} + \mathbf{V})\mathbf{P} = \lambda_i\mathbf{P} \tag{5.1}$$

and $\mathbf{P}$ is given in the resolvent formalism in terms of $\mathbf{P}^{o}$, $\mathbf{Q}^{o}$, and $\mathbf{V}$ by[11-13,22]

$$\mathbf{P} = \mathbf{P}^{o} + \sum_{n=1}^{\infty}(-1)^{n-1}\Sigma\mathbf{S}^{k_1}\mathbf{V}\mathbf{S}^{k_2}\mathbf{V}\cdots\mathbf{V}\mathbf{S}^{k_{n+1}} \tag{5.2}$$

where the restriction on the second sum is $k_i \geq 0$, $\sum_1^{n+1}k_i = n$, and where, as noted earlier,

$$\mathbf{S}^{\circ} = -\mathbf{P}^{o} \text{ and } \mathbf{S} = \mathbf{Q}^{o}\frac{1}{\mathbf{A}^{o} - \lambda_i^{o}}\mathbf{Q}^{o} \quad \text{for } \mathbf{S} \neq \mathbf{S}^{o} \tag{5.3}$$

No assumption regarding symmetric nature of $\mathbf{A}^{o}$ is made in obtaining eqs 5.2 and 5.3.

The perturbed eigenvector x_i corresponding to the eigenvalue λ_i is obtained from x_i^{o} via $\mathbf{P}$:

$$x_i = \mathbf{P}x_i^{o} \tag{5.4}$$

As a normalization for x_i we again use the linear relation eq 2.8, and again use the biorthogonal set of eigenvectors x_i^{o} and y_j^{o} introduced earlier.

The projection operators $\mathbf{P}^{o}$ and $\mathbf{Q}^{o}$ operating on an arbitrary vector Φ can be written in terms of this biorthogonal basis set as

$$\mathbf{P}^{o}\Phi = (y_i^{oT}\Phi)x_i^{o},\ \mathbf{Q}^{o}\Phi = \sum_{j\neq i}(y_j^{oT}\Phi)x_j^{o} \tag{5.5}$$

where the parentheses denote a scalar product. One can verify, thereby, that $\mathbf{P}^{o2} = \mathbf{P}^{o}$, $\mathbf{Q}^{o2} = \mathbf{Q}^{o}$, and $\mathbf{P}^{o}\mathbf{Q}^{o} = \mathbf{Q}^{o}\mathbf{P}^{o} = 0$. Applying y_j^{oT} to the left of eq 5.1 and using eq 5.5 we obtain

$$\lambda_i = x_i^{o} + y_i^{oT}\mathbf{V}\mathbf{x} \tag{5.6}$$

Using eqs 5.2 and 5.5, the first few terms in the series for the perturbed eigenvalues, namely eqs 2.17 and 2.18, are again obtained, but now the higher-order terms are also obtained.

The extension to the degenerate case is also straightforward using the resolvent formalism applied to the nonsymmetric case: Using the projection operator $\mathbf{P}$ operating on the vector $x_{i\alpha}^{o}$, we have

$$\mathbf{A}\mathbf{P}x_{i\alpha}^{o} = (\mathbf{A}^{o} + \mathbf{V})\mathbf{P}x_{i\alpha}^{o} = \lambda_{i\alpha}\mathbf{P}x_{i\alpha}^{o} \tag{5.7}$$

where α is a degeneracy index ($\alpha = 1, ..., m$) in the unperturbed subspace Ω^{o}. Thereby,[13]

$$\mathbf{P}^{o}\mathbf{A}\mathbf{P}\mathbf{P}^{o}x_{i\alpha}^{o} = \lambda_{i\alpha}\mathbf{P}^{o}\mathbf{P}\mathbf{P}^{o}x_{i\alpha}^{o} \tag{5.8}$$

$\mathbf{P}^{o}\mathbf{P}\mathbf{P}^{o}$ is positive definite, and eq 5.8 is a generalized eigenvalue equation. On multiplication on the left by a member of the reciprocal basis set $y_{j\beta}^{oT}$, one obtains a series of equations, such that det $(\mathbf{P}^{o}\mathbf{A}\mathbf{P}\mathbf{P}^{o} - \lambda_{i\alpha}\mathbf{P}^{o}\mathbf{P}\mathbf{P}^{o}) = 0$ in this representation. Introducing the expressions given earlier for the expansion of $\mathbf{P}$, one obtains the $\lambda_{i\alpha}$ to the desired order of approximation. Another expansion has also been given.[13]

6. Discussion

In the present article, prompted by a different problem,[7,8] a simple extension is given of early literature results,[1-4] using a well established formalism.[9-12] A simple direct method for diagonalization of nonsymmetric matrices has been applied to the perturbation theory for both $\Delta\mathbf{G}$ and $\Delta\mathbf{F}$. Independently of this derivation, a presymmetrization of the **GF** matrix is introduced for the $\Delta\mathbf{G}$ case (it had been for the $\Delta\mathbf{F}$ case), so permitting as an alternative the standard method for perturbation of symmetric matrices to be applied.

Acknowledgment. It is a pleasure to acknowledge the support of the National Science Foundation. It is a pleasure to congratulate Bill Miller on the occasion of his 60th birthday. Bill's skill, insight, and mastery of mathematics and physics applied to chemical dynamics and collisions are a pleasure to behold and serve as a remarkable trademark of his work. There is perhaps also an added personal touch, in that these brief comments on **GF** perturbation theory deal with a tiny corner of the classic book by Bill's eminent mentor, E. Bright Wilson,

Jr., whose contributions have been a source of stimulation to numerous researchers.

References and Notes

(1) Wilson, E. B., Jr.; Decius, J. C.; Cross, P. C. Molecular Vibrations. *The Theory of Infrared and Raman Vibrational Spectra*; Mc-Graw-Hill: New York, 1955; Dover Publications: New York, 1980.

(2) Mills, I. M. *J. Mol. Spectrosc.* **1960**, *5*, 334. Mills, I. M. *J. Mol. Spectrosc.* **1965**, *17*, 164.

(3) Califano, S. *Vibrational States*; Wiley: New York, 1976; pp. 95−96.

(4) Edgell, W. F. *J. Chem. Phys.* **1945**, *13*, 539. See also Knyzazev, D. A.; Denisov, D. A.; Ivlev, A. A. *Opt. Spectrosc.* **1970**, *29*, 432 (Engl. Transl. of *Opt. Spektrosk*.)

(5) Wilson, E. B., Jr.; Decius, J. C.; Cross, P. C. Molecular Vibrations. *The Theory of Infrared and Raman Vibrational Spectra*; Mc-Graw-Hill: New York, 1955; Dover Publications: New York, 1980; pp 188−189.

(6) Califano, S. *Vibrational States*; Wiley: New York, 1976; pp. 95−96, p 224

(7) Hathorn, B. C.; Marcus, R. A. *J. Phys. Chem.*, in press. Gao, Y. Q.; Marcus, R. A. (submitted).

(8) Hathorn, B. C.; Marcus, R. A. *J. Chem. Phys.* **1999**, *111*, 4087.

(9) Kato, T. *Perturbation Theory for Linear Operators*; Springer: New York, 1980; pp. 62−83.

(10) Messiah, A. *Quantum Mechanics*; North-Holland: Amsterdam, 1963; Vol. 2, p 687.

(11) Kato, T. *Prog. Theor. Phys.* **1949**, *4*, 514, which is formulated in terms of a self-adjoint operator but that assumption is not needed for **P** and **AP**.

(12) Bloch, C. *Nucl. Phys.* **1958**, *6*, 329 (same remark as for ref 11).

(13) Messiah, A. *Quantum Mechanics*; North-Holland: Amsterdam, 1963; Vol. 2; pp 718−720. Although the argument uses a self-adjoint operator, the Hamiltonian, as an example it is equally applicable to a nonsymmetric matrix.

(14) Baym, G. *Lectures on Quantum Mechanics*; W. A. Benjamin: New York, 1969; p 226.

(15) Sakurai, J. J. *Modern Quantum Mechanics*; Addison-Wesley: New York, 1996; p 290.

(16) Friedman, B. *Principles and Techniques of Applied Mathematics*; John Wiley: New York, 1956; pp 85, 111−113.

(17) Kato, T. *Perturbation Theory for Linear Operators*; Springer: New York, 1980; pp 12, 43−44.

(18) Actually, the introduction of a biorthogonal set of vectors is made tacitly in the **GF** literature whenever the symbols $\mathbf{L}$ and $\mathbf{L}^{-1}$ are introduced.

(19) Wilson, E. B., Jr.; Decius, J. C.; Cross, P. C. Molecular Vibrations. *The Theory of Infrared and Raman Vibrational Spectra*; Mc-Graw-Hill: New York, 1955; Dover Publications: New York, 1980; p 309.

(20) Hedberg, L.; Mills, I. M. *J. Mol. Spectrosc.* **1993**, *160*, 117. Mills, I. M.; Aldous, J. *Spectrochim. Acta* **1962**, *18*, 1073. A related factorization is found in Miyazawa, T. *J. Chem. Phys.* **1958**, *29*, 246.

(21) If we introduced the operator $\mathbf{S}^{o}$, defined in eq 5.3, on either side of the matrices in the trace in eq 3.13, the restriction that the *ii*th element be selected can be omitted, but this modification is a minor detail.

(22) Kato, T. *Perturbation Theory for Linear Operators*; Springer: New York, 1980; pp 74−84.