Theor Chem Acc (2000) 104:163–166
DOI 10.1007/s002140000124

Theoretical
Chemistry Accounts

*Letter*

# Unitary perturbation theory: a generalization of two-by-two rotations

**I. Mayer**

Institute of Chemistry, Chemical Research Center, Hungarian Academy of Sciences, P.O. Box 17, 1525 Budapest, Hungary

Received: 15 November 1999 / Accepted: 30 January 2000 / Published online: 19 April 2000

**Abstract.** A generalization of the two-by-two rotation technique is proposed, permitting a whole row (column) of the matrix to be treated simultaneously. The method is based on the explicit analytical evaluation of the matrix exponent representing a symmetric combination of the individual rotations. Besides constructing the unitary transformation matrices, a new orthogonalization algorithm is also proposed. The resulting "unitary perturbation theory" and orthogonalization method may be useful in different areas.

**Key words:** Unitary transformations – Two-by-two rotations – Perturbation theory

## 1 Introduction

The method of two-by-two rotations is used in different areas of quantum chemistry in order to calculate iteratively functions satisfying some desired requirements: the classical Jacobi matrix diagonalization algorithm is a well-known prototype. When doing two-by-two rotations, the orthogonality of the functions considered is conserved in every step, so their sequence produces a unitary global transformation matrix. This scheme is, however, hardly applicable in analytical studies because of the inherent asymmetry by which one treats the different functions.

In many cases one has an acceptable zero-order estimate of the vector required and seeks improved approximations for it – this is the essential starting point for perturbational approaches. Standard perturbation methods suffer from the drawback that the orthogonality of the different functions is lost after the perturbative correction of the solution considered is introduced. (Orthogonality is restored only in infinite order and if the perturbative corrections for all the vectors are computed.)

The purpose of the present letter is to present a simple formalism which combines, as far as possible, the advantages of these schemes. It is based on the explicit analytical evaluation of the matrix exponent representing the symmetric combination of the individual rotations.

## 2 Theory

Let us consider the "reference" state denoted "0" and an "excited" one denoted "$i$", and let the respective state vectors (wave functions) be $\psi_i$ and $\psi_0$. The two-by-two rotation leading to functions

$$\begin{aligned}\psi'_0 &= \psi_0 \cos\varphi_i + \psi_i \sin\varphi_i \\ \psi'_i &= -\psi_0 \sin\varphi_i + \psi_i \cos\varphi_i\end{aligned} \tag{1}$$

is described by the rotation matrix

$$\mathbf{U}_i = \begin{pmatrix} \cos\varphi_i & -\sin\varphi_i \\ \sin\varphi_i & \cos\varphi_i \end{pmatrix} . \tag{2}$$

It is easy to see that the unitary matrix, $\mathbf{U}_i$, can be presented as a matrix exponent of the antisymmetric (anti-Hermitian) matrix, $\mathbf{A}_i$:

$$\mathbf{U}_i = \exp(\mathbf{A}_i); \qquad \mathbf{A}_i = \begin{pmatrix} 0 & -\varphi_i \\ \varphi_i & 0 \end{pmatrix} . \tag{3}$$

If there are more than two states, the matrices $\mathbf{U}_i$ and $\mathbf{A}_i$ describing a given two-by-two rotation should be defined in a space of larger dimensions: they can be written down by trivial padding with zero and unit submatrices in the respective subspaces. We will actually need the padded form of $\mathbf{A}_i$ only:

$$\mathbf{A}_i = \begin{pmatrix} 0 & 0 & \dots & -\varphi_i & \dots & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 \\ \dots & & & & & \dots \\ \varphi_i & 0 & \dots & 0 & \dots & 0 \\ \dots & & & & & \dots \\ 0 & 0 & \dots & 0 & \dots & 0 \end{pmatrix} . \tag{4}$$

If there are several "excited" states to be considered, the sequential two-by-two rotations will involve their nonsymmetric treatment, which is hardly adequate if an analytical investigation is desired. However, one may treat them on an equal footing if, instead of performing the (noncommutative) multiplication of matrices $\mathbf{U}_i$, one performs the summation of their “logarithms” $\mathbf{A}_i$ and seeks the overall unitary transformation in the form

$$\mathbf{U} = \exp(\mathbf{A}) \ , \tag{5}$$

with

$$\mathbf{A} = \begin{pmatrix} 0 & -\varphi_1 & -\varphi_2 & \dots & -\varphi_N \\ \varphi_1 & 0 & 0 & \dots & 0 \\ \varphi_2 & 0 & 0 & \dots & 0 \\ \dots & & & & \dots \\ \varphi_N & 0 & 0 & \dots & 0 \end{pmatrix} . \tag{6}$$

Matrix $\mathbf{U}$ defined in this manner describes the simultaneous effect of the individual rotations in a completely symmetric manner. In order for Eqs. (5) and (6) to be actually useful, one has to sum over the series expansion of the matrix exponent

$$\exp(\mathbf{A}) = \sum_{n=0}^{\infty} \frac{1}{n!} \mathbf{A}^n = \mathbf{1} + \mathbf{A} + \frac{1}{2}\mathbf{A}^2 + \frac{1}{6}\mathbf{A}^3 \dots \ . \tag{7}$$

We define the quantities

$$\varphi = \left( \sum_i \varphi_i^2 \right)^{\frac{1}{2}} \tag{8}$$

and

$$x_i = \frac{\varphi_i}{\varphi} \tag{9}$$

and the auxiliary matrices expressed in terms of $x_i$-s:

$$\mathbf{A}' = \begin{pmatrix} 0 & -x_1 & -x_2 & \dots & -x_N \\ x_1 & 0 & 0 & \dots & 0 \\ x_2 & 0 & 0 & \dots & 0 \\ \dots & & & & \dots \\ x_N & 0 & 0 & \dots & 0 \end{pmatrix} \tag{10}$$

and

$$\mathbf{B} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & x_1^2 & x_2 x_1 & \dots & x_N x_1 \\ 0 & x_1 x_2 & x_2^2 & \dots & x_N x_2 \\ 0 & x_1 x_3 & x_2 x_3 & \dots & x_N x_3 \\ \dots & & & \dots & \\ 0 & x_1 x_N & x_2 x_N & \dots & x_N^2 \end{pmatrix} . \tag{11}$$

Note that by definition

$$\sum_i x_i^2 = 1 \ . \tag{12}$$

Using this result, one can check that $\mathbf{B}^2 = \mathbf{B}$, so the Hermitian matrix $\mathbf{B} = \mathbf{B}^\dagger$ is a projection matrix. As it is easy to see, it projects on the two-dimensional subspace of the orthonormalized vectors

$$\mathbf{c}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \dots \\ 0 \end{pmatrix} ; \quad \mathbf{c}_2 = \begin{pmatrix} 0 \\ x_1 \\ x_2 \\ \dots \\ x_N \end{pmatrix} . \tag{13}$$

Now, it is easy to check by induction that

$$\begin{aligned} \mathbf{A}^{2n} &= (-1)^n \varphi^{2n} \mathbf{B} \\ \mathbf{A}^{2n+1} &= (-1)^n \varphi^{2n+1} \mathbf{A}' \ . \end{aligned} \tag{14}$$

Using this recurrence formula, one can trivially sum up the matrix exponent $\exp(\mathbf{A})$ as

$$\mathbf{U} = \begin{pmatrix} \cos\varphi & -x_1 \sin\varphi & -x_2 \sin\varphi & \dots & -x_i \sin\varphi & \dots \\ x_1 \sin\varphi & 1 + x_1^2(\cos\varphi - 1) & x_2 x_1(\cos\varphi - 1) & \dots & x_i x_1(\cos\varphi - 1) & \dots \\ x_2 \sin\varphi & x_1 x_2(\cos\varphi - 1) & 1 + x_2^2(\cos\varphi - 1) & \dots & x_i x_2(\cos\varphi - 1) & \dots \\ \dots & & \dots & \dots & \dots & \\ x_i \sin\varphi & x_1 x_i(\cos\varphi - 1) & x_2 x_i(\cos\varphi - 1) & \dots & 1 + x_i^2(\cos\varphi - 1) & \dots \\ \dots & & \dots & \dots & & \\ x_N \sin\varphi & x_1 x_N(\cos\varphi - 1) & x_2 x_N(\cos\varphi - 1) & \dots & x_i x_N(\cos\varphi - 1) & \dots \end{pmatrix} . \tag{15}$$

This means that the reference vector $\mathbf{c}_1$ defined in Eq. (13) transforms to a linear combination

$$\mathbf{c}_1' = \cos\varphi \mathbf{c}_1 + \sin\varphi \mathbf{c}_2 \tag{16}$$

where the angle $\varphi$ is given by Eq. (8). However, Eq. (15) also gives the unitary transformation for any value of $\varphi$, not only for that defined in Eq. (8); one may therefore consider $\varphi$ in Eq. (16) as an optimization parameter and determine the best (e.g. variational) linear combination of vectors $\mathbf{c}_1$ and $\mathbf{c}_2$. (Parameters $x_i$ should not be recalculated.)

This scheme can also be utilized as an orthogonalization algorithm (replacing, for example, Schmidt orthogonalization) when the task is to generate a full set of orthonormalized vectors orthogonal to some vector

$$\mathbf{q} = \begin{pmatrix} q_0 \\ q_1 \\ q_2 \\ \dots \\ q_N \end{pmatrix} . \tag{17}$$

Assuming that vector $\mathbf{q}$ is normalized ($\sum_{i=0}^{N} |q_i|^2 = 1$), one can present it in the form of the first column of Eq. (15) by introducing the notations $\cos\varphi = q_0$, $\sin\varphi = (\sum_{i=1}^{N} |q_i|^2)^{\frac{1}{2}}$, $x_i = q_i / \sin\varphi$ $(i = 1, 2, \dots, N)$. Then the other columns of Eq. (15) give the orthonormalized vectors requested.

## 3 Illustrations

We think the scheme proposed may appear useful in both analytical and numerical investigations. To illustrate its viability, we performed some very simple calculations. Table 1 contains some results for the

**Table 1.** Some results calculated for the helium full configuration interaction (*CI*) problem using a $10/5s2p1d$ basis and Møller–Plesset (*MP*) perturbation theory (*PT*) and Epstein–Nesbet (*EN*) partitioning

| Method | Energy value |
|---|---|
| Self-consistent-field energy | −2.861673 |
| Second-order PT – MP2 | −2.889217 |
| Fourth-order PT – MP4 | −2.895485 |
| Second-order PT – EN2 | −2.894155 |
| Expectation value for the first-order EN perturbational wave function | −2.895355 |
| Expectation value – present scheme | −2.895360 |
| Expectation value – present scheme with optimized $\varphi$ | −2.895402 |
| Full CI | −2.895719 |

| Method | Overlap with the exact eigenvector |
|---|---|
| Hartree–Fock wave function | 0.995975 |
| First-order EN perturbational wave function | 0.999926 |
| Present scheme | 0.999927 |
| Present scheme with optimized $\varphi$ | 0.999939 |

correlational problem of the He atom in the small $10/5s2p1d$ basis defined in Ref. [1] – including the full configuration interaction (CI) solution, as a reference. It can be seen that the results of the present scheme, as it is applied to the CI matrix, are significantly better than those given by second-order Møller–Plesset (MP) perturbation theory (PT), but are only slightly better than those obtained by second-order PT using Epstein–Nesbet (EN) partitioning. (This is not the case in the $H_2$ example discussed later.)

In the case of problems for which the conditions of the applicability of the traditional PT are also fulfilled, the approximation provided by the present scheme can be expected to be about as good as those of ordinary second-order PT. If, however, there are some large off-diagonal matrix elements making the usual PT inapplicable, the present algorithm still works acceptably, which is again not surprising as it utilizes the exact solutions of the individual two-by-two rotations, i.e., the individual rotational angles, $\varphi_i$, are determined by solving exactly the respective two-by-two eigenvalue problems. We shall illustrate this by two examples.

Some characteristic data collected for the following $10 \times 10$ matrix are displayed in Table 2.

$$\begin{pmatrix}
0 & 0.8 & 0.62 & 0.5 & 0.33 & 0.1 & 0 & 0.52 & -0.5 & 0.08 \\
0.8 & 1 & 0 & 0 & 0.41 & 0 & 0.4 & -0.21 & 0 & -0.18 \\
0.62 & 0 & 1.5 & 0 & 0 & 0 & -0.35 & 0 & 0 & 0 \\
0.5 & 0 & 0 & 2 & 0 & 0.15 & 0.09 & 0 & 0.32 & 0 \\
0.33 & 0.41 & 0 & 0 & 2.3 & 0 & -0.13 & 0 & 0 & 0.15 \\
0.1 & 0 & 0 & 0.15 & 0 & 3 & 0 & 0 & 0 & 0 \\
0 & 0.4 & -0.35 & 0.09 & -0.13 & 0 & 3.6 & 0 & 0 & 0 \\
0.52 & -0.21 & 0 & 0 & 0 & 0 & 0 & 4.1 & 0.09 & 0 \\
-0.5 & 0 & 0 & 0.32 & 0 & 0 & 0 & 0.09 & 4.4 & 0 \\
0.08 & -0.18 & 0 & 0 & 0.15 & 0 & 0 & 0 & 0 & 4.8
\end{pmatrix}$$

**Table 2.** Some characteristics calculated for the $10 \times 10$ matrix shown in the text

| Method | Energy value |
|---|---|
| Second order EN-type PT | −1.19605 |
| Expectation value for the first-order EN perturbational wave function | −0.60640 |
| Expectation value – present scheme | −0.67284 |
| Expectation value – present scheme with optimized $\varphi$ | −0.75704 |
| Exact eigenvalue | −0.77504 |

| Method | Overlap with the exact eigenvector |
|---|---|
| First-order EN perturbational wave function | 0.96950 |
| Present scheme | 0.98416 |
| Present scheme with optimized $\varphi$ | 0.99659 |

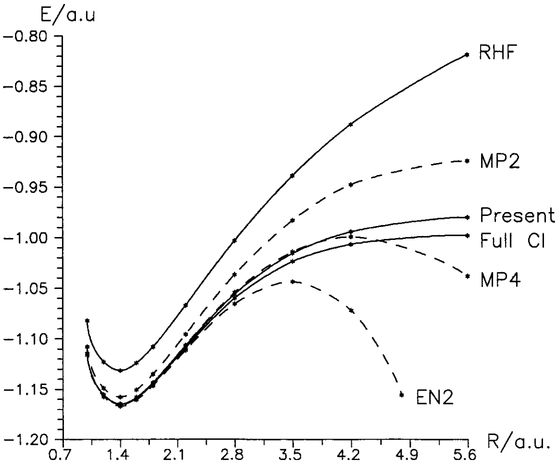

**Fig. 1.** Potential curves calculated for the $H_2$ molecule with the 6-31G** basis set

The results show that the present scheme gives quite good results, especially if the optimization of angle $\varphi$ is also performed.

Finally, the potential curves calculated for the $H_2$ molecule calculated in the 6-31G** basis set are displayed in Fig. 1; the present scheme is again applied to the correlation problem and the results are compared with those of the exact calculation (the full CI results) and different perturbational schemes. Inspection of the curves shows that both MP and EN PTs fail when the bond length increases, but our approximate scheme behaves quite acceptably.

*Acknowledgement*. The author acknowledges support from the Hungarian Research Fund (grant OTKA no. 29716).

## Reference

1. Mayer I, Vibók Á, Valiron P (1994) Chem Phys Lett 224: 166