JOURNAL OF COMPUTATIONAL PHYSICS 26, 236–242 (1978)

Note

A Partial Transformation for Application to Perturbation Theory Configuration Interaction

INTRODUCTION

In the application of first-order perturbation theory to the implementation of configuration interaction (CI) based on SCF orbitals, one is concerned only with double excitations of the type $ij \rightarrow uv$, where i, j are in the occupied set of orbitals and u, v are in the virtual set of orbitals. Single excitations do not contribute to the second-order energy correction for closed-shell systems.

If Slater determinants ϕ_k and ϕ_l differ by two spin orbitals such that spin orbitals i and j of Slater determinant ϕ_k are replaced by spin orbitals u and v of ϕ_l, the matrix element between ϕ_k and ϕ_l is given in general by

$$\langle \phi_k \mid \mathscr{H} \mid \phi_l \rangle = \mathscr{P}(\langle iu \mid \mid jv \rangle - \langle iv \mid \mid ju \rangle), \tag{1}$$

where

$$\langle iu \mid \mid jv \rangle = \iint \chi_i^*(1)\, \chi_u(1)(1/r_{12})\, \chi_j^*(2)\, \chi_v(2)\, d\tau_1\, d\tau_2 .^1$$

($\mathscr{P}$ is a permutation operator which brings ϕ_k and ϕ_l into maximum coincidence.) The integration implied in Eq. (1) is over space and spin coordinates; thus, spin integration can lead to a zero contribution from the second term. To evaluate these matrix elements, one need not complete the full four-index transformation from the starting atomic orbital basis set to the molecular orbital basis set [1, 2]; rather, the special subset $\langle iu \mid \mid jv \rangle$ of the completely transformed molecular orbital list is all that is required for the off-diagonal elements. The present paper describes an efficient procedure for achieving this special subset transformation (partial transformation).

The method described herein may also be of interest for Moller–Plesset theory calculations since the partial transformation method generates all the two-electron integrals needed for the second-order treatment. Pople *et al.* [2] have discussed the application of this perturbation theory approach and have shown for a series of small molecules that second-order Moller–Plesset theory describes a number of energy dependent properties better than conventional Hartree–Fock theory.

---

[1] This notation is similar to that of R. G. PARR, D. P. CRAIG, AND I. G. ROSS, *J. Chem. Phys.* **18** (1950), 1561, and is to be differentiated from the notation of R. J. BARTLETT AND D. M. SILVER, *Chem. Phys. Lett.* **29** (1974), 199.

0021-9991/78/0262–0236$02.00/0

Copyright © 1978 by Academic Press, Inc.
All rights of reproduction in any form reserved.

RATIONALE

This partial transformation method is an adaptation of a previously reported full transformation method [3]. An important element in the procedure is the use of partial summations [4]. If we let $\langle pq \mid \mid rs\rangle$ be the set of starting atomic basis integrals and $\langle iu \mid \mid jv\rangle$ be the final list, the following partial summations are necessary:

$$\text{(a)}\quad \langle iq \mid \mid rs\rangle = \sum_{p=1}^{N} c_{pi}\langle pq \mid \mid rs\rangle; \tag{2}$$

$$\text{(b)}\quad \langle iq \mid \mid js\rangle = \sum_{r=1}^{N} c_{rj}\langle iq \mid \mid rs\rangle; \tag{3}$$

$$\text{(c)}\quad \langle iu \mid \mid js\rangle = \sum_{q=1}^{N} c_{qu}\langle iq \mid \mid js\rangle; \tag{4}$$

$$\text{(d)}\quad \langle iu \mid \mid jv\rangle = \sum_{s=1}^{N} c_{sv}\langle iu \mid \mid js\rangle. \tag{5}$$

In the following discussion, N is the length of the expansion basis, n_C is the number of closed shell orbitals, and n_T is the total number of occupied orbitals. The orbitals are assumed to be serially occupied ($1 \leqslant$ index of the doubly occupied orbital $\leqslant n_C$, $n_C + 1 \leqslant$ index of the singly occupied orbital $\leqslant n_T$).

As discussed by Bender [5], Diercksen [3], and Pendergast and Fink [6], the summations in Eqs. (2)–(5) cannot be carried out as written without some manipulation of the lists input to, or produced by, these equations. The reason for this will become apparent in later discussions. In addition, the list shortening possible by retention of only those items necessary as opposed to retention of all N^4 items in the complete lists and the possible exclusion of zero valued integrals are important considerations. For the input atomic basis integral list, the first of these savings makes use of the symmetric permutations of integral indices:

$$\begin{aligned}\langle ab \mid \mid cd\rangle &= \langle ab \mid \mid dc\rangle \\ &= \langle ba \mid \mid dc\rangle \\ &= \langle ba \mid \mid cd\rangle \\ &= \langle cd \mid \mid ab\rangle \\ &= \langle cd \mid \mid ba\rangle \\ &= \langle dc \mid \mid ba\rangle \\ &= \langle dc \mid \mid ab\rangle.\end{aligned} \tag{6}$$

Thus only unique ab and cd pairs need be retained, resulting in

$$N^4/8 + N^3/4 + 3N^2/8 + N/4 \tag{7}$$

entries. The ranges of the indices for this set are given by

$$\begin{aligned} &1 \leqslant [cd] \leqslant [ab] \qquad \text{for every } [ab], \\ &1 \leqslant [ab] \leqslant [NN], \end{aligned} \tag{8}$$

where

$$[ab] = P(ab) = a(a-1)/2 + b \tag{9}$$

is the pair position of the pair ab in a list of unique pairs of a and b,

$$1 \leqslant b \leqslant a, \qquad 1 \leqslant a \leqslant N, \tag{10}$$

in which there are $(N^2 + N)/2$ unique pairs.

The list $\langle iq \mid \mid rs \rangle$ created through the use of Eq. (2) should have the order

$$\begin{gathered} 1 \leqslant r \leqslant N \qquad \text{for every iqs}, \qquad 1 \leqslant [qs] \leqslant [NN] \qquad \text{for every } i, \\ 1 \leqslant i \leqslant n_T, \end{gathered} \tag{11}$$

resulting in $n_T(N^3 + N^2)/2$ terms produced. The second partial summation, Eq. (3), then becomes the straightforward task of N multiplications and additions to form each member of a list $\langle iq \mid \mid js \rangle$ where

$$\begin{gathered} 1 \leqslant j \leqslant i \qquad \text{for every } iqs, \qquad 1 \leqslant [qs] \leqslant [NN] \qquad \text{for every } i, \\ 1 \leqslant i \leqslant n_T. \end{gathered} \tag{12}$$

This list consists of $n_T(n_T + 1) \cdot N(N+1)/4$ total terms.

Some manipulation of the list defined by Eq. (12) in concert with the application of Eq. (4) can produce the list of $\langle iu \mid \mid js \rangle$,

$$\begin{gathered} 1 \leqslant s \leqslant N \qquad \text{for every } iuj, \qquad n_c + 1 \leqslant u \leqslant N \qquad \text{for every } ij, \\ 1 \leqslant [ij] \leqslant [n_T n_T], \end{gathered} \tag{13}$$

which consists of $N(N - n_c)\, n_T(n_T + 1)/2$ terms. The last partial summation, Eq. (5), then becomes a straightforward exercise of N multiplications and additions to form each member of the output list $\langle iu \mid \mid jv \rangle$,

$$\begin{gathered} 1 \leqslant [jv]^0 \leqslant [iu]^0 \qquad \text{for every } iu, \\ 1 \leqslant [iu]^0 \leqslant [n_T N]^0. \end{gathered} \tag{14}$$

In this last equation,

$$[ab]^0 = P^0(ab) = a(N - n_c) + b - N, \tag{15}$$

is the pair position of pair ab in the list generated by this method, with

$$n_c + 1 \leqslant b \leqslant N, \qquad 1 \leqslant a \leqslant n_T. \tag{16}$$

The position of any $\langle iu \mid \mid jv \rangle$ in the partially transformed list is given by

$$I(i, u, j, v) = \tfrac{1}{2}\{[iu]^0([iu]^0 - 1)\} + [jv]^0. \tag{17}$$

The number of operations required to accomplish the desired transformation via this algorithm for all $\langle iu \mid \mid jv \rangle$ is

$$\tfrac{1}{2}\{N^4 n + 3N^3(n_{\mathrm{T}}^2 + n_{\mathrm{T}}) + N^2(n_{\mathrm{T}} - 3n_{\mathrm{T}}^2 n_{\mathrm{C}} - n_{\mathrm{C}} n_{\mathrm{T}}) + N(n_{\mathrm{T}}^2 n_{\mathrm{C}}^2 - n_{\mathrm{T}} n_{\mathrm{C}})\}, \tag{18}$$

which reduces to

$$\tfrac{1}{2}\{N^4 n_T + 3N^3(n^2 + n) + N^2(n - 3n^3 - n^2) + N(n^4 - n^2)\} \tag{19}$$

if $n_{\mathrm{T}} = n_{\mathrm{C}} = n$. This can be compared with the full transformation for which the total number of operations is

$$\tfrac{11}{8}N^5 + \tfrac{7}{4}N^4 + \tfrac{5}{8}N^3 + \tfrac{1}{4}N^2. \tag{20}$$

Here, the "Golden Rule" [4] has been followed as long as $(N - n_{\mathrm{C}}) \geqslant n_{\mathrm{T}}$. It should be noted that the second terms in Eqs. (17) and (18) may contribute significantly to the number of operations. For $N = 60$ and n values of 10, 20, and 30, (partial operations/full operations) · 100 values are 9.9, 23.5, and 38.7, respectively. Thus, the worst case considered (number of occupied orbitals $=\frac{1}{2}$ of basis) requires less than 40% of the number of operations involved in the full transformation.

The final list $\langle iu \mid \mid jv \rangle$ has the range $\langle 1n_{\mathrm{C}} + 1 \mid \mid 1n_{\mathrm{C}} + 1 \rangle$ to $\langle n_{\mathrm{T}} N \mid \mid n_{\mathrm{T}} N \rangle$, and has the following number of terms,

$$n_{\mathrm{T}}(N - n_{\mathrm{C}})\{n_{\mathrm{T}}(N - n_{\mathrm{C}}) + 1\}/2, \tag{21}$$

as opposed to the approximately $N^4/8$ unique terms in the fully transformed list. For $N = 60$, and n values of 10, 20, and 30 (the same values considered in the previous examples), the percent of partial to full list members is 7.5, 19.1, and 24.2, respectively.

METHOD

A flow scheme of the algorithm described herein is given in Fig. 1. The individual elements are described below.

(1) *Reorder input basis.* An input integral list which contains more than just the unique set of integral values specified by Eq. (8) is required. Thus, such a list $\langle pq \mid \mid rs \rangle$ must be expanded to include all integrals whose indices satisfy the following conditions:

$$\begin{gathered} 1 \leqslant p \leqslant N, \ 1 \leqslant r \leqslant N \qquad \text{for every } [qs], \\ 1 \leqslant [qs] \leqslant [NN]. \end{gathered} \tag{22}$$

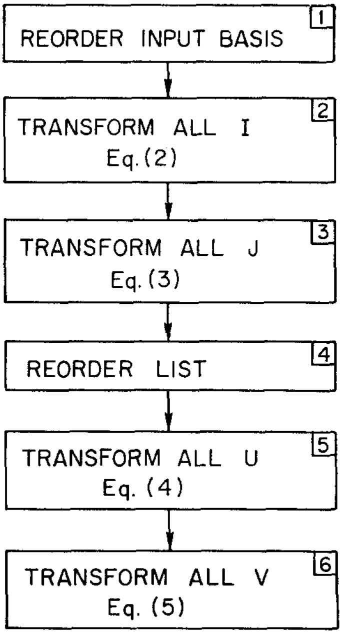

FIG. 1. Flow scheme of transformation algorithm

The unusual pairing of indices p with r (and q with s) rather than the more common pq (rs) pairing is required so that the "Golden Rule of Multiple Summation" [4] can be followed. The reordering uses the algorithm reported by McLean [4] which calls for the formation of blocks of integrals that contain all integrals with the unique set of indices pr for a given number of pairs qs.

(2) *Transform all i, Eq.* (2). The expanded list is input to Eq. (2) to form a list $\langle iq \,|\,|\, rs \rangle$ whose indices satisfy Eq. (11). One needs to retain only part of the list generated by Eq. (2) in computer memory at any given time while producing those integrals with indices $\langle iq \,|\,|\, rs \rangle$ for the range of qs values in the block being treated. This reduces the amount of storage required for the transformation, so that the core requirement is proportional to N^2. Only one reading of the input integral list is required to complete this partial summation. The ordering of Eq. (11) is maintained through the use of random access disk files.

(3) *Transform all j, Eq.* (3). The second partial summation is accomplished easily due to the ordering of the previously produced list; again the entire list $\langle iq \,|\,|\, js \rangle$ satisfies Eq. (12).

(4) *Reorder list.* The output list from Eq. (3) is reordered to satisfy

$$1 \leqslant q \leqslant N \qquad \text{for every } ijs, \qquad 1 \leqslant s \leqslant N, \qquad 1 \leqslant [ij] \leqslant [n_T n_T]. \tag{23}$$

Blocks of integrals which contain all integrals with indices q and s for a given number of unique pairs ij are formed to facilitate the next steps of the transformation.

(5) *Transform all u.* The reordered list generated in the previous step is transformed according to Eq. (4). This is essentially the same type of operation as step 3.

(6) *Transform all v.* The partial summation indicated by Eq. (5) may be applied as written due to the ordering of the list defined by Eq. (13) to form the final partially transformed list, Eq. (14).

As discussed previously, the unique pairings of p with r and q with s are necessary for the "Golden Rule" to be followed. For this pairing the amount of work is roughly proportional to $n_T N^4/2$. If one retains the usual pairing convention of p with q and r with s and transforms in the index order i, u, j, and v, the number of operations is proportional to $n_T N^4$.

CONCLUSION

The total number of terms produced by the partial transformation is a small fraction of those produced in the full transformation. The reduction in the total number of terms is substantial and this is important for expeditious formation of matrix elements within the Perturbation Theory Configuration Interaction program.

This method may be extended to include more than one core configuration; however, the techniques required depend upon the nature of the core configurations and the excitations allowed. In some cases the required integrals may be obtained just by extending the range of the occupied set to include all orbitals with fractional occupation in the set of core configurations. Such a case can be handled by a simple change in the definition of n_T. However, when several core configurations are used, the interaction matrix elements may require matrix elements between configurations which differ by only a single excitation. In such cases, the matrix elements are more complicated and a simple extension of the present approach is not possible.

The method calculates all the integrals required for off-diagonal matrix elements in the method described by Siu and Hayes [1]. The diagonal matrix elements required for the method involve the integral sets $\langle ii \mid\mid jj \rangle$ and $\langle ij \mid\mid ij \rangle$ for every ij, $1 \leqslant [ij]^0 \leqslant [NN]^0$. But there are only N^2 total of these and they can be handled separately from the partially transformed list.

We have not discussed savings available through symmetry blocking of the *input list* [7]. Even so, the work savings and reduction of list size make this method a desirable way of handling perturbation theory configuration interaction calculations. This particular method has been implemented as part of the MOLE quantum chemistry system [8] and is an important element in realizing efficient PTCI calculations.

ACKNOWLEDGMENTS

This research was supported in part by grants to Rice University from the Robert A. Welch Foundation, Grant C-389, and the National Science Foundation, Grant CHE 75-00937-A03. The authors wish to thank Dr. Glenn O. Morrell, Jr. for help in some preliminary calculations.

REFERENCES

1. A. K. Q. SIU AND E. F. HAYES, *J. Chem. Phys.* **61** (1974), 37.
2. J. A. POPLE, J. S. BINKLEY, AND R. SEEGER, *Int. J. Quantum Chem.* XS (1976), 1.
3. G. H. F. DIERCKSEN, *Theor. Chim. Acta* **33** (1974), 1.
4. A. D. MCLEAN, Potential energy surfaces from *ab initio* computation: Current and projected capabilities of the ALCHEMY computer program, *in* "Proceedings of the Conference on Potential Energy Surfaces in Chemistry" (W. A. Lester, Jr., Ed.), 1971.
5. C. F. BENDER, *J. Computational Phys.* **9** (1972), 547.
6. P. PENDERGAST AND W. H. FINK, *J. Computational Phys.* **14** (1974), 286.
7. I. SHAVITT, The method of configuration interaction, *in* "Modern Theoretical Chemistry" Vol. II, "Electronic Structure: *ab initio* Methods" (H. F. Schaefer III, Ed.), Plenum, New York, 1976.
8. (a) S. ROTHENBERG, P. KOLLMAN, M. E. SCHWARTZ, E. F. HAYES, AND L. C. ALLEN, *Int. J. Quantum Chem.* IIIS (1970), (b) P. Pendergast, E. F. Hayes, R. C. Liedtke, M. E. Schwartz, S. Rothenberg, and P. A. Kollman, *Int. J. Quantum Chem.* XS (1976), 77.

RECEIVED: August 31, 1976; REVISED: March 4, 1976.

PHIL PENDERGAST AND EDWARD F. HAYES
Department of Chemistry
William Marsh Rice University
Houston, Texas 77001