Theor Chim Acta (1991) 79: 115–122

Theoretica Chimica Acta

Exploiting non-abelian point group symmetry in direct two-electron integral transformations

Marco Häser,[1],* Jan Almlöf,[2] and Martin W. Feyereisen[2],**

[1] Institut für Physikalische Chemie (Theoret. Chemie), Universität Karlsruhe (TH), Kaiserstr. 12, W-7500 Karlsruhe 1, Federal Republic of Germany

[2] Department of Chemistry and Minnesota Supercomputer Institute, University of Minnesota, Minneapolis, Minnesota 55415, USA

Received January 24, 1990; received in revised form October 16, 1990/Accepted October 23, 1990

Summary. A simple and general scheme to exploit any discrete point group symmetry in two-electron integral transformations is introduced. It has been implemented together with integral pre-screening techniques in direct two-electron integral transformations. Its application has also been extended to subsequent MO integral processing steps like MP2 or solution of the coupled-perturbed Hartree–Fock equations (CPHF). Timings for representative applications are presented.

Key words: Symmetry – Integral transformation – Direct methods – Molecular orbitals – Electron repulsion integrals – Two-electron integrals – MP2 – CPHF

1. Introduction

Orbital and geometry optimizations at the SCF level of theory can today be carried out routinely and cheaply for moderately sized molecules (10–30 atoms) with the aid of workstation computers [1], or for bigger molecules (30–100 atoms) on supercomputers [2]. These applications take advantage of the rapid development of CPU technology by employing direct schemes which avoid input/output by calculating two-electron integrals over atomic orbitals (AOs) whenever needed. This idea has been introduced as the "direct" SCF approach by Almlöf et al. [3], and is in contrast to the conventional methods, which calculate and store integrals once to retrieve them in each iteration.

The direct approach has recently been extended to the transformation of two-electron integrals from an AO basis set to a molecular orbital basis set (MO) [4, 5]. The availability of transformed integrals greatly simplifies calculations which account for electron correlation, or which demand many coupled Hartree–Fock equations (CPHF) to be solved, like the simultaneous evaluation of infrared frequencies and intensities [6].

---

* *Present address*: Minnesota Supercomputer Institute, 1200 Washington Ave. S., Minneapolis, MN 55415, USA

** *Present address*: Department of Chemistry, University of Utah, Salt Lake City, UT 84112, USA

Since the direct SCF idea was first introduced, improved algorithms for the evaluation of two-electron integrals [7, 8], enhanced integral pre-screening [9, 10], and efficient methods to exploit molecular symmetry [3, 9], have advanced the use of direct SCF programs.

Naturally one asks which of the improvements of the direct SCF scheme can be taken over to a direct transformation. Of these the use of molecular symmetry is non-trivial and deserves special attention. It is well known that the point group D_{2h} and its subgroups are readily utilized in transformation procedures, whereas going to higher groups becomes a tedious task, even though the theory for general point groups is known [11]. An algorithm which retains some simplicity and at the same time accounts for some of the more important non-abelian groups has been suggested by Čarsky et al. [12]. Their method can be applied to direct transformations, but it lacks generality. Other schemes which go beyond D_{2h} require an AO → SO (symmetry orbital) transformation first. This hampers integral pre-screening, since SOs tend to be more delocalized than AOs. Such algorithms seem less suited to a direct scheme.

Here we introduce a simple and general method to use symmetry in a direct AO → MO transformation of two-electron integrals.

We finally obtain "packed integrals" in the spirit of Slater–Condon integrals [13] or of Pitzer's equal contribution electron repulsion integrals [14], but we arrive there without resorting to rigorous coupling algebra. Our approach is based upon projecting the integrand in a two-electron MO integral onto its totally symmetric part which can be represented as a linear combination of symmetry-related integrands. These projectors can easily be constructed from the representation matrices. An analysis of their properties (sparsity and low rank) provides the tools to set up an efficient but simple algorithm to handle the integrals. We present applications of the algorithm to the calculation of the MP2 correlation energy and the solution of CPHF equations.

2. Theory

In our approach complex numbers are completely avoided. Given the molecular point group $\mathscr{G}$ of order g we choose a complete set of non-equivalent real matrix representations $\boldsymbol{D}^{\Gamma_1}, \boldsymbol{D}^{\Gamma_2}, \ldots$ so that Γ is either irreducible or reducible to a pair of complex conjugate one-dimensional representations only.

The molecular orbitals (MOs) can be chosen to constitute a basis for these representations:

$$\hat{R}|i\Gamma\gamma\rangle = \sum_{\varepsilon=1}^{\dim(\Gamma)} |i\Gamma\varepsilon\rangle D^{\Gamma}_{\varepsilon\gamma}(R), \tag{1}$$

where $\hat{R}$ denotes a symmetry operator, $\{|i\Gamma\gamma\rangle,\ \gamma = 1, \ldots, \dim(\Gamma)\}$ is a set of degenerate MOs forming a basis for the irreducible representation Γ with the corresponding matrices $\boldsymbol{D}^{\Gamma}$.

Similarly, the quadruples of MOs $\{|i_1\Gamma_1\gamma_1\rangle, |i_2\Gamma_2\gamma_2\rangle, |i_3\Gamma_3\gamma_3\rangle, |i_4\Gamma_4\gamma_4\rangle),$ $\gamma_1 = 1, \ldots, \dim(\Gamma_1), \ldots, \gamma_4 = 1, \ldots, \dim(\Gamma_4)\}$ define a basis in a (in general reducible) representation space S^{Π} of the Kronecker product representation:

$$\Pi := \Gamma_1 * \Gamma_2 * \Gamma_3 * \Gamma_4, \tag{2}$$

such that for all $R \in \mathscr{G}$:

$$\boldsymbol{D}^{\Pi}_{\sigma\pi}(R) = \boldsymbol{D}^{\Gamma_1}_{\varepsilon_1\gamma_1}(R)\boldsymbol{D}^{\Gamma_2}_{\varepsilon_2\gamma_2}(R)\boldsymbol{D}^{\Gamma_3}_{\varepsilon_3\gamma_3}(R)\boldsymbol{D}^{\Gamma_4}_{\varepsilon_4\gamma_4}(R), \tag{3}$$

where $\sigma = (\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)$, $\pi = (\gamma_1, \gamma_2, \gamma_3, \gamma_4)$.

Turning now to two-electron integrals formed over quadruples of MOs, we introduce as a shorthand notation:

$$I^{\Pi}_{\pi} := (i_1\Gamma_1\gamma_1, i_2\Gamma_2\gamma_2 \mid i_3\Gamma_3\gamma_3, i_4\Gamma_4\gamma_4), \tag{4}$$

where I is a collective symbol for (i_1, i_2, i_3, i_4).

It follows from Eq. (1) that (in symbolic notation, i.e. with the symmetry operator applied under the integral sign):

$$\hat{R}I^{\Pi}_{\pi} = \sum_{\sigma} I^{\Pi}_{\sigma}\boldsymbol{D}^{\Pi}_{\sigma\pi}(R). \tag{5}$$

We are interested only in non-zero integrals $I^{\Pi}_{\pi} \neq 0$ and we can thus trivially dismiss all product representations Π which do not contain the totally symmetric representation A. This can quickly be established by character operations.

If however Π contains the totally symmetric representation A n^{Π}_{A} times ($n^{\Pi}_{A} > 0$), then we need to investigate the product representation space S^{Π}. Our tool is the operator which projects S^{Π} onto its totally symmetric subspace. This operator has the matrix representation:

$$\boldsymbol{P}^{\Pi}_{\sigma\pi} := \frac{1}{g}\sum_{R\in\mathscr{G}} \boldsymbol{D}^{\Pi}_{\sigma\pi}(R). \tag{6}$$

Henceforth we will refer to $\boldsymbol{P}^{\Pi}$ as the totally symmetric projector.

Since only the totally symmetric part of the integrand influences the value of I^{Π}_{π}, we can substitute the integrand by its projection onto the totally symmetric subspace of S^{Π}:

$$I^{\Pi}_{\pi} = \sum_{\sigma=1}^{\dim(\Pi)} I^{\Pi}_{\sigma}P^{\Pi}_{\sigma\pi}. \tag{7}$$

From Eq. (7) we can infer a more stringent necessary condition[1] for an integral I^{Π}_{π} to be non-zero:

$$\varrho^{\Pi}_{\pi} := \sum_{\sigma=1}^{\dim(\Pi)} |\boldsymbol{P}^{\Pi}_{\sigma\pi}| > 0. \tag{8}$$

It is convenient to define an index set O^{Π}, which consists of all (quadruple) indices $\pi = (\gamma_1, \gamma_2, \gamma_3, \gamma_4)$ for which $\varrho^{\Pi}_{\pi} > 0$, that is, which correspond to integrals that contain totally symmetric contributions. Integrals I^{Π}_{π} with $\pi \notin O^{\Pi}$ are zero and can be discarded beforehand.

From (7) we can also conclude that integrals I^{Π}_{π} and I^{Π}_{ϱ} are equal if $\boldsymbol{P}^{\Pi}_{\sigma\pi} = \boldsymbol{P}^{\Pi}_{\sigma\varrho}$ for all $\sigma \in O^{\Pi}$. Again we define a subset R^{Π} of O^{Π}, containing the indices of representative non-equal and non-zero integrals only. Equality to within a $\pm$ sign may be treated similarly.

So far we have exploited the low rank n^{Π}_{A} of the projector $\boldsymbol{P}^{\Pi}$ only partially (in many cases $n^{\Pi}_{A} = 1$). We can take full advantage of it, writing $\boldsymbol{P}^{\Pi}$ in terms of its spectral resolution (in matrix notation):

$$\boldsymbol{P}^{\Pi} = \boldsymbol{U}^{\Pi}\lambda^{\Pi}\boldsymbol{U}^{\Pi+}, \tag{9}$$

---

[1] In general (all four sets of MOs being different) this is also a sufficient condition for the integral to be non-zero on symmetry grounds

where λ^{Π} is a diagonal matrix with the first n_A^{Π} diagonal elements = 1, the remaining being zero.

It follows that:

$$P_{\sigma\varrho}^{\Pi} = \sum_{\tau=1}^{n_A^{\Pi}} U_{\sigma\tau}^{\Pi} U_{\varrho\tau}^{\Pi}, \qquad \sigma, \varrho = 1, \ldots, \dim(\Pi). \tag{10}$$

Equation (7) now splits into two equations:

$$\tilde{I}_{\tau}^{\Pi} := \sum_{\sigma \in O^{\Pi}} I_{\sigma}^{\Pi} U_{\sigma\tau}^{\Pi}, \qquad \tau = 1, \ldots, n_A^{\Pi}, \tag{11}$$

$$I_{\varrho}^{\Pi} = \sum_{\tau=1}^{n_A^{\Pi}} \tilde{I}_{\tau}^{\Pi} U_{\varrho\tau}^{\Pi}, \qquad \varrho = 1, \ldots, \dim(\Pi). \tag{12}$$

We shall refer to the intermediate quantities $\tilde{I}_{\tau}^{\Pi}$, $\tau = 1, \ldots, n_A^{\Pi}$, as packed integrals. While their number, n_A^{Π}, is usually significantly smaller than $\dim(\Pi)$, all the original integrals can always be recovered – unpacked – by Eq. (12). In actual applications only the representative integrals of R^{Π} are unpacked.

Packing and unpacking are simple procedures with negligible computing cost – as will be shown in our numerical analysis. Here we do not go into details of the various applications, but instead we show how packing aids the symmetrization step in a direct AO → MO two-electron integral transformation.

The straightforward transformation of the symmetry-distinct and properly weighted [15] AO integrals yields what we call skeleton MO integrals $\hat{I}_{\sigma}^{\Pi}$, $\sigma \in O^{\Pi}$. These still need to be symmetrized!

As noticed by Lazzeretti et al. [16], this can be achieved by the formula:

$$I_{\pi}^{\Pi} = \sum_{\sigma \in O^{\Pi}} \hat{I}_{\sigma}^{\Pi} P_{\sigma\pi}^{\Pi}, \qquad \pi \in O^{\Pi} \tag{13}$$

(the formula in Ref. [16] differs slightly in that σ, π still run through all quadruples $\sigma = (\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)$, $\pi = (\gamma_1, \gamma_2, \gamma_3, \gamma_4)$; we also note that the normalization of the projector differs).

The similarity between Eqs. (7) and (13) is a consequence of Pitzer's equal contribution theorem [17]. Lazzeretti et al. simplified Eq. (13) further for D_{2h} and its subgroups, but they did not regard the general case in any more detail. Čarsky et al. [12] extended the application of Eq. (13) to some of the more important non-abelian point groups.

We now apply to Eq. (13) the same procedure as previously done to Eq. (6) and obtain:

$$\tilde{I}_{\tau}^{\Pi} = \sum_{\sigma \in O^{\Pi}} \hat{I}_{\sigma}^{\Pi} U_{\sigma\tau}^{\Pi}, \qquad \tau = 1, \ldots, n_A^{\Pi} \tag{14}$$

in close analogy to Eq. (11). That's our symmetrization procedure – it's simply packing the skeleton integrals! (Note that $\tilde{I}_{\tau}^{\Pi}$ has a totally symmetric integrand and Pitzer's theorem thus applies.)

3. Numerical analysis and discussion

We present detailed timings for an SCF force constant calculation of hexamethylenetetramine $(CH_2)_6N_4$ with the coupled Hartree–Fock step (CPHF) carried out in the MO basis. The calculations have been done with the program system TURBOMOLE [6] on an Hewlett–Packard HP 835 SRX workstation computer. The basis set employed is a 3–21 G, giving rise to 114 orbitals, one-third of which are occupied. The symmetry group is T_d.

The following steps have been timed separately:

(1) Find all product representations $\Pi = \Gamma_1 * \Gamma_2 * \Gamma_3 * \Gamma_4$ which contain the totally symmetric representation. For these, evaluate $\boldsymbol{P}^{\Pi}$ according to Eq. (6), determine O^{Π}, R^{Π}, and eigenvectors U^{Π}, which are stored as sparse matrices (a few kB).

(2) Evaluate symmetry-distinct AO integrals $(\nu\mu \mid \chi\lambda)$ and transform them to "skeleton" MO integrals $\hat{I}^{\Pi}_{\sigma}$ ($\sigma \in O^{\Pi}$: all-internals, one-externals, two-externals).

(3) Symmetrize $\hat{I}^{\Pi}_{\sigma}$ according to Eq. (14) and store the packed integrals $\tilde{I}^{\Pi}_{\tau}$ on disk.

(4) Construct the A-matrix for CPHF: $A_{iajb} = 4(ia \mid jb) - (ib \mid ja) - (ij \mid ab)$. That is, read packed two-externals, unpack them, form A-matrix, pack A-matrix according to Eq. (11) (note that the same procedure works for integrals *and* A-matrix elements!), and store it on disk. (The packed A-matrix takes only 1.5 MB and could be kept in main memory if desired.)

(5) Evaluate and process derivatives of one-electron AO integrals [18].

(6) Evaluate and process derivatives of two-electron AO integrals [18].

(7) Solve CPHF equations using the out of core method proposed by Ahlrichs et al. [5], which involves 18 MB I/O in the present case (the equations are solved for 3 right-hand sides (which correspond to x, y, z displacements of one nucleus) simultaneously: there are 3 symmetry-distinct nuclei and in each case 4 iterations sufficed for convergence to 7 significant digits – $3*4*1.5 = 18$ MB).

(8) Process CPHF solutions

The computational cost of these steps are given in Table 1. Symmetrization, packing and unpacking of two-electron MO integrals or A-matrix elements can be seen to be a matter of seconds while the whole calculation takes 105 min. It should also be recognized that the most expensive step is not yet the transformation of the two-electron integrals, but the calculation of AO integral derivatives. The CPHF step is comparatively fast.

We further tested the performance of the force constant evaluation in the symmetry groups D_{2d} and D_2, comparing computation cost to T_d. The respective timings listed in Table 2 indicate a symmetry speed up for the transformation which is somewhat less than the order of the group: it scales more like the ratio of symmetry-distinct AO pairs to total number of pairs. This behavior is what one would expect for a direct AO → MO transformation and could be improved

Table 1. Timings of individual steps in a force constant calculation of hexamethylenetetramine in T_d-symmetry on a HP 835 SRX workstation computer. The steps are explained in the text. Steps marked with an asterisk (*) have been performed without using symmetry

Step	1	1	3	4	5*	6	7	8*
Time(s)	$\leqslant 7$	1412	19	37	689	2615	486	1104

if an AO → SO transformation were done first. There are three reasons to avoid an AO → SO transformation in a direct algorithm:

(1) From their calculation SO integrals do not come in an order convenient for the immediate SO → MO transformation; as a consequence more partially transformed integrals have to be stored.

(2) The AO → SO transformation becomes expensive as soon as a multiple path through the AO integral evaluation is necessary in order to cut down storage demand in the subsequent transformation to MO integrals (this applies for instance to the amount of half-transformed integrals which are written to direct access file in TURBOMOLE [6], or to fully transformed MO integrals to be kept in memory in DISCO [19]).

(3) Integral pre-screening, to be applied for instance to the half-transformed integrals $(\nu\mu \mid ij)$ by means of the bound [6, 9, 10]

$$|(\nu\mu \mid ij)| \leqslant (\nu\mu \mid \nu\mu)^{1/2}(ij \mid ij|^{1/2} \tag{15}$$

performs better if ν, μ are AOs than if ν, μ are SOs and are thus more delocalized. Careful analysis shows that the AO → MO transformation becomes competitive to the SO → MO transformation if pre-screening is used and the molecule is large enough. This does not hold for the last quarter transformation, since the canonical Hartree–Fock MOs are so delocalized. The last quarter transformation, however, is inexpensive as long as the four-external MO integrals are not needed (a solution appropriate for the evaluation of four-external MO integrals would be a symmetrization of the three-quarter transformed integrals according to Eq. (14) after an AO → SO transformation of the remaining AO).

Table 2. Comparison of performance of a force constant calculation with MO-based CPHF, including AO → MO transformation of two-electron integrals, for hexamethylenetetramine in T_d-symmetry versus subgroups D_{2d} and D_2. Timings refer to a Hewlett–Packard HP 835 SRX workstation computer

(Sub) Group	T_d	D_{2d}	D_2
Transformation	24′35″	45′10″	74′28″
A-matrix (packed)	1.5 MB	4.4 MB	8.5 MB
Second derivatives & CPHF	80′31″	164′47″	292′20″
I/O in CPHF	18 MB	88 MB	297 MB

4. Conclusion

The construction of projection operators, which project quadruples of molecular orbitals onto their totally symmetric part, Eqs. (3, 6, 7), provides a means of identifying those two-electron MO integrals, A-matrix elements or similar quantities, which are zero or equal on symmetry grounds. The factorization of these projection operators according to Eqs. (9, 10) serves as a substitute for rigorous coupling analysis and makes possible the construction of packed quantities (MO integrals), Eq. (11), which are similar to irreducible tensor elements. Packing not only reduces the number of quantities to be handled but also results in their symmetrization, Eq. (14), at no extra cost. Unpacking, Eq. (12) supplies the (symmetrized) original quantities (MO integrals). Packing and unpacking have been shown to take only a negligible amount of computing time and can be vectorized by putting molecular shell indices into the innermost loops. These techniques have been used to facilitate the symmetrization step in our direct AO → MO transformation of two-electron repulsion integrals. Together with pre-screening of near-zero integrals the direct AO → MO transformation becomes competitive with SO-based transformations in the limit of large molecules. Packing and unpacking have also been applied to the A-matrix in the coupled Hartree–Fock step of a force constant evaluation in order to reduce I/O or memory requirements. The relative simplicity of the method allows its implementation in existing programs. Presently it is used in the program systems TURBOMOLE [6] and DISCO [19]. Some recent applications are listed in Table 3. It can be seen how the use of symmetry in the direct AO → MO transformation and the subsequent handling of two-electron MO integrals advances the applicability of MO-based methods to more than 800 basis functions.

Table 3. Applications of the direct AO → MO two-electron integral transformations in TURBOMOLE (=T) and DISCO (=D) which use the symmetry approach proposed in this publication. Note that for CPHF one-external, two-external, and all-internal MO integrals have to be evaluated, while for MP2 only exchange-type two-externals are needed. For the MP2 applications only active orbitals have been included in the counting of occupied molecular orbitals. The computers employed are Hewlett–Packard HP 835 SRX (a), Silicon Graphics IRIS-4D/70T (b), SUN microsystems SUN-4/280 (c), and Cray-2 (d). The timings for TURBOMOLE include the time for AO integral evaluation, whereas DISCO timings are for the transformation part only

Molecule	Symmetry	AOs	occ. MOs	Application	Time for transform.	Comp.	Program
Hexamethylene-tetramine	T_d	114	38	CPHF	1/2 h	a	T
Cyclododeca-tetraene [20]	C_{4v}	152	44	CPHF	4 h	b	T
B_6H_{14} [21]	C_s	160	22	CPHF	10 h	b	T
$Cu_2(N_3H_2)_2$ [22]	D_{2h}	160	52	CPHF	9 h	b	T
$Ni(CO)_4$	T_d	147	29	MP2	100 min	b	T
$Ce(C_8H_8)_2$ [23]	D_{8h}	359	46	MP2	24 h	c	T
$C_{34}H_{36}N$ [24]	T_d	391	88	MP2	20 min	d	D
$(C_6H_6)_2$ [25]	D_{6h}	636	30	MP2	2 h	d	D
C_{18} [26]	D_{9h}	828	36	MP2	5 h	d	D

Acknowledgements. The authors thank Prof. N. C. Handy and Prof. H. F. Schaefer for organizing the conference "Forty Years of Quantum Chemistry" where the authors met for the first time and found that they had independently developed the same ideas. One of the us (MH) is particularly thankful to his teacher, Prof. R. Ahlrichs, for this continuous support and interest. This work was supproted by Fonds der Chemischen Industrie and by the National Science Foundation, grant No. CHE-8915629.

References

1. Ahlrichs R, Häser M, Schnöckel H, Tacke M (1989) Chem Phys Lett 154:104–110
2. Lüthi HP, Almlöf J (1987) Chem Phys Lett 135:357–360
3. Almlöf J, Faegri K, Korsell K (1982) J Comput Chem 3:385–399
4. Almlöf J, Saebø S (1989) Chem Phys Lett 154:83–89
5. Head-Gordon M, Pople JA, Frisch MJ (1988) Chem Phys Lett 153:503–506
6. Ahlrichs R, Bär M, Häser M, Horn H, Kömel C (1989) Chem Phys Lett 162:165–169
7. Obara S, Saika A. (1986) J Chem Phys 84:3963–3974
8. Gill PMW, Head-Gordon M, Pople JA (1989) Int J Quant Chem Symp 23:269–280
9. Häser M, Ahlrichs R (1989) J Comput Chem 10:104–111
10. Maxwell JC (1904) A treatise on electricity and magnetism, Vol 1, Clarendon Press, London; see also e.g., Power SD, Pitzer RM (1974) Chem Phys Lett 24:478–483; Whitten, JL (1973) J Chem Phys 58:4496–4501
11. Davidson ER (1975) J Chem Phys 62:400–409
12. Čarsky P, Hess BA, Schaad LJ (1984) J Comput Chem 5:280–287
13. Condon EU, Shortley GH (1935) The theory of atomic spectra, Cambridge Univ Press, Cambridge
14. Pitzer RM (1973) J Chem Phys 59:3308–3312
15. Dupuis M, King HF (1977) Int J Quant Chem 11:613–625
16. Lazzeretti P, Zanasi R, Rossi E (1980) Chem Phys Lett 75:392–394
17. Pitzer RM (1973) J Chem Phys 58:3111–3112
18. Subroutines taken from program, HONDO VII: Dupuis M, Watts SD, Villars HO, Hurst GJB (1988) IBM Corp, Kingston (New York)
19. DISCO is a vectorized code for Direct SCF and MP2 calculation, written by: Feyereisen MW, Korsell K, Faegri K, Almlöf J
20. Krause A, Musso H, Boland W, Ahlrichs R, Gleiter R, Boese R, Bär M (1989) Angew Chem 101:1401–1402
21. Horn H, Ahlrichs R, Kölmel C (1988) Chem Phys Lett 150:263–267
22. Kömel C, Ahlrichs R (1990) J Phys Chem 94:5536–5542
23. Dolg M private communication
24. Liu JJ, Feyereisen MW, Almlöf J, Saebø S, Rohlfing CM (to be published)
25. Almlöf J, Chen BT, Feyereisen, MW (to be published)
26. Parasuk V, Almlöf J, Feyereisen MW (1991) J Am Chem Soc (in press)