

On the use of spatial symmetry in ab initio calculations. Transformation of the two-electron integrals from atomic orbital to localized molecular orbital basis

Ede Kapuy^{a,b} and Cornelia Kozmutza^b

^a*Department of Theoretical Physics, József Attila University, 6720 Szeged,
Aradi vértanúk tere 1, Hungary*

^b*Quantum Theory Group, Physics Institute, Technical University of Budapest,
Budafoki út 8, 1521 Budapest, Hungary*

Received 31 October 1991; revised 15 May 1992

An m^5 -dependent integral transformation procedure from atomic orbital basis to localized molecular orbitals is described for spatially extended systems with some Abelian symmetry groups. It is shown that exploiting spatial symmetry, the number of non-redundant integrals for normal saturated hydrocarbons can be reduced by a factor of 2.5–3.5, depending on the size of the system and on the basis. Starting from a list of integrals over basis functions in canonical order, the number of multiplications of the four-index transformation is reduced by a factor of 2.8–3.5 as compared to that of Diercksen's algorithm. It is pointed out that even larger reduction can be achieved if negligible integrals over localized molecular orbitals are omitted from the transformation in advance.

1. Introduction

There exist various methods beyond the Hartree–Fock level (CI, MBPT, CCA, etc.) which require two-electron integrals over molecular orbitals (canonical or localized).

Several attempts have been made in this field during the last two decades to elaborate efficient procedures. The straightforward m^8 process was modified by Tang and Edmiston, who presented an m^5 algorithm [1]. Their method was further developed by other authors [2–4]. The utilization of symmetry was discussed by Bender [2], Winter et al. [5], and by Pitzer [6].

The method of the latter author was extended by several papers (Diercksen [7], Almöf [8], Davidson [9], Pople et al. [10], Saunders and van Lenthe [11], Carsky et al. [12]). Another approach for taking into account spatial symmetry was suggested by Dacre [13]. His method to generate symmetry-distinct two-electron integrals was extended by Dupuis and King [14] and by Taylor [15]. Dacre's

suggestions were utilized in a program package for spatially extended systems with Abelian symmetry by one of the present authors (CK) [16–19].

The four-index transformations were reviewed by Wilson [20].

It was pointed out that the use of localized molecular orbitals (LMOs) in spatially extended systems has some advantage over the canonical ones when the many-body perturbation theory (MBPT) is applied [21–25]. A special algorithm has been published recently [26].

In this paper, we study the exploitation of spatial symmetry when LMOs are used.

2. Integral transformation: from atomic orbital basis to LMOs

In the localized representation of the MBPT, the occupied φ_i and the virtual orbitals φ_a are separately localized by unitary transformations U, V ,

$$\varphi_i = \sum_{j=1}^N U_{ij} \varphi_j, \quad i, j = 1, 2, \dots, N_i,$$

$$\varphi_a = \sum_{b=N+1}^m V_{ab} \varphi_b, \quad a, b = N+1, N+2, \dots, m,$$

where N is the number of occupied orbitals and m is the number of basis functions. Multiplying the direct sum of U and V with the coefficients C_μ^k of the canonical orbitals, we obtain the matrix elements of $T_{k\kappa}$ of the transformation from atomic basis to LMOs. The transformation of the two-electron integrals can be expressed in the following form:

$$(\varphi_i \varphi_j | \varphi_k \varphi_l) = \sum_{\kappa, \lambda, \mu, \nu} T_{i\kappa} T_{j\lambda} T_{k\mu} T_{l\nu}, \quad (\kappa\lambda | \mu\nu), \quad (1)$$

where

$$(\kappa\lambda | \mu\nu) = \int \chi_\kappa(1) \chi_\lambda(1) r_{12}^{-1} \chi_\mu(2) \chi_\nu(2) dv_1 dv_2$$

are the integrals over basis functions χ_κ , $\kappa = 1, 2, \dots, m$.

As is well known, direct application of (1) is an m^8 process, but carried out in four steps, it becomes an m^5 procedure [1]. The number of required operations can be further reduced when the permutational symmetry of the integrals and/or the spatial symmetry of the system are taken into account [2–13]. We consider only spatially extended systems (chain molecules, oligomers, etc.). The spatial symmetry of these systems is usually low: Abelian groups (C_{2v} , C_{2h} , etc.).

As to the symmetry of the LMOs, there exist two different cases.

Case (a). All (or most) of the LMOs transform according to the irreducible representations of a subgroup of the symmetry group of the system. The integrals can be partitioned into blocks of lower dimension, which are transformed separately.

Case (a) seldom occurs for polymers and an efficient reduction of the number of necessary integrals is usually not possible. For example, for all-trans polyacetylene, when σ and π orbitals are separately localized, the plane of the molecule is a common symmetry of all the LMOs. As a consequence, integrals $(\pi\sigma|\sigma\sigma)$ vanish due to the symmetry, but they amount only to 30–45% of the total number.

Case (b). The LMOs of the system can be partitioned into subsets containing one, two and four equivalent functions transforming into each other under the symmetry operations of the group of the system. This is usually true for normal saturated hydrocarbons.

(It should be noted that in some cases the LMOs might be completely unsymmetrical. These cases will not be considered here.)

The list of the nonredundant integrals is not determined by the symmetry alone, but it is intimately connected to the distribution of the numbers of singly-, doubly- and quadruply-equivalent subsets M_1 , M_2 , and M_4 , respectively. The latter depend upon the system and upon the basis set used. The partitioning of the LMOs for normal saturated hydrocarbons with conventional basis sets is shown in table 1. (In the actual calculation, Boys' procedure was used.)

Table 1

The number of subsets M_1, M_2, M_3 containing 1, 2, 4 equivalent LMOs for some normal saturated hydrocarbons (C_nH_{2n+2}) with conventional basis sets. (m is the dimension of the bases.)

	Basis	m	M_1	M_2	M_3
C_3H_8	STO-3G	23	1	7	2
	6-31G	43	1	13	4
	6-31G*	61	3	17	6
C_5H_{12}	STO-3G	37	1	10	4
	6-31G	69	1	18	8
	6-31G*	99	3	24	12
C_7H_{16}	STO-3G	51	1	13	6
	6-31G	95	1	23	12
	6-31G*	137	3	31	18

In what follows, we describe a simple but appropriate procedure which makes it possible to collect the list of non-redundant integrals. We use capital letters I, J, K , and L to denote the subsets of LMOs, the members of which merely permuted (except for a sign) by the symmetry operations of the system \hat{A}_α , $\alpha = 1, 2, \dots, g$. The functions within each subset I are indexed by lower case letters $i, j, k, l = 1, 2, \dots, v_I$ as ϕ_{Ii} . The symmetry operations affect only the second indices and never the first ones:

$$\hat{A}_\alpha \varphi_{li} = \mp \varphi_{li'}. \quad (2)$$

Having only one LMO, φ_{li} , in subset I , we can generate all members of that subset by applying some (or all) symmetry operations of the group

$$\hat{A}_{li} \varphi_{li} = \varphi_{li}, \quad i = 1, 2, \dots, \nu_I. \quad (3)$$

We call operation \hat{A}_{li} the generator of LMO φ_{li} . As an example, in point group C_{2v} we have four symmetry operations: E , C_2 , σ_1 , σ_2 . The singly-equivalent functions are invariant (except for a sign) under all symmetry operations. The generators of the doubly-equivalent subset are E , C_2 . The quadruply-equivalent subsets are generated by the operations E , C_2 , σ_1 , σ_2 .

The set of the $\nu_I \nu_J \nu_K \nu_L$ integrals

$$\begin{aligned} (\varphi_{li} \varphi_{lj} | \varphi_{kk} \varphi_{ll}), \quad & i = 1, 2, \dots, \nu_I; \quad j = 1, 2, \dots, \nu_J; \\ & k = 1, 2, \dots, \nu_K; \quad l = 1, 2, \dots, \nu_L; \end{aligned} \quad (4)$$

I, J, K, L fixed,

are related to each other. They are not all different, i.e. the set contains redundant integrals. To pick out the non-redundant ones, we utilize the property that the value of an integral is invariant (except for a sign) under any operation of the symmetry group of the system:

$$\hat{A}_\alpha (\varphi_{li} \varphi_{lj} | \varphi_{kk} \varphi_{ll}) = p_{li}^\alpha p_{lj}^\alpha p_{kk}^\alpha p_{ll}^\alpha (\varphi_{li} \varphi_{lj} | \varphi_{kk} \varphi_{ll}), \quad (5)$$

where $p_{li}^\alpha, p_{lj}^\alpha, p_{kk}^\alpha, p_{ll}^\alpha = \mp 1$ are the parities of the corresponding LMOs upon operation \hat{A}_α . By applying the inverse of the generator of φ_{li} to all integrals of the set (4), we obtain

$$\hat{A}_{li}^{-1} (\varphi_{li} \varphi_{lj} | \varphi_{kk} \varphi_{ll}) = \mp (\varphi_{li} \varphi_{lj'} | \varphi_{kk} \varphi_{ll'}). \quad (6)$$

Consequently, the set can be partitioned into n_I groups which, apart from the sign of the integrals, are equal. This means that each of the groups contains, apart from the sign, all the $\nu_J \nu_K \nu_L$ non-redundant integrals of the set (4). The factor of reduction due to symmetry is at most 4. Since the factor of reduction due to the permutational symmetry is larger: 8, they should be utilized together. A suitable order of the LMO indices is the following: singly-equivalent subsets, doubly-equivalent subsets, and quadruply-equivalent subsets. We obtain a unique non-redundant list of integrals,

$$(\varphi_{li} \varphi_{lj} | \varphi_{kk} \varphi_{ll}), \quad (7)$$

if

$$L \leq K, \quad K \leq I, \quad J \leq I, \quad \text{for all } j, k, l.$$

Table 2

The number of non-redundant integrals I_{NR} when spatial symmetry is exploited, and the number of integrals I_{TOT} when only permutational symmetry is utilized for some normal saturated hydrocarbons. (R is the ratio I_{NR}/I_{TOT} in percent.)

	Basis	I_{TOT}	I_{NR}	R (%)
C_3H_8	STO-3G	38226	15354	40.17
	6-31G	447931	154410	34.47
	6-31G*	1788886	585203	32.32
C_5H_{12}	STO-3G	247456	84781	34.26
	6-31G	2917320	926831	31.48
	6-31G*	12253725	3746993	29.22
C_7H_{16}	STO-3G	879801	278004	31.60
	6-31G	10411580	3048516	29.28
	6-31G*	44684331	13071721	28.10

The number of non-redundant integrals over LMOs for some normal saturated hydrocarbons is shown in table 2. They can be compared with those obtained by taking into account the permutational symmetry only

$$I_{TOT} = (m^4 + 2m^3 + 3m^2 + 2m) / 8.$$

The factor of reduction due to the spatial symmetry is 2.5–3.5, increasing slowly with the dimension of the basis.

Table 3

The number of multiplications in four-index transformation when the spatial symmetry of the LMOs is taken into account: M_S compared to that in Diercksen's algorithm M_D [7] for some normal saturated hydrocarbon with conventional basis sets. The procedure starts from a permutationally non-redundant list of integrals. (Q is the ratio M_S/M_D in percent.)

	Basis	M_S	M_D	Q (%)
C_3H_8	STO-3G	9347430	3316600	35.48
	6-31G	208169665	69146924	33.22
	6-31G*	1185692929	388946797	32.80
C_5H_{12}	STO-3G	98659723	31742152	32.17
	6-31G	2190417075	669793350	30.58
	6-31G*	13244826375	3834253863	28.95
C_7H_{16}	STO-3G	486332379	149513436	30.74
	6-31G	10782564600	3097791920	28.73

The first partial sum is calculated for the "incomplete" I_{vj} indices, the second one for all Jj indices compatible with inequalities (7), etc. Starting from a list of

integrals over basis functions in canonical order, the number of multiplications in the four steps M_S for some normal saturated hydrocarbons are shown in table 3. These can be compared with the number of multiplications of Diercksen's algorithm [7]:

$$M_D = (11m^5 + 14m^4 + 5m^3 + 2m^2) / 8,$$

where the spatial symmetry is not exploited. It can be seen that the factor of reduction is again 2.8–3.5, depending on the basis. Starting from a symmetrically non-redundant list of integrals over basis functions, the reduction can be further increased [19].

3. Discussion

For small molecules, the symmetry-adapted LMOs have advantages over the LMOs giving a shorter non-redundant integral list due to spatial symmetry. The factor of reduction can exceed 6. For LMOs the upper bound is 4, which can be approached only for larger systems. The advantage of the LMOs manifests itself for spatially extended systems because using conventional basis sets, the LMOs are localized into different spatial regions. As a consequence, many integrals over LMOs can be neglected. It is unnecessary to transform out those integrals which will be omitted.

For linear chain systems, the LMOs can be partitioned into groups and to any pair of groups, a unique "order of neighbourhood" can be assigned. Having partitioned the LMOs, the classification of the integrals $(\varphi_i \varphi_j | \varphi_k \varphi_l)$ was carried out in the following way: if j is the μ th neighbour of i and l is the ν th neighbour of k , and $\mu \geq \nu$, then we consider the integral as the μ th neighbour effect [23–25]. In the terms of the energy corrections of the MBPT, the integrals of largest order determine the order of the term. It has been shown that, for normal saturated hydrocarbons with STO-3G basis set including only zeroth, first and second neighbours, we obtained more than 99% of the total second-order correlation energy correction [25]. The total number of integrals necessary for the calculation of the second-order correction when the permutational symmetry is taken into account is

$$I^{(2)} = \frac{1}{2} M_o^2 M_v^2 + \frac{1}{2} M_o M_v,$$

where M_o and M_v are the number of the occupied and the virtual LMOs, respectively. To recover 99% of the second-order correlation energy correction, we need for C_5H_{12} 66% of $I^{(2)}$, for C_7H_{16} 35% of $I^{(2)}$, for C_9H_{20} 22% of $I^{(2)}$, etc. Since the number of LMOs up to a certain "neighbourhood order" is fixed (and independent of the size of the system), the number of necessary integrals is proportional to m^2 instead of m^4 . The results obtained for all-trans polyenes show that it is also true for weakly localizable systems [25]. (In the latter case, fourth neighbours have to be taken into account to recover 99% of the second-order correlation energy correction.) Preliminary calculations show that this is essentially valid for larger basis sets, too, but some of the virtual orbitals are only weakly localized.

Acknowledgement

This work was supported by Grant No. OTKA 192(1991). Comments by the referee are also acknowledged.

References

- [1] K.C. Tang and C. Edmiston, *J. Chem. Phys.* 52(1970)997.
- [2] C.F. Bender, *J. Comput. Phys.* 9(1972)547.
- [3] M. Yoshimine, *J. Comput. Phys.* 11(1973)449.
- [4] Ph. Pendergast and W.H. Fink, *J. Comput. Phys.* 14(1974)286.
- [5] N.W. Winter, W.C. Ermler and R.M. Pitzer, *Chem. Phys. Lett.* 19(1973)179.
- [6] R.M. Pitzer, *J. Chem. Phys.* 58(1973)3111.
- [7] G.H.F. Diercksen, *Theor. Chim. Acta* 33(1974)1.
- [8] J. Almlöf, *Int. J. Quant. Chem.* 8(1974)915.
- [9] E.R. Davidson, *J. Chem. Phys.* 62(1974)400.
- [10] J.A. Pople, R. Seeger and R. Krishnan, *Int. J. Quant. Chem.* S11(1977)149.
- [11] V.K. Saunders and J.H. van Lenthe, *Mol. Phys.* 48(1983)923.
- [12] P. Carsky, B.A. Hess, Jr. and L.J. Schaad, *J. Comput. Chem.* 5(1984)280.
- [13] P.D. Dacre, *Chem. Phys. Lett.* 7(1970)47.
- [14] M. Dupuis and H.F. King, *Int. J. Quant. Chem.* 11(1977)613.
- [15] P.R. Taylor, *Int. J. Quant. Chem.* 27(1985)89.
- [16] C. Kozmutza, *Theor. Chim. Acta* 60(1981)53.
- [17] F. Bartha, E. Kapuy and C. Kozmutza, *J. Mol. Struct. THEOCHEM* 122(1985)205.
- [18] C. Kozmutza, *J. Mol. Struct. THEOCHEM* 123(1985)391.
- [19] C. Kozmutza, *J. Comput. Chem.* 8(1987)1179.
- [20] S. Wilson, in: *Methods in Computational Chemistry*, ed. S. Wilson (Plenum Press, New York, 1987), pp. 251–309.
- [21] E. Kapuy, Z. Csépes and C. Kozmutza, *Int. J. Comput. Chem.* 23(1983)981.
- [22] E. Kapuy, Z. Csépes and C. Kozmutza, *Croatica Chem. Acta* 57(1984)855.
- [23] E. Kapuy, F. Bartha, F. Bogár and C. Kozmutza, *Theor. Chim. Acta* 72(1987)337.
- [24] E. Kapuy, F. Bartha, C. Kozmutza and F. Bogár, *J. Mol. Struct. THEOCHEM* 170(1988)59.
- [25] E. Kapuy, F. Bartha, F. Bogár, Z. Csépes and C. Kozmutza, *Int. J. Quant. Chem.* 38(1990)139;
E. Kapuy, F. Bogár, F. Bartha and C. Kozmutza, *J. Mol. Struct. THEOCHEM* 233(1991)61.
- [26] E. Tfirst, C. Kozmutza and E. Kapuy, *J. Mol. Struct. THEOCHEM* 227(1991)93.