Theor Chim Acta (1993) 86: 285–296

Theoretica
Chimica Acta

Ab initio molecular orbital calculations on large lattice cluster models: Use of translational symmetry

Tapani A. Pakkanen and Juha Muilu

Department of Chemistry, University of Joensuu, P.O. Box 111, SF-80101, Joensuu, Finland

Received September 21, 1992/Accepted December 29, 1992

Summary. Translational symmetry has been shown to be useful in the calculation of electronic structures of large lattice models. The number of unique integrals has been derived for cases of different dimensionality. For the unique integrals zero screening and approximation methods are described. The method has been applied to arrays of hydrogen atoms and to a zincblende surface model. When the size of the system is increased the translationally unique integrals are shown to become either zero or they can be calculated by simple coulombic approximations.

Key words: Translational symmetry – Two-electron integrals – Integral approximation – Cluster model – Hartree–Fock method

1. Introduction

Ab initio Hartree–Fock calculations of chemical systems have become a useful tool for chemists. Due to the development of computational techniques and computer hardware molecules of moderate complexity can be studied and the accuracy of the calculated properties is in many cases comparable to the measurement. The main difficulty of the Hartree–Fock method is the rapidly increasing number of electron repulsion integrals to be evaluated when the size of the system increases. Integral evaluation has been speeded up by various computational techniques [1–6] and the storage problem has been avoided by a direct SCF method, where the integrals are evaluated as needed in the SCF step [7].

Computational studies of solid surfaces pose particularly difficult problems. A good model for a surface site requires many atoms in order to reproduce the properties of the real system. The atoms of the solid are often heavy atoms with many active valence electrons leading to models where the number of basis functions is very large. Practical approaches to the problem are either the limitation of the model to a few atoms or attempt to model the surroundings of the site with approximate methods. There are numerous examples of such studies [8–12].

An alternative approach to surface calculations is to use the translational symmetry of the solid system. This leads to a different formalism, but allows chemical properties to be evaluated also on surface sites. The obvious limitation is the restriction to periodic systems, where the asymmetric unit is repeated. This method has been successfully used for several solids and surface systems [13]. The symmetry properties of the wavefunctions of a system are determined by the symmetry of the Hamiltonian. In an infinite period system the symmetry can be accounted for with the properties of the space group of the system. In a nonperiodic system the symmetry properties are found from the point group, respectively. In cluster surface models the latter situation applies, and the symmetry of the system is seldom high enough to reduce the computational effort significantly. The symmetry properties of the electron repulsion integrals are not limited to the symmetry of the point group, however. In a surface cluster model the translational global symmetry is broken by the edges of the model or by an adsorbate or defect in a surface site. The great majority of the two-electron integrals in the periodic region of the system are still related by translational symmetry [14]. Despite its obvious advantages, the translational symmetry of the integrals has not been exploited in molecular orbital studies of clusters until recently. Head and Dillon have reported "pseudo-lattice symmetry" calculations, where the unique integrals are identified by lattice displacement vectors [15].

In the present work we report another method for the identification of the translationally unique integrals. We also demonstrate how the computational effort of the unique set can be further reduced by error bound and coulombic approximation techniques.

2. Theory

The starting point of the study is a cluster of atoms cut from a solid with maintaining the geometry of the bulk. Surface relaxations and site perturbations may be introduced, but we begin with the ideal case.

The aim of our approach is to provide simple rules for the identification of the translationally unique integrals. Integrals, related by the point group symmetry are not considered, since standard methods can be used. The rules for the cases of 1D-, 2D-, and 3D-symmetry are: (the proofs can be found in [14])

I) Chain Translationally unique integrals have at least one basis function on the end atom of the chain.

II) Sheet Translationally unique integrals have at least one basis function on two adjacent edges of the sheet.

III) Parallelpiped ABC Translationally unique integrals have at least one basis function on the three adjacent faces (A, B, and C) of the parallelepiped.

The present method of classification applies to any system with translational symmetry, since the translation unit is not restricted to a basis function or an atom. The lattice spacings in a, b, and c directions as well as the angles of the lattice vectors are arbitrary. The point group symmetry of the system may be used to relate translationally unique integrals. The method is also not restriced to three dimensions, and in fact there are several cases where considerations of higher dimensions can be useful. If in a solid there are subunits with internal translational symmetry and uniform orientation, higher dimensions of translation can be assigned for the subunit. An example of such a system is a diamond lattice with two *fcc* lattices are displaced from each other by one quarter of the body diagonal. A translation along the fourth dimension transforms one *fcc* lattice to another.

The number of translationally unique integrals in a general three-dimensional case is:

$$\begin{aligned}
&+ Q[ijkm] \\
&- Q[ij(k-1)m] - Q[i(j-1)km] - Q[(i-1)jkm] \\
&+ Q[i(j-1)(k-1)m] + Q[(i-1)j(k-1)m] + Q[(i-1)(j-1)km] \\
&- Q[(i-1)(j-1)(k-1)m]
\end{aligned} \tag{1}$$

where $Q[n] = \{[n(n+1)/2][n(n+1)/2+1]\}/2$ and m is the number of basis functions in the translational unit, i, j, and k are the maximum numbers of translations to a, b, and c directions.

An illustration of translational symmetry relationships can be seen in Table 1, where the number of integrals for different systems of one basis function atoms are given.

2.1 Properties of the symmetry unique integrals

Thus far it has been assumed that all of the symmetry unique integrals are calculated and stored. The translationally unique integrals may be divided into classes which have interesting properties: As the number of the translational units increases the new integrals are either small or easy to approximate. The case of the chain illustrates this classification. If the integrals for a chain of n atoms have been evaluated, the number of additional translationally unique integrals for a chain of $n+1$ atoms is $p = Q[nm] - 2Q[(n-1)m] + Q[(n-2)m]$, where p increases rapidly. All p integrals have a common feature, namely, of the four basis functions in the integral $\langle ijkl \rangle$:

$$\langle ijkl \rangle = \left\langle i(1)j(1) \left| \frac{1}{r_{12}} \right| k(2)l(2) \right\rangle \tag{2}$$

there is at least one basis function on each end of the chain. When these two functions belong to the same particle the integral diminishes rapidly with the increasing distance because of the small overlap and the short range nature of the exchange interactions. For example, letting i and j denote basis functions on opposite ends of the chain, the integral $\langle ijkl \rangle$ is small irrespective of the location of k and l since:

$$\left| \left\langle ij \left| \frac{1}{r_{12}} \right| kl \right\rangle \right| \leqslant \left\langle ij \left| \frac{1}{r_{12}} \right| ij \right\rangle^{\frac{1}{2}} \left\langle kl \left| \frac{1}{r_{12}} \right| kl \right\rangle^{\frac{1}{2}} \tag{3}$$

where the ij exchange integral is small. In the other case, when the end atom basis functions belong to different particles (e.g. $i(1)k(2)$ with i and k on the opposite ends of the chain), integrals can be significant in magnitude, but only if they are of coulombic type. If the electron repulsion integral is considered as an interaction of two charge clouds defined by basis functions i, j, k, and l in an integral belonging to the translationally unique set p with large n either one of the two charges is very small since its basis function are in the ends of the chain or both charges are localized in the opposite ends of the chain. This feature gives possibilities to prescreen zero integrals from the calculation and approximate coulombic integrals with multipole approximations. In summary, on increasing the chain length, the number of nonnegligible new integrals is much smaller than given by the formula for p, and the integrals are primarily of long range coulombic type. The short-range interaction integrals do not proliferate with the increasing size of the system, if translational symmetry is used. Since this type of integral is the most difficult to calculate, symmetry recognition complements other evaluation procedures.

3. Computational techniques

The use of the translational symmetry in *ab initio* Hartree–Fock calculations is rather straightforward. Symmetry can be used for one-electron and two-electron integrals by implementing the principles described above. The principle can be also extended to valence calculations, where symmetry is used in the core-valence and valence integrals [16]. It is useful to prescreen the two-electron integrals with the exchange integral test in order to avoid the accurate computation of the zero integrals.

An option to exclude atoms from the SCF step was built to the program for systems where the desired model is a subset of a translationally ideal system.

In the SCF part all nonzero integrals are regenerated for the construction of the Fock matrix. Since the number of integrals to be stored is relatively small, they can be kept in the main memory.

3.1 Monopole approximation

A simple monopole approximation for the long-range coulombic integrals was tested as well. The charge densities of the integrals were approximated by point charges and the magnitude of the integral was computed from Coulombs law. A practical way to use the monopole approximation is to study pairs of charge distributions. At certain distance the monopole approximation becomes accurate enough. Once this limit is found all integrals with the same charge distributions but with longer distance can be approximated.

4. Numerical applications

4.1 Array of hydrogen atoms

4.1.1 One-dimensional array. The simplest case to demonstrate the use of translational symmetry is a linear array of one basis function atoms. The atoms were chosen hydrogens and the basis is a standard five component lobe basis. The distance of adjacent atoms is 1.5 a.u. (78 pm), where the atoms are within a bonding distance. The model simulates metal lattices where the basis functions overlap extensively neighboring centers.

The first system considered was a chain of 900 hydrogens. In this 900 basis function system there are 8.22×10^{10} two electron integrals of which 3.64×10^{8} are symmetry unique according to (I) as shown in Table 1. This is about 0.4% of the total two-electron integrals.

Table 1. Number of total (*T*), unique (*U*), unique nonzero (*N*) and exact unique (*E*) two-electron integrals of hydrogen chain, slab, and parallelepiped bulk

H	*T*	*U*	N^a	E^a	$(U \times 100)/T$	$(N \times 100)/T$	$(N \times 100)/U$	$(E \times 100)/N$
Chain								
1	1	1	1	1	100.00	100.00	100.00	100.00
2	6	5	5	5	83.33	83.33	100.00	100.00
5	120	65	65	65	54.17	54.17	100.00	100.00
10	1540	505	363	276	32.79	23.57	71.88	76.03
50	813450	62525	2603	904	7.69	0.32	4.16	34.73
100	12753775	500050	5403	974	3.92	0.04	1.08	18.03
500	7843843875	62500250	27803	974	0.80	0.00	0.04	3.50
900	82195053975	364500450	50203	974	0.44	0.00	0.01	1.94
Slab								
5 × 5	52975	17981	17962	17828	33.94	33.91	99.89	99.25
10 × 10	12753775	1496836	405004	134798	11.74	3.18	27.06	33.28
20 × 20	3216060100	110431421	4219824	338834	3.43	0.08	2.19	14.00
30 × 30	82195053975	1320604756	6116644	482528	1.61	0.01	0.46	7.89
Bulk								
2 × 2 × 2	666	518	518	518	77.78	77.78	100.00	100.00
5 × 5 × 5	31011750	6337445	6189278	4939966	20.44	19.96	97.66	79.81
10 × 10 × 10	125250375250	5088989890	382997364	42684079	4.06	0.31	7.53	11.14

[a] Integral rejection threshold is 10^{-7}

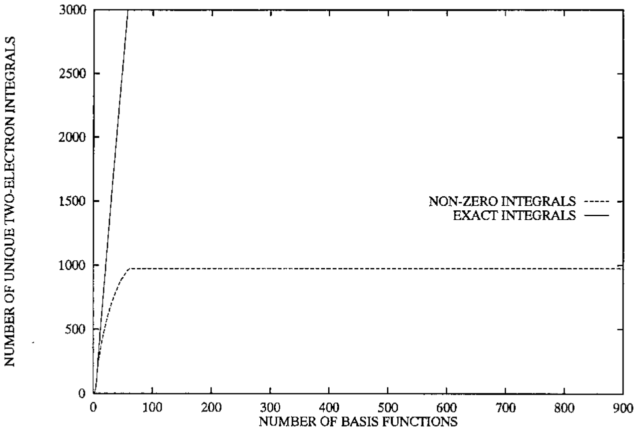

Fig. 1. Number of unique nonzero and exact two-electron integrals as the function of the size of the hydrogen chain. Integral rejection threshold if 10^{-7}

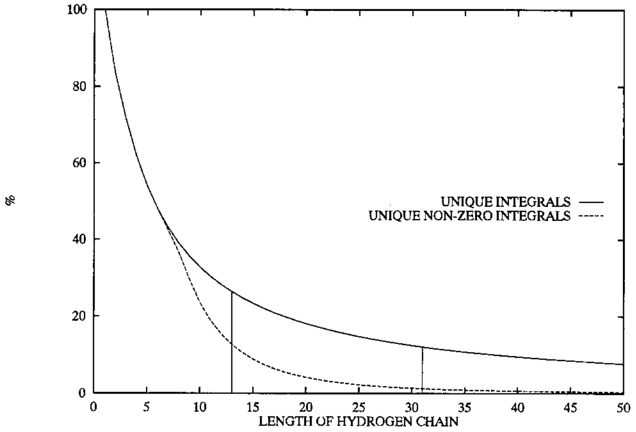

Fig. 2. Percentage of the unique and nonzero unique two-electron integrals of total. Chain of hydrogen atoms. Integral rejection threshold is 10^{-7}

Further and a more dramatic decrease in the integrals can be obtained by using the exchange bound (3). When the threshold is 10^{-7} only 50203 integrals remains to be computed. A majority of these integrals can be computed by the monopole approximation, leaving only 974 integrals for the full evaluation. These integrals have been already encountered within the chain length of 65 atoms. The number of integrals vs. number of basis functions is shown in Fig. 1.

A closer inspection of the shorter chain lengths shows the importance of the exchange error bound. In Fig. 2 the percentage of translationally unique and nonzero integrals of total are shown. At the chain length of 13 atoms 50% of the translationally unique integrals are zero, and at 31 atoms the corresponding proportion is 90%.

The definition of zero integral was studied by varying the criteria of rejecting integrals as zero. Figure 3 shows how the percentage of the zero integrals out of the translationally unique set varies as the function of the chain length. An order of magnitude change in the criteria shifts the curve to the left by about 2 "atoms". A more noticeable effect is seen if the distance of the lattice atoms is increased. In Fig. 4 there are three cases, namely 1.5 a.u., 1.75 a.u., and 2.0 a.u.

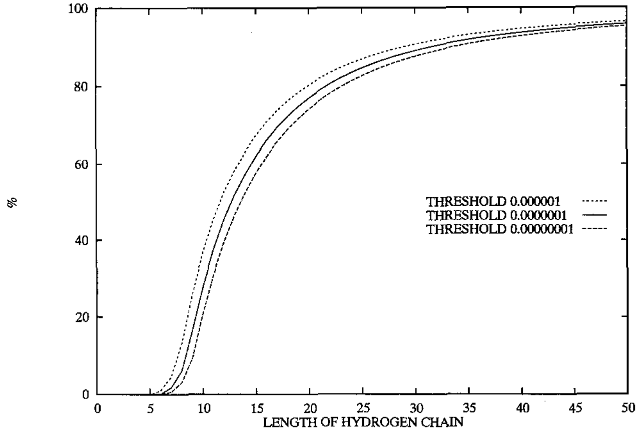

Fig. 3. Percentage of the zero integrals out of translational unique set. Chain of hydrogen atoms. Influence of the threshold.

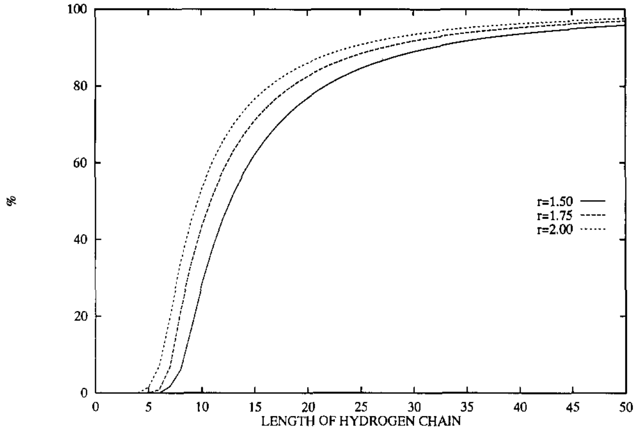

Fig. 4. Percentage of the zero integrals out of the translational unique set. Chain of hydrogen atoms. Influence of the bond length. Integral rejection threshold is 10^{-7}

The same shape of the curve is again reproduced, but now the shift is towards shorter chain lengths as expected.

For practical computational purposes we have shown that the exact integrals for chains of any length can be easily computed, and for the remaining coulombic integrals can be calculated with the simple monopole formula.

4.1.2 Two- and three-dimensional arrays. The computational effort increases considerably when two and three-dimensional systems are studied. The main reason is the high concentration of atoms in a small space, where there the overlaps of the basis functions remain large. The translational symmetry is also quite as useful as in the chain case. As a two-dimensional case an orthogonal slab of 30 × 30 hydrogen atoms was computed. According to Table 1 the 900 (30 × 30) basis functions generate 8.22×10^{10} two electron integrals. The number of symmetry unique integrals is now 1.32×10^{9}, which is about 1.6% of the total. This is about four times more than in the chain case. Number of unique nonzero integrals is 6.11×10^{6} which is about 0.01% of the total. Monopole approximation can be used for 92% of these leaving 482528 integrals for a full calculation. Figure 5 shows the dependence of the integrals on the number of basis functions.

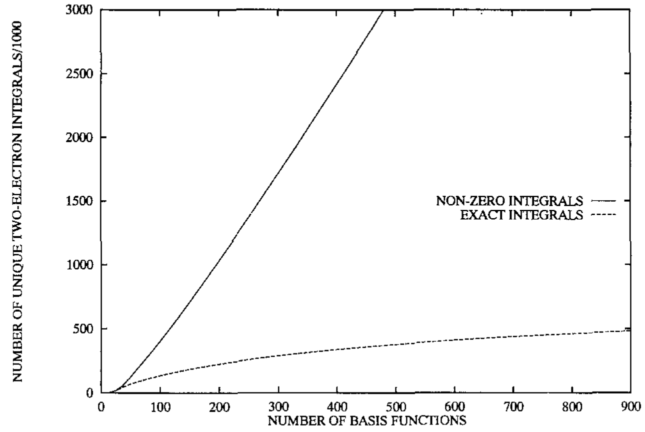

Fig. 5. Number of unique nonzero and exact two-electron integrals as the function of the size of the hydrogen slab. Integral rejection threshold is 10^{-7}

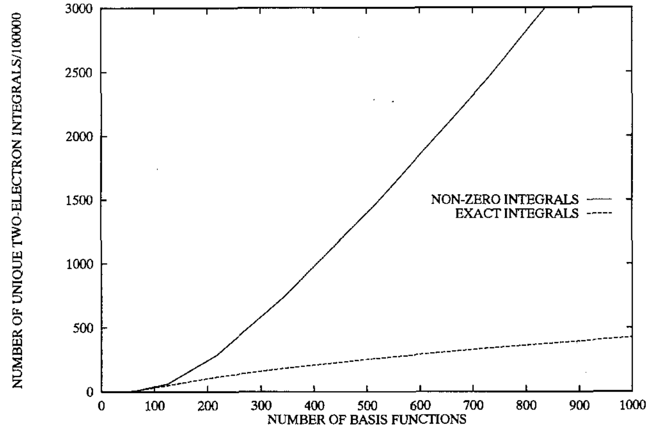

Fig. 6. Number of unique nonzro and exact two-electron integrals as the function of the size of the hydrogen bulk. Integral rejection threshold is 10^{-7}

The nonzero integrals depend almost linearly and the "exact" integrals level off approaching a constant number as in the chain case.

A three-dimenisonal system comprising $10 \times 10 \times 10$ atoms was also considered. As shown in Table 1 the 1000 basis functions produce 1.25×10^{11} integrals of which 5.10×10^{9} (4%) are symmetry unique and 3.83×10^{8} (0.3%) are nonzero. Of the latter 89% can be approximated by the monopole approximation leaving 4.2×10^{7} integrals for full evaluation. Figure 6 shows again the almost linear dependence of the integrals on the number of basis functions. The amount of integrals demanding full calculation is again growing slowly and approaching a constant as in the previous cases.

In order to compare results of different dimensionality the number of basis functions was chosen as the common x-axis. In Fig. 7 the percentage of the unique and unique nonzero integrals of total is shown. The figure indicates that the proportion of the translationally unique and nonzero integrals decreases much faster in the cases of lower dimensionality.

If the size of the system is considered as the variable the proportion of the zero integrals within the symmetry unique set starts to dominate faster in higher

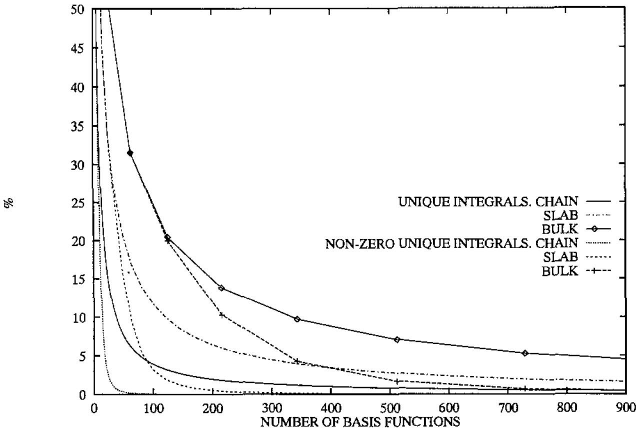

Fig. 7. Percentage of unique and unique nonzero integrals of total. All hydrogen models. One-, two- and three-dimensional hydrogen models are compared as the function of the size of the system. Integral rejection threshold is 10^{-7}

dimensions. About 50% of the integrals become zero at the slab size of 8 × 8 and 90% are zero when the size is 13 × 13.

In the three-dimensional case the 50% proportion of nonzero integrals is at 6 × 6 × 6, and the 90% proportion is at model size 9 × 9 × 9, as compared to the respective chain results of 13 and 31.

4.2 Zinc sulfide surface model

As a test of a real lattice system zinc sulfide surface was chosen. The surface properties of II–VI compounds are of considerable practical interest in view of their electrical and electro-optical properties [16–18]. In the present work only the computational aspects of the surface models are considered, the electronic structures will be reported separately. The models represent the (111) surface of zincblende (Fig. 10). The *ab initio* calculations were limited to the valence region of the system, the influence of the atomic cores were taken into account by the stepwise approximation method based on density matrix expansions [8, 9, 16].

In zinc sulfide the translational symmetry was used as follows: ZnS unit was chosen to be the translational unit, which was transferred along two vectors which are in a 120° angle (Fig. 8). Minimal basis was chosen, since it has been shown to reproduce the surface properties of the system at least qualitative-ly [17]. The core and valence integrals were computed using translational symmetry.

The results for the valence integrals of surface models up to $Zn_{100}S_{100}$ are shown in Table 2 and in Fig. 9. The largest 10 × 10 model has 500 s- and p-basis

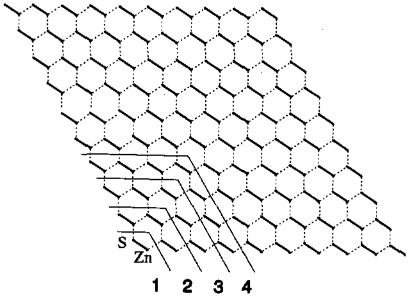

Fig. 8. Two-dimensional zinc sulfide ($Zn_{100}S_{100}$) model. Models of 1 × 1, 2 × 2, 3 × 3, and 4 × 4 zinc sulfide units are indicated

Table 2. Number of total (T), unique (U), nonzero unique (N) and exact unique (E) two-electron integrals as increasing the zinc sulfide slab

Zn_xS_x	S	T	U	N^a	E^a	$(U \times 100)/T$	$(N \times 100)/T$	$(N \times 100)/U$	$(E \times 100)/N$
Zn_1S_1	1	120	120	120	120	100.00	100.00	100.00	100.00
Zn_4S_4	2	22155	19195	17982	13521	86.64	81.16	93.68	75.12
Zn_9S_9	3	536130	341595	151367	87346	63.71	28.23	44.31	57.70
$Zn_{16}S_{16}$	4	5250420	2435820	411083	197871	46.39	7.83	16.88	48.12
$Zn_{25}S_{25}$	5	31011750	10754620	785381	336330	34.68	2.53	7.30	42.82
$Zn_{36}S_{36}$	6	132690195	35434995	1279767	501719	26.71	0.96	3.61	39.20
$Zn_{49}S_{49}$	7	454074180	95898195	1894341	691177	21.12	0.42	1.98	36.49
$Zn_{64}S_{64}$	8	1318950480	225349720	2628803	903636	17.09	0.20	1.17	34.37
$Zn_{81}S_{81}$	9	3379694220	476179320	3483453	1137405	14.09	0.10	0.73	32.65
$Zn_{100}S_{100}$	10	7843843875	926260995	4458191	1391198	11.81	0.06	0.48	31.20

[a] Integral rejection threshold is 10^{-7}

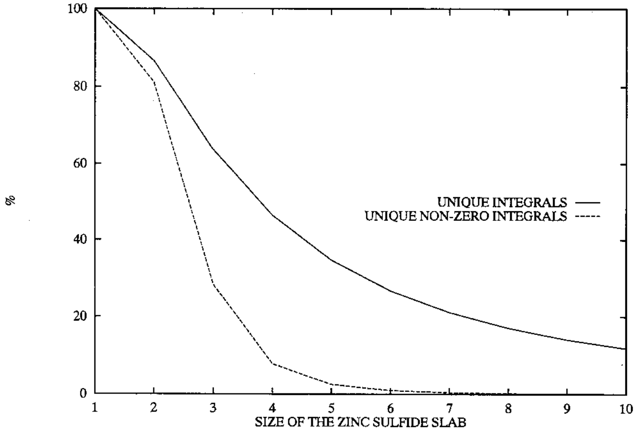

Fig. 9. Percentage of unique and unique nonzero two-electron integrals of total. Slab of zinc sulfide. Integral rejection threshold is 10^{-7}

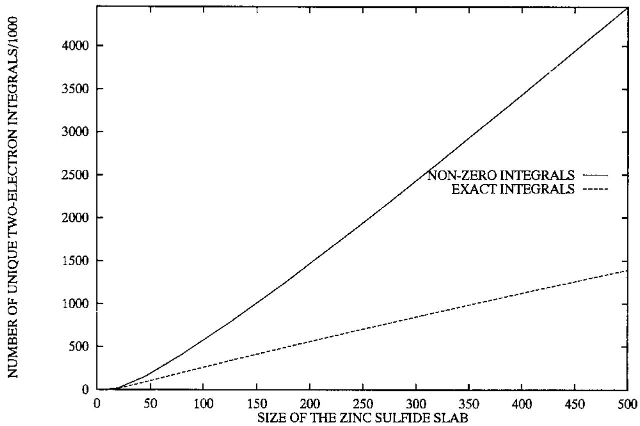

Fig. 10. Number of unique nonzero and exact two-electron integrals as the function of the zinc sulfide slab. Used threshold is 10^{-7}

functions, which generate 7.84×10^9 two electron integrals of which 9.26×10^8 are translationally unique (11.8%). The number of nonzero integrals is 4.45×10^6 which is only 0.1% of the total. Coulombic approximation reduces the number of exact integral further to 1.2×10^6. As a function of the basis the number of integrals increases linearly as in the hydrogen case (Fig. 10). 90% of the translationally unique integrals are zero at the system size of 5 × 5. In the hydrogen case the same level was found for model size 15 × 15. This shows the power of the method in a real application with more than one basis function in the translational unit. In the zinc sulfide case the lattice is more open and the basis functions do not overlap as well as in the hydrogen case, which represents a tightly packed metallic case. The power of translational symmetry is less in the cases of many basis functions in the translational unit and high dimensionality but the exchange bound and the monopole approximation gives a more favorable overall reduction in the computational task.

5. Conclusions

Translational symmetry together with integral approximations has been shown to be an efficient approach for *ab initio* studies of large lattice models. The number of unique nonzero integrals form only a small fraction of the total integrals. Large lattice models allow studies of local properties vs. model size, which has been a topic of extensive discussions in the literature.

Acknowledgements. Discussions with Prof. Jerry Whitten at an early stage of the work are acknowledged.

References

1. McMurchie LE, Davidson ER (1978) J Comput Phys 26:218
2. Pople JA, Hehre WJ (1978) J Comput Phys 27:161
3. Rys J, Dupuis M, King HF (1983) J Comput Chem 4:154
4. Obara S, Saika A (1986) J Chem Phys 84:3963
5. Head-Gordon M, Pople JA (1988) J Chem Phys 89:5777
6. Lindh R, Ryu U, Liu B (1991) J Chem Phys 95:5889
7. Almlöf J, Faegri Jr K, Korsell K (1981) J Comput Chem 3:385
8 Pakkanen TA, Whitten JL (1978) J Chem Phys 69:2168
9. Whitten JL, Pakkanen TA (1980) Phys Rev B 21:4357
10. Whitten JL (1981) Phys Rev B 24:1810
11. Madhavan PV, Whitten JL (1981) Surf Sci 112:38
12. Cremaschi P, Whitten JL (1981) Surf Sci 112:343
13. Pisani C, Dovesi R, Roetti C (1988) Hartree–Fock *ab initio* treatment of crystalline systems. Lecture Notes in Chemistry 48. Springer-Verlag, Berlin
14. Pakkanen TA (1977) Ph.D. thesis, State University of New York at Stony Brook
15. Head JD, Dillon IP (1991) Theor Chim Acta 78:231
16. Pakkanen TA, Linblad M, Laitinen RS (1986) Int J Quant Chem XXIX:1789
17. Lindblad M, Pakkanen TA (1988) J Comput Chem 9:581
18. Pakkanen TA, Nevalainen V, Lindblad M, Makkonen P (1978) Surf Sci 188:456