The Generalized Slater–Condon Rules

JACOB VERBEEK AND JOOP H. VAN LENTHE

Theoretical Chemistry Group, University of Utrecht. Padualaan 14, 3584 CH Utrecht, The Netherlands

Abstract

By relating the blocking structure of the relevant matrix of overlap-integrals to its cofactors, the Slater–Condon rules for the evaluation of an element of a matrix representation of an electronic Hamiltonian in a Slater determinant basis are generalized to the case where not all orbitals are orthogonal. This yields a set of 33 rules, which allows for an efficient implementation of the valence bond theory.

1. Introduction

The evaluation of elements of a matrix representation of one- and two-electron operators in a Slater determinant basis is a straightforward exercise in case the orbitals are orthogonal. The differences between the occupation numbers of the spin orbitals in the two Slater determinants in question tell what integrals should be added to yield the matrix element. These so-called Slater–Condon rules [1, 2] are a cornerstone of ab initio quantum chemistry and have contributed greatly to the success of orthogonal orbitals. If the orbitals are nonorthogonal, the occupation numbers are not so easily related to the structure of the matrix element in terms of the one- and two-electron integrals.

In this paper, we show that the underlying principle of the Slater–Condon rules is the difference in the numbers of orbitals per mutually orthogonal subspace. In the orthogonal case, each spin orbital coincides with such a subspace.

To make our point clear, let us look at Löwdin’s general formula for a matrix element [3, 4]:

$$\langle A | H | B \rangle = Tr[\mathbf{h}_{ab} \cdot \text{adj}(\mathbf{S}_{ab})] + Tr[\mathbf{G}_{ab} \cdot \text{adj}^{(2)}(\mathbf{S}_{ab})].$$

(1)

A and B are normalized Slater determinants constructed from the spin orbitals \{a\} and \{b\}, respectively. H is the electronic Hamiltonian, which consists of a one-electron part \(h\) and a two-electron part \(g\). Tr denotes the trace of a matrix, \(\mathbf{h}_{ab}\) (one-electron integrals), \(\mathbf{S}_{ab}\) (overlap integrals), and \(\text{adj}(\mathbf{S}_{ab})\) (first-order adjugate) are \(N\)-dimensional matrices, and \(\mathbf{G}_{ab}\) (two-electron integrals) and \(\text{adj}^{(2)}(\mathbf{S}_{ab})\)
(second-order adjugate) are \( N \cdot (N - 1)/2 \) dimensional matrices, where \( N \) is the number of electrons. These matrices are defined as

\[
\begin{align*}
(h_{ab})_k & = \langle a(1)|h(1)|b_k(1) \rangle \\
(G_{ab})_{ij} & = \langle a(1)a_j(2)|g(1, 2)|b_i(1)b_1(2) \rangle \\
- \langle a(1)a_j(2)|g(1, 2)|b_i(2)b_1(1) \rangle \\
(S_{ab})_k & = \langle a(1)|b_k(1) \rangle \\
(adj(S_{ab})) & = S_{ab}^{(k, 0)} \\
(adj^2(S_{ab}))_{i} & = S_{ab}^{(k, l, i, j)}.
\end{align*}
\]

(2)

\( S_{ab}^{(k, 0)} \) and \( S_{ab}^{(k, l, i, j)} \) are first- and second-order cofactors of the matrix of overlap-integrals \( S_{ab} \). The \((i, j)\)-th cofactor is defined as the signed determinant (minor) of the submatrix obtained from a matrix by deleting the \( i \)-th row and the \( j \)-th column, the sign being \((-1)^{i+j}\). Likewise, the \((i, j, k, l)\)-th cofactor corresponds to deleting the \( i \) and \( j \)-th rows and the \( k \) and \( l \)-th columns and assigning \((-1)^{i+j+k+l}\) [5]. The matrices carrying four labels are supermatrices; the first two labels, \( i \) and \( j \), refer to rows; the second two labels, \( k \) and \( l \), refer to columns. Typically, \( ij \) corresponds to \((j - 1) \cdot (j - 2)/2 + i \), with \( i < j \).

Löwdin's formula tells what the roles of the one- and two-electron integrals are. The matrix element is a linear function of every integral, the weight of each integral being its cofactor. This also holds in the orthogonal case, of course, but then almost all weights are zero, and if not zero, they are one. The sole function of the adjugates in the orthogonal case is to identify the relevant integrals that should be added to yield the matrix element. This selection can be done right away using the Slater–Condon rules, so that, in the orthogonal case, the adjugates can be dispensed with [4]. One also might look at it this way: The Slater–Condon rules predict the relation between \( S_{ab} \) and its first- and second-order adjugates. If we want to generalize these rules to the nonorthogonal case, we have to relate the structure of \( S_{ab} \) to the structure of the adjugates. By structure we mean the systematic distribution of nonzeros. Once this relation is established, by inspecting the structure of \( S_{ab} \) and using the generalized Slater–Condon rules, the summations in Löwdin's formula can be restricted to the nonzero contributions, which greatly improves the efficiency.

In general, the matrix of overlap-integrals consists of two blocks belonging to alpha and beta spin. To deal with spin orthogonality properly, this structure must be recognized, because the overlap-integrals usually refer to orbital, not to spin orbitals. This has led other workers [6–9] (see also [13]) to formulate rules that also occur in our set (cf. section 2.6).

The exploitation of the block-diagonal structure of \( S_{ab} \) has been hinted at by Prosser and Hagstrom [10] in 1968. To our knowledge, the rules have not yet been derived in extenso.
2. The Relation between the Sparsity of a Matrix and the Sparsities of Its First- and Second-Order Adjugates

From now on, the subscript \(ab\) will be omitted from \(S\), but it must be remembered that \(S\) is different for different combinations of Slater determinants. It is assumed that the square matrix \(S\) has a block-diagonal form. The blocks sometimes will be referred to as matrices, e.g., "the determinant of a block." All blocks are taken to be decomposed by elimination [4,10], which is the first step in the cofactor evaluation, and reveals the rank of each block.

To establish the desired relation between \(S\) and its first- and second-order adjugates precisely, 33 distinct cases will be recognized, which fall into three broad categories: \(S\) having a nullity of 0, 1, or 2. The nullity is dimension minus rank. For electronic Hamiltonian matrix elements, a nullity of three or more implies the matrix element is zero.

2.1. The Nullity of \(S\) Equals 0

If \(S\) is nonsingular, all the blocks of nonzeros must be square [5]. Therefore, the only choices of \(i\) and \(j\) that will lead to a nonzero first-order cofactor will be those that involve one block at a time. (See Fig. 1.)

The first-order adjugate is seen to have the same block-diagonal structure as \(S\). Each nonzero first-order cofactor now equals a first-order cofactor of the block at hand times the product of determinants of the other blocks. In actual computation, this means that the blocks can be decomposed by elimination separately, and given the blocking structure of \(S\), the first-order cofactors that lie outside a block simply do not have to be considered.

Similarly, the choices of \(i, j, k,\) and \(l\) that will yield a nonzero second-order cofactor are those that leave behind a submatrix having square blocks. One then can distinguish two kinds:

1. Those that involve one block only (first kind) and therefore are equal to a second-order cofactor of the block at hand times the product of the determinants of the other blocks.

![Figure 1. Symbolic representation of first-order cofactors of block-diagonal non-singular matrices.](image-url)
2. Those that involve two blocks \((i \text{ and } k \text{ in block } a, j \text{ and } l \text{ in block } b\), second kind) and therefore are equal to the first-order cofactor \((i,k)\) of block \(a\) times the first-order cofactor \((j,l)\) of block \(b\) times the product of the determinants of the other blocks. (See Fig. 2.)

The second-order cofactors of the first kind can be found by means of the Jacobi ratio theorem [3, 4, 5, 10], considering each block as a separate matrix. The second kind of second-order cofactors are easily constructed from the first-order cofactors. In practice, those of the second kind are (much) more numerous, so it is advantageous that they take just one operation each, as opposed to the three operations brought along by the Jacobi ratio theorem.

The case of nullity 0 will be referred to as “case (0).” A cofactor of a submatrix that is multiplied by the determinants of the other blocks, will be called a “weighted cofactor.”

2.2 A Graphical Representation of Singular Block Diagonal Matrices

The principle of restricted choice, as outlined in the previous section, already greatly reduces the number of terms in Löwdin's formula. However, there is much more to be gained, once the consequences of the singularity of \(S\) are appreciated.

If \(S\) is singular, things get more complicated and a simplification of the graphical representation is helpful. Blocks that are not singular will not be depicted. Lines symbolize rows or columns, which, when deleted, cause the cofactors to be nonzero for the following reasons (below, the numbers in parentheses refer to Fig. 3):

(a) A row (column) consisting of only zeros must be deleted, and this forces us to fix the corresponding index, as other terms of the pertinent sum will have zero cofactors (1 and 2).
(b) If a block of the nonzero overlap integrals is rectangular, one has to cross out as many rows (columns) as to get a square submatrix that can have a nonzero determinant; this forces us to take the corresponding index(es) from the given block; otherwise, the cofactor vanishes (3, 4, 6, and 7).
(c) If a square \(m \times m\) block has rank \(m - k\) \((k = 1\text{ or } 2)\), the cofactor can differ from zero only if \(k\) rows and \(k\) columns are striked out. Thus, \(k\)

![Figure 2. Symbolic representation of second-order cofactors of block-diagonal nonsingular matrices.](image-url)
deleted rows and \( k \) deleted columns must be within the block (5 and 10).

(d) If a rectangular \( m \times n \) matrix \( (m = n + 1) \) has rank \( n - 1 \), then two rows and one column must be deleted, which restricts the indices of two rows and one column to be within the block (8).

e) The same as (d) but with rows and columns interchanged (9).

The ten structures just described are shown in Figure 3. These pictures represent the fundamental structures occurring in singular block-diagonal matrices with nullity up to two. They will be referred to as “shapes.”

There are two rules that should be fulfilled by combinations of shapes.

Rule 1. The total number of restrictions on row indices (i.e., horizontal lines) must be equal to the total number of restrictions on column indices (vertical lines).

Rule 2. The total number of restrictions on row (or column) indices equals the nullity of \( S \).

Note on rule 1: If \( S \) consists of just one block, an index may be restricted to the entire region of the block, e.g., if only the last row consists of zeros, one row index is restricted to that row, whereas the column index/indices may be chosen freely. In actual computation, one will have alpha and beta electrons, so that their will be two blocks at least, due to the spin orthogonality. Then, if, say, the last row of the alpha block consists of zeros, one column index indeed is restricted to the alpha block: The weighted second-order cofactors of the beta block are zero.

2.3. The Nullity of \( S \) is 1

Rule 2 shows that in case \( S \) has a nullity of 1 only shapes 1 to 5 can occur in \( S \). Using rule 1, five proper combinations can be found (see Fig. 4). In Figure 4, the numbers between the parentheses refer to the numbering of the shapes in Figure 3. The proper combinations of shapes will be referred to as “cases.” It is noted that the cases may be defined in different ways, thereby possibly yielding a different number of cases. For instance, by interchanging the Slater determinants, case (2, 3) may be converted to case (1, 4), without changing the value of the matrix element. We will not allow for such conversions. We will allow for moving shapes within a spin block, without changing case.

Per case, the nonzero first-order cofactors of \( S \) are those that correspond to the choices of the rows and columns as indicated in Figure 4. It is seen that for

![Figure 3. Ten basic shapes in singular block-diagonal matrices.](image-url)
cases (1, 2), (1, 4), (2, 3), and (5), these cofactors are weighted first-order cofactors of the singular block. For (3, 4), the first-order cofactors are the weighted products of the first-order cofactors of the two rectangular blocks. By the cofactors of a rectangular block, we mean the signed minors that can be obtained from the block via the deletions of the rows or columns that leave behind square blocks (see Section 2.5 about the parity signs).

The second-order cofactors now are (1) the weighted second-order cofactors of the singular substructure, if any, and (2) the weighted products of the first-order cofactors of the substructure and the first-order cofactors of the other, nonsingular blocks. Again, those of the second kind are the cheapest and the most numerous.

2.4 The Nullity of $S$ is 2

Now, many more combinations of shapes arise, though they can be enumerated without great difficulty. All first-order cofactors are zero.

First, the five singular cases, shown in Figure 4, can be combined among themselves, yielding 15 \([-5(5+1)/2\)] proper combinations. All nonzero second-order cofactors are weighted products of the first-order cofactors of the two subcases that are combined in a given case. Note that the cases (1, 2)(3, 4) and (1, 4)(2, 3) essentially are the same. We regard them as different because each subcase may reside in a different spin block. This then exhausts the possibilities involving shapes 1 to 5 only. As far as shapes 6 to 10 are concerned, there is just one proper combination, namely, 6 with 7. In that case, the second-order cofactors are the weighted products of the second-order cofactors of both rectangular blocks. The remaining of the possibilities is found by combining shapes 6 to 9 each with shapes 1 to 5, using rule 1 and rule 2. This adds another 10 cases to the list, in all of which the second-order cofactors are the weighted products of the cofactors of the blocks involved. Finally, one has the case of shape 10, which is a case on its own. Adding up all possibilities, we arrive at 27 different cases for matrices with a nullity of two (see Table I). The numbering in the table refers to Figures 3 and 4.

<table>
<thead>
<tr>
<th>Table I. The 27 cases of doubly singular block-diagonal matrices.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1,2),(1,2)</td>
</tr>
<tr>
<td>(1,4),(1,4)</td>
</tr>
<tr>
<td>(2,3),(3,4)</td>
</tr>
<tr>
<td>(2),(2),(6)</td>
</tr>
<tr>
<td>(3),(3),(7)</td>
</tr>
<tr>
<td>(6),(7)</td>
</tr>
</tbody>
</table>
2.5 The Extra, Overall Parity

Above the relative, alternating parity signs, which arise naturally while calculating the cofactors of the substructures, an extra, overall parity sign should be imposed on the value of the matrix element as the matrix is not treated as a whole: A change of the relative positions of the blocks, and the rows and columns of zeros, may alter the sign of all the "true" cofactors simultaneously, whereas the schemes given in the previous sections after such a change still would yield the same cofactors.

The extra sign conveniently can be deduced according to the following reasoning: The "true" cofactors, which could be evaluated by constructing the entire $S$ matrix and applying a brute force technique, will be called "global cofactors." They consist of two parts: a parity sign, which is determined by the indices and alternates with the indices, and a value, which is equal to the determinant of the submatrix, left over after striking out the rows and columns. The cofactors of the squares and the rectangles will be called "local cofactors." In our schemes, the global cofactors are evaluated as the products of one to four local cofactor(s) and the determinants of the other blocks.

While looping over one index, it will strike out one row or one column at a time, thereby involving one block at a time. The set of local cofactors, belonging to this block, will alternate the parity sign by itself, so that, once one cofactor, which has resulted from our schemes, has been calibrated (is known to have the right parity sign), all the others must also be right. We therefore have to bother about only one cofactor, say the first. Let us define the first cofactor as the nonzero cofactor corresponding to the lowest values of the indices, i.e., the starting positions of the indices in the full matrix.

In case (0) (only squares, no rows or columns of zeros), the signs must be correct because each block can be seen to start at a diagonal position (even parity, $+$), and the first cofactors per block also have even parity. In the other cases, we have to be careful, because then the blocks do not necessarily start at diagonal positions, and for rectangles, the local cofactors may start at odd parity. In our current implementation of the algorithms that yield the cofactors of the rectangles, it is assumed implicitly that the rectangles are supplemented by rows or columns of zeros, so as to result in a square. Thereby, one or two indices are fixed immediately right from, or below, the rectangle, thus influencing the parity signs.

We do know that the submatrices corresponding to nonzero cofactors must be block-diagonal and must have each block start at a diagonal position: They are case (0) matrices. The determinant of each submatrix therefore is equal to the product of the determinants of the blocks of the submatrix. Some of these determinants may actually be local cofactors. Once care has been taken to let the first cofactors of the blocks begin at even parity, the parity sign of the first cofactor completely is determined by the starting positions of the indices in the full matrix.

The extra, overall parity thus properly is taken into account by multiplying the value of the matrix element with $(-1)^{c_{ij+k+l}}$, where $i, j, k$ and $l$ are the starting positions of the indices in the full matrix, and by letting the first cofactor of each block start at even parity. The first cofactor of a square always starts at even, but for a rectangle, this generally must be imposed.
2.6. The Generalized Slater–Condon Rules

Adding up the number of cases for a nullity of \( S \) up to 2, we arrive at 
\[ 1 + 5 + 27 = 33 \] cases of matrix elements, the recognition of which allows for 
efficient matrix element evaluation as follows: The total orbital space is divided 
in mutually orthogonal subspaces by inspecting the orbital metric. The Slater 
determinants are set up such that the orbitals appear in groups belonging to the 
same subspace, and groups occur in the same order always. By matching the 
numbers of orbitals belonging to each subspace, per matrix element the blocking 
structure of the matrix of overlap-integrals of the orbitals that comprise the 
Slater determinants is revealed. Each block of overlap-integrals is decomposed 
via elimination. The structure then is one of the 33 cases (cf. Fig. 5). The choices 
of the rows and columns to be deleted from the matrix of overlap integrals in 
order to yield nonzero cofactors follow from the case and the dimensions of the 
blocks. In general, this means the indices in Löwdin’s formula are restricted to 
certain ranges. The nonzero cofactors then can be calculated via the cofactors of 
the relevant blocks, the determinants of the other blocks, and the overall parity.

Usually, the integrals are available in the spatial part of the spin-orbital basis 
only. The number of orthogonality groups then is twice the number of groups, 
that is apparent from the orbital metric. The matrix of overlap integrals belong-
ing to two Slater determinants then will consist of two global diagonal blocks, 
corresponding to alpha and beta spin, that each have a block diagonal structure 
due to spatial orthogonality. The classification of the blocking structure still 
stands, though while performing the summations of Löwdin’s formula, care has 
to be taken not to include two-electron integrals that vanish through spin inte-
gration. If \( \langle j | k \rangle \) appears, \( \langle j | k \rangle \) should be subtracted from it only if electron 1 
and electron 2 have the same spin. Rule 1 should be applied for the alpha and 
beta blocks separately.

If all orbitals are orthogonal, the blocks at the diagonal of the matrix of over-
lap integrals consist of just one overlap-integral each. Using the rules given, the 
indices in Löwdin’s formula then will be restricted so as to result in the usual 
(“orthogonal”) Slater–Condon rules, which are the cases \( (0), (1, 2), \) and \( (1, 2)(1, 2) \). 
The rules given by the authors mentioned in the Introduction [6–9] are the cases 
\( (0), (5), (5)(5), \) and \( (10) \). (See Fig. 5.)

3. Conclusion

By recognition of the relation between the blocking structure of the relevant 
overlap matrix and the sparsity of the first- and second-order adjugates, the terms 
in Löwdin’s formula that will be zero because of the zero cofactors can be 
skipped. In general, indices of the summations will be fixed, or restricted to cer-
tain ranges, thereby greatly reducing the number of terms that have to be taken 
into account.

The rules we derived converge to the Slater–Condon rules as more orbitals be-
come orthogonal. They allow for implementation [11] of ab initio methods based 
on the use of nonorthogonal orbitals, like the \( \text{VBSCF} \) method [12], which can be-
come competitive with equivalent currently standard methods, like MCSCF, CAS, and GVB, both from the conceptual and the economical viewpoint.

Doubly occupied orbitals usually can be orthogonalized and projected out of the other orbitals without changing the wave function. The use of the generalized Slater–Condon rules then excludes the doubly occupied orbitals from the nonorthogonality problem and focuses the computational effort on the singly occupied orbitals.

Acknowledgments

We are indebted to Dr. F.B. van Duijneveldt, Dr. P.J.A. Rutting, P.H. Kruiskamp, and the referee for their comments on the manuscript.

Bibliography

[11] TURTLE, an ab initio VB/VSCF/VCI program by J. Verbeek and J. H. Langenberg,

University of Utrecht.

Received February 12, 1990
Accepted for publication July 11, 1990