

## A NEW METHOD FOR LARGE-SCALE CI CALCULATIONS

B. ROOS

*Institute of Theoretical Physics, University of Stockholm, Stockholm, Sweden*

Received 8 May 1972

A new method for obtaining the coefficients in a large CI expansion is proposed. Instead of constructing an interaction matrix, the expansion coefficients are obtained directly from the list of two-electron integrals by means of an iterative procedure. At present, the method can be applied to systems with a closed-shell ground state including all singly and doubly excited configurations. The number of terms used in the expansion of the wavefunction may be as large as  $10^4$ – $10^5$ . Wavefunctions for  $\text{NH}_3$  with 4336 and  $\text{H}_2\text{O}$  with 2458 configurations are presented as test cases.

## 1. Method

The basis for all CI calculations is a set of one-electron functions

$$\{\varphi_p\} = \{\varphi_1, \dots, \varphi_p, \dots, \varphi_M\}. \quad (1)$$

These have usually been obtained from a previous SCF calculation or are the result of a unitary transformation of the canonical molecular orbitals. Other choices are permissible, for instance pseudonatural orbitals obtained, either by the MC SCF [5], IEPA [6, 8] APSC [7] or the iterative natural orbital CI [1, 2] method. We will assume that the basis set is orthonormalized and that it forms a basis for the irreducible representations of the molecular point group and thus consists of complete shells of one-electron functions. The set  $\{\varphi_p\}$  is divided into two subsets

$$\{\varphi_i\} = \{\varphi_1, \dots, \varphi_m\} \quad (2a)$$

and

$$\{\varphi_a\} = \{\varphi_{m+1}, \dots, \varphi_M\}, \quad (2b)$$

where  $n = 2m$  is the number of electrons in the system. The first set consists of molecular orbitals with occupation numbers close to two. These are the occupied orbitals, if the basis set has been obtained from an SCF calculation. Otherwise they are the strongly occupied natural orbitals. Such a division of the basis set

is usually quite normal. Orbitals belonging to  $\{\varphi_i\}$  will, in what follows be called occupied orbitals and those in  $\{\varphi_a\}$  "virtual" orbitals. Indices  $i, j, k$  and  $l$  will be used to label orbitals in  $\{\varphi_i\}$  and indices  $a, b, c$  and  $d$  for  $\{\varphi_a\}$ .

The one-electron functions  $\{\varphi_i\}$  are used to construct a closed-shell determinant  $\phi_0$ , called thereafter the "zeroth-order wavefunction". The determinant  $\phi_0$  need not be the dominant term in the CI expansion. It is, however, a basic determinant from which the other configurations are formed by replacement of one or two spin orbitals with orbitals from the set  $\{\varphi_a\}$ . In this way we obtain the following different types of spin symmetrized ( $S = 0$ ) basis functions for the CI expansion:

$$\phi_{i \rightarrow a} = 2^{1/2} \hat{O}[\varphi_i \bar{\varphi}_a], \quad (3a)$$

$$\phi_{ii \rightarrow aa} = [\varphi_a \bar{\varphi}_a], \quad (3b)$$

$$\phi_{ij \rightarrow aa} = 2^{1/2} \hat{O}[\varphi_i \bar{\varphi}_a \varphi_a \bar{\varphi}_j], \quad (3c)$$

$$\phi_{ii \rightarrow ab} = 2^{1/2} \hat{O}[\varphi_a \bar{\varphi}_b], \quad (3d)$$

$$\phi'_{ij \rightarrow ab} = -3^{1/2} \hat{O}[\varphi_i \varphi_a \bar{\varphi}_b \bar{\varphi}_j], \quad (3e)$$

$$\phi''_{ii \rightarrow ab} = \hat{O} \{ [\varphi_i \bar{\varphi}_a \varphi_j \bar{\varphi}_b] + [\varphi_a \bar{\varphi}_i \varphi_j \bar{\varphi}_b] \}. \quad (3f)$$

The square brackets are used to denote a normalized determinant, where doubly occupied orbitals in  $\phi_0$ ,

which are not replaced have been excluded.  $\bar{O}$  is a spin projection operator for singlet spin states

$$\hat{O} = \prod_{S \neq 0} \left( 1 - \frac{\hat{S}^2}{S(S+1)} \right). \quad (4)$$

The wavefunction for the  $n$ -electron system is expanded in  $\phi_0$  and the functions (3a)–(3f) according to,

$$\Psi = \phi_0 + \sum_{\mu} C_{\mu} \phi_{\mu}, \quad (5)$$

where  $\phi_{\mu}$  is an excited configuration.

Only single- and double-excited configurations are included in this expansion. The method used to obtain the expansion coefficients makes it difficult to include higher excitations in (5) although possible in principle. For molecular systems in a closed-shell ground state an expansion of the type given above is known to yield most of the correlation energy (probably more than 90%). Shavitt et al. have recently published results from a minimal basis MO SCF CI study of the water molecule, which included all single- to tetra-excited configurations [3]. The contributions to the correlation energy from tri- and tetra-excited configurations was found to be 0.8 and 4.4%. Meyer found, in a very accurate study of the water molecule, that the unlinked cluster correction to the correlation energy is 4.5% [4].

The electrostatic hamiltonian for the  $n$ -electron system may be divided into a one- and two-electron part,

$$\hat{H} = \sum_{i=1}^n \hat{h}(i) + \sum_{i < j} 1/r_{ij}. \quad (6)$$

In order to obtain matrix elements which are easy to handle in the CI calculation, the hamiltonian (6) is partitioned according to the following procedure.

First, define a Hartree–Fock type operator

$$\hat{F}(i) = \hat{h}(i) + \sum_{j=1}^m \{ 2\hat{J}_j(i) - \hat{K}_j(i) \}, \quad (7)$$

where  $\hat{J}_j$  and  $\hat{K}_j$  are the ordinary Coulomb and exchange operators, defined in the set of strongly occupied basis orbitals  $\{\varphi_i\}$ . The operator  $\hat{F}$  will be the or-

dinary HF operator if the set  $\{\varphi_i\}$  are the HF molecular orbitals.

The total hamiltonian  $\hat{H}$  may now be rewritten as

$$\hat{H} = \hat{F} + \hat{V}, \quad (8)$$

where

$$\hat{F} = \sum_{i=1}^n \hat{f}(i) \quad (9)$$

and  $\hat{V}$  is the “fluctuation operator”,

$$\hat{V} = \sum_{i=1}^n \left\{ \sum_{j=1}^{i-1} (1/r_{ij}) - \sum_{j=1}^m [2\hat{J}_j(i) - \hat{K}_j(i)] \right\}. \quad (10)$$

Application of the linear variational method to the trial function (5) leads to the usual matrix equation for the expansion coefficients  $C$ ,

$$(\mathbf{H} - E\mathbf{1})C = 0. \quad (11)$$

The main problem in all CI calculations is how to obtain the lowest eigenvalues and the corresponding eigenvectors from this equation. The upper limit to the number of configurations in (5) is set by the possibility to construct and save the hamiltonian matrix in computer auxiliary memory and its subsequent diagonalization. The time and space necessary for this will increase as  $N_{\text{conf}}^2$ , where  $N_{\text{conf}}$  is the dimension of the CI vector. The possible number of configurations, which can actually be used in a CI calculation is therefore heavily restricted both by the computer time and available space in auxiliary memory devices. According to published results the limit thus placed on the number of configurations is about 2000 at least for cases where most of the matrix elements are different from zero. Also calculations with more than 1000 configurations tend to be very time consuming. In the present paper we shall avoid all these problems by presenting a method for extracting the eigenvalues and eigenvectors directly from (11). The basic feature of this method is that the results are obtained directly from a list of molecular two-electron integrals, without any need to construct the  $\mathbf{H}$  matrix explicitly.

Several methods for diagonalizing (11) are available. The most commonly used method is probably the iterative scheme suggested by Nesbet [9]. The benefits and drawbacks of the different diagonalization methods, especially with respect to the convergence

problem, will be discussed in a later paper [11]. In the present paper we use the Rayleigh-Schrödinger perturbation treatment: the choice is, however, arbitrary since our method can be applied to any iterative diagonalization scheme.

In the perturbation treatment  $\phi_0$  is the zeroth-order function, while the higher-order contributions  $\Psi_n$  are expanded in the excited configurations

$$\Psi_n = \sum_{\mu \neq 0} C_{\mu}^{(n)} \phi_{\mu}. \quad (12)$$

The zeroth-order hamiltonian will contain the diagonal part of the operator  $\hat{F}$ . Using Dirac's notation we can write it as

$$\hat{H}_0 = \sum_{\mu} |\phi_{\mu}\rangle \langle \phi_{\mu}| \hat{F} |\phi_{\mu}\rangle \langle \phi_{\mu}|. \quad (13)$$

The configurations  $\phi_{\mu}$  will be eigenfunctions of this operator with the corresponding eigenvalues

$$E_{\mu,0} = \langle \phi_{\mu} | \hat{F} | \phi_{\mu} \rangle. \quad (14)$$

The perturbation  $\hat{H}_1$  is according to (4) given by

$$\hat{H}_1 = \left\{ \hat{F} - \sum_{\mu} E_{\mu,0} |\phi_{\mu}\rangle \langle \phi_{\mu}| \right\} + \hat{V}. \quad (15)$$

Using the usual perturbation expansions

$$\Psi = \phi_0 + \sum_{n=1}^{\infty} \lambda^n \Psi_n, \quad (16a)$$

$$E = E_0 + \sum_{n=1}^{\infty} \lambda^n E_n, \quad (16b)$$

we obtain the following equation for the  $n$ th order contribution to the wavefunction,

$$(\hat{H}_0 - E_0) \Psi_n = \sum_{k=1}^{n-2} E_{n-k} \Psi_k - (\hat{H}_1 - E_1) \Psi_{n-1}. \quad (17)$$

Inserting expansion (12) and the perturbation operator (15) we get an equation for the expansion coefficients of the  $n$ th order,

$$(E_{\mu,0} - E_{0,0}) C_{\mu}^{(n)} = \sum_{k=1}^{n-2} E_{n-k} C_{\mu}^{(k)} - \sum_{\nu \neq \mu} F_{\mu\nu} C_{\nu}^{(n-1)} - \sum_{\nu} (V_{\mu\nu} - E_1 \delta_{\mu\nu}) C_{\nu}^{(n-1)}, \quad (18)$$

where  $F_{\mu\nu}$  and  $V_{\mu\nu}$  are matrix elements of the operators  $\hat{F}$  and  $\hat{V}$ , respectively. The contribution to the correlation energy pertaining to the  $(n+1)$ th order is then

$$E_{n+1} = \sum_{\mu} (F_{0\mu} + V_{0\mu}) C_{\mu}^{(n)}. \quad (19)$$

The energy up to the order  $2n$  can easily be calculated from the  $n$ th-order wavefunction. Explicit formulae for this can be found in ref. [10].

The two first terms in (18) are trivial and can easily be evaluated since the matrix elements of  $\hat{F}$  are proportional to the matrix elements of the one-electron operator  $\hat{f}$ . These can be kept in the computer random access memory. What remains is then the last term. A term of this type appears in all types of iterative diagonalization procedures and our way of handling it is therefore independent of the particular diagonalization method used here. It was noted that the matrix elements  $(V_{\mu\nu} - E_1 \delta_{\mu\nu})$  consist of simple linear combinations of two-electron integrals

$$(pq|rs) = \int \varphi_p^*(1) \varphi_q(1) (1/r_{12}) \varphi_r^*(2) \varphi_s(2) dV_1 dV_2. \quad (20)$$

It is therefore possible to rewrite the last sum in (18) as a sum over such two-electron integrals. Introduce the notation

$$\sigma_{\mu}^{(n)} = \sum_{\nu} (V_{\mu\nu} - E_1 \delta_{\mu\nu}) C_{\nu}^{(n-1)}. \quad (21)$$

The contribution to  $\sigma_{\mu}^{(n)}$  from the interaction between doubly excited configurations ( $\mu = ij \rightarrow ab$  and  $\nu = kl \rightarrow cd$ ) is given by

$$\begin{aligned}
\Delta\sigma_{ij \rightarrow ab}^{(n)} = & \sum_{k,l} \{A_{(\mu)(\nu)}(ik|jl) + A'_{(\mu)(\nu)}(il|kj)\} C_{kl \rightarrow ab}^{(n-1)} \\
& + \sum_{c,d} \{A_{(\mu)(\nu)}(ac|bd) + A'_{(\mu)(\nu)}(ad|bc)\} C_{ij \rightarrow cd}^{(n-1)} \\
& + \sum_{k,c} \{B_{(\mu)(\nu)}(ai|ck) + B'_{(\mu)(\nu)}(ac|ki)\} C_{kj \rightarrow cb}^{(n-1)} \\
& + \sum_{k,c} \{B_{(\mu)(\nu)}(bj|ck) + B'_{(\mu)(\nu)}(bc|kj)\} C_{ik \rightarrow ac}^{(n-1)} \\
& + \sum_{k,c} \{C_{(\mu)(\nu)}(bi|ck) + C'_{(\mu)(\nu)}(bc|ki)\} C_{kf \rightarrow ac}^{(n-1)} \\
& + \sum_{k,c} \{C_{(\mu)(\nu)}(aj|ck) + C'_{(\mu)(\nu)}(ac|kj)\} C_{ik \rightarrow cb}^{(n-1)}. \quad (22)
\end{aligned}$$

From the interaction between single- and double-excited configurations ( $\mu = ij \rightarrow ab$  and  $\nu = k \rightarrow c$ ) the following contribution to  $\sigma_{ij \rightarrow ab}^{(n)}$  is obtained

$$\begin{aligned}
\Delta\sigma_{ij \rightarrow ab} = & \sum_c \{D_{(\mu)}(ac|bj) + D'_{(\mu)}(aj|bc)\} C_{i \rightarrow c}^{(n-1)} \\
& + \sum_c \{D_{(\mu)}(ai|bc) + D'_{(\mu)}(ac|bi)\} C_{j \rightarrow c}^{(n-1)} \\
& - \sum_k \{D_{(\mu)}(ai|kj) + D'_{(\mu)}(aj|ki)\} C_{k \rightarrow b}^{(n-1)} \\
& - \sum_k \{D_{(\mu)}(bj|ki) + D'_{(\mu)}(bi|kj)\} C_{k \rightarrow a}^{(n-1)}. \quad (23)
\end{aligned}$$

A similar expression holds for  $\Delta\sigma_{i \rightarrow a}^{(n)}$ .

The interaction between singly excited states yields the following contributions

$$\Delta\sigma_{i \rightarrow a}^{(n)} = \sum_{j,b} \{2(ai|bj) - (ab|ij)\} C_{j \rightarrow b}^{(n-1)}. \quad (24)$$

In the above expressions indices  $i, j, k$  and  $l$  have been used to label occupied orbitals in  $\phi_0$ , while  $a, b, c$  and  $d$  label the virtual orbitals. The numerical constants

$A_{(\mu)(\nu)}$  to  $D_{(\mu)}$  will depend upon the types of configurations  $\mu$  and  $\nu$  [according to the definitions (3a)–(3f)] and upon the permutation necessary in order to

put equal orbitals in  $\phi_\mu$  and  $\phi_\nu$  into equivalent positions. A list of these coupling constants can be found in table 1.

The sum over configurations in the general expression (21) has been replaced by simple sums over two-electron integrals by means of expressions (22)–(24). The computation of the CI wavefunction has thus been made much more efficient; namely, the tedious construction of the interaction matrix has been avoided and further the number of multiplications necessary to obtain one set of  $\sigma_\mu$ 's is reduced to a minimum. The number of terms in (21) is of the order of  $N_{\text{conf}} \propto m^2(M-m)^2$ , while the largest sum in (22) contains at most  $\approx (M-m)^2/2$  terms. The computation of the  $\sigma_\mu$ 's is in practice done during one reading of the two-electron integral list. Each integral is identified with respect to its type as defined in (22)–(24). It is then multiplied with the proper coupling constant and a CI coefficient. Finally the product is added to the appropriate  $\sigma_\mu$ . The procedure is quite fast, the time required is comparable to time needed for an SCF iteration. An analysis of the number of terms in the different parts of  $\sigma_\mu$  shows that the computer time needed for one iteration should be proportional to

$$m^2(M-m)^2 M^2 \propto N_{\text{conf}}(N_{\text{int}})^{1/2}, \quad (25)$$

where  $N_{\text{int}}$  is the number of non-zero two-electron integrals, that is, larger than a given threshold. The latter expression should be approximately valid even when  $N_{\text{conf}}$  and  $N_{\text{int}}$  are reduced by symmetry. The computer time necessary for a direct computation of  $\sigma_\mu$  from (21) is, on the other hand, proportional to  $N_{\text{conf}}^2$ .

The above considerations hold for the case when all singly and doubly excited configurations are included in the CI expansion. Since the total computer time depends only weakly on the number of configurations, there is little need to reduce these, except perhaps, when limitations in random access memory have to be taken into consideration. The two vectors  $C^{(n)}$  and  $C^{(n-1)}$  and an index vector have to be kept simultaneously in core.

If the available core storage is around 100 k words, the maximum number of configurations will be around 40000.

A computer program – CICLOS – based on the procedure presented above has been written by the present author. CICLOS will be included as the CI

Table 1  
Coupling constants for excited configurations. The indices  $\kappa$  and  $\lambda$  ( $= 1, \dots, 5$ ) refer to the different types of double-excited configurations as defined by the formulae (3b)–(3f)

$A_{\kappa\lambda}$					$A'_{\kappa\lambda}$				
1	2	3	4	5	1	2	3	4	5
1 1					0				
2 0	1				0	0			
3 $2^{1/2}$	0	1			0	0	1		
4 0	$(3/2)^{1/2}$	0	1		0	0	0	$1/2$	
5 0	$2^{-1/2}$	0	0	1	0	0	0	$3^{1/2}/2$	$-1/2$

  

$B_{\kappa\lambda}^a$					$B'_{\kappa\lambda}$				
1	2	3	4	5	1	2	3	4	5
1 $1/2$					$-1$				
2 $2^{-1/2}$	$1/2$				$-2^{1/2}$	$-1$			
3 $2^{-1/2}$	$1/2$	$1/2$			$-2^{1/2}$	$-1$	$-1$		
4 0	0	0	0		$-3^{1/2}$	$-(3/2)^{1/2}$	$-(3/2)^{1/2}$	$-\frac{1}{4}(p-3)$	
5 2	$2^{1/2}$	$2^{1/2}$	0	$\frac{1}{2}(1-3p)$	$-1$	$-2^{-1/2}$	$-2^{-1/2}$	$-(3^{1/2}/4)(p+1)$	$\frac{1}{4}(3p-1)$

  

$C_{\kappa\lambda}$					$C'_{\kappa\lambda}$				
1	2	3	4	5	1	2	3	4	5
1 $1/2$					$-1$				
2 $2^{-1/2}$	$1/2$				$-2^{1/2}$	$-1$			
3 $2^{-1/2}$	$1/2$	$1/2$			$-2^{1/2}$	$-1$	$-1$		
4 $3^{1/2}$	$(3/2)^{1/2}$	$(3/2)^{1/2}$	$\frac{3}{4}(1+p)$		$-3^{1/2}$	$-(3/2)^{1/2}$	$-(3/2)^{1/2}$	$-\frac{1}{4}(3+p)$	
5 $-1$	$-2^{-1/2}$	$-2^{-1/2}$	$(3^{1/2}/4)(1-3p)$	$\frac{1}{2}(3p-1)$	$-1$	$-2^{-1/2}$	$-2^{-1/2}$	$(3^{1/2}/4)(p-1)$	$\frac{1}{4}(3p+1)$

  

$D_{\kappa}$		$D'_{\kappa}$	
1	2	1	2
1 $2^{1/2}/4$		$2^{1/2}/4$	
2 $1/2$		$1/2$	
3 $1/2$		$1/2$	
4 0		$-(3/2)^{1/2}$	
5 $2^{1/2}$		$-2^{-1/2}$	

a)  $p = (-1)^q$ , where  $q$  is the parity of the permutation which brings equal indices into equivalent positions (example: if  $c = a$  in the last term of (22) a transposition is necessary in order to obtain the excitation  $\nu = ik \rightarrow ab$ , in which  $i, a$  and  $b$  take the same positions as in  $\mu = ij \rightarrow ab$ . In this case  $q$  is odd and  $p = -1$ ).

part in the program system MOLECULE for MO LCAO SCF CI calculations, presently being developed at this institute [17].

## 2. Application to $\text{NH}_3$ and $\text{H}_2\text{O}$

In this section we present the results of two cal-

culations, one for ammonia and one for the water molecule. These calculations should, however, be looked upon as test samples rather than as an attempt at obtaining accurate wavefunctions.

Two calculations for  $\text{NH}_3$  have been made. Both used a basis set of 58 primitive gaussian functions (10s, 6p and 1d for nitrogen; 5s and 1p for hydrogen), contracted to (N:4s, 2p, 1d/H:2s, 1p) in calculation I.

Table 2

Timing data for CI calculations with the computer program  
CICLOS (central processing time on an IBM 360/75)

	NH <sub>3</sub> I <sup>a)</sup>	NH <sub>3</sub> II	H <sub>2</sub> O
number of conf.	2473	4335	2459
number of two-el. int.	42436	113247	65797
$T_{\text{transf.}}$ b)	16 min	50 min	43 min
$T_{\text{CI}}$ (per iteration) c)	0.9 min	2.0 min	1.2 min
$T_{\text{total}}$	25 min	71 min	57 min
number of iterations	10	10	12

a) I refers to a calculation with 31 atomic basis functions and 28 MO's, while II has 38 basis functions and 36 MO's.

b) CPU time for the transformation of the two-electron integrals.

c) CPU time for one iteration in the actual CI calculation.

and to (N:5s, 3p, 1d/H:3s, 1p) in II.

The basis set used for water consists of 60 primitive gaussians (1 1s, 7p and 1d for oxygen; 5s and 2p for hydrogen). This set was contracted to (O:5s, 3p, 1d/H:3s, 2p). All calculations were carried out at experimental geometries.

All single- and double-excited configurations have been included (excluding, however, in all cases excitations from the 1s-orbitals), which for NH<sub>3</sub> (I), NH<sub>3</sub> (II) and H<sub>2</sub>O means 2473, 4335 and 2459 configurations. Canonical molecular orbitals were used as the basic one-electron functions. Computation times for the CI part of the calculations are presented in table 2. As evident only a minor part of the total time is used for obtaining the CI vectors from the list of two-electron integrals. The major part of the central processing time is used for the transformation of the two-electron integrals from an atomic to a molecular orbital basis. Since the time when this computation was done, much faster routines have been developed and will be implemented in the next version of the program [12, 13]. Nevertheless, the upper limit to the atomic basis set size is determined in this part of the calculation and can at present be estimated to be around 100. The correlation energy had in all three cases converged to  $10^{-6}$  after 10 iterations.

Energy results for NH<sub>3</sub> (II) and H<sub>2</sub>O are presented in table 3. Kari and Csizmadia have published results from a CI study of this molecule which included 918 configurations [14]. They obtained a total energy of -56.3747 au and a correlation energy of

Table 3

Total and pair energies for H<sub>2</sub>O and NH<sub>3</sub> (in atomic units)

	NH <sub>3</sub> (II)		H <sub>2</sub> O <sup>a)</sup>	
$E_{\text{SCF}}$	-56.2122		-76.0518	(76.0628)
$E_{\text{corr}}$	-0.2033		-0.2099	(0.2497)
$E_{\text{tot}}$	-56.4155		-76.2617	(76.3683)
pair energies:				
$2a_1-2a_1$	-0.0101	$2a_1-2a_1$	-0.0104	(0.0114)
$3a_1-2a_1$	-0.0164	$3a_1-2a_1$	-0.0173	(0.0219)
$3a_1-3a_1$	-0.0194	$3a_1-3a_1$	-0.0191	(0.0223)
$1e_x-2a_1$	-0.0199	$1b_1-2a_1$	-0.0161	(0.0235)
$1e_x-3a_1$	-0.0258	$1b_1-3a_1$	-0.0280	(0.0353)
$1e_x-1e_x$	-0.0203	$1b_1-1b_1$	-0.0173	(0.0223)
$1e_y-2a_1$	-0.0199	$1b_2-2a_1$	-0.0210	(0.0249)
$1e_y-3a_1$	-0.0258	$1b_2-3a_1$	-0.0318	(0.0340)
$1e_y-1e_x$	-0.0251	$1b_2-1b_1$	-0.0269	(0.0317)
$1e_y-1e_y$	-0.0203	$1b_2-1b_2$	-0.0218	(0.0224)

a) Values given within parentheses refer to the very accurate calculation of Meyer [4].

-0.1631 au [15]; this should be compared to the present values -56.4155 and -0.2033 au. (The computer time was reported to be 6 hours on an 7094-II.) The non-relativistic limit of the total energy of ammonia has been estimated to -56.554 au and the corresponding correlation energy to be -0.329 au [15]. Of this approximately 0.05 au can be taken as inner shell correlation. Thus the valence shell should contribute around 0.28 au of the correlation energy. The present calculation gives 73% of this value.

Pair energies are also given in table 3. They have been obtained from the following division of the total correlation energy,

$$E_{\text{corr}} = \sum_{i < j} \left\{ \sum_{a < b} H_{0,ij \rightarrow ab} C_{ij \rightarrow ab} \right\} = \sum_{i < j} E_{ij} \quad (26)$$

Even if this formula does not have a direct physical significance it is nevertheless useful for the interpretation of the results. A study of the pair energies for NH<sub>3</sub> indicates that the electron pair in the most stable orbital  $2a_1$  is less well correlated compared to the other interpair correlations. This orbital has a large N2s character. Probably the basis set is not sufficiently flexible to give enough angular correlation in

Table 4

Occupation numbers for natural orbitals in  $\text{NH}_3$  and  $\text{H}_2\text{O}$ . Only orbitals with occupation numbers larger than 0.01 are given

$\text{NH}_3$ (II)		$\text{H}_2\text{O}$	
$2a_1$	1.9815	$2a_1$	1.9852
$3a_1$	1.9685	$3a_1$	1.9684
$4a_1$	0.0190	$4a_1$	0.0231
$5a_1$	0.0144	$5a_1$	0.0112
$1e_x$	1.9678	$1b_1$	1.9735
$2e_x$	0.0210	$2b_1$	0.0173
$1e_y$	1.9678	$1b_2$	1.9671
$2e_y$	0.0210	$2b_2$	0.0245

this region. More basis functions of p and d type with relatively high exponents should have been needed for this purpose. This deficiency of the basis set is also manifested in the natural orbital occupation numbers presented in table 4. The orbital  $2a_1$  has the largest occupation number of all the strongly occupied orbitals.

A number of CI calculations have been reported for the water molecule [3, 4, 16]. The most accurate wavefunction is presented by Meyer [4], using as a basis of one-electron functions a set of non-orthogonal pseudonatural orbitals which was obtained by an independent pair approach. He reports a total energy of  $-76.3683$  au as compared to our computed value of  $-76.2617$ . The non-relativistic equilibrium energy has been estimated by Meyer to be  $-76.4314$  au and the correlation energy to  $-0.369$  au. He obtains a correlation energy of  $-0.2497$  au for the valence shell electron pairs while the present study gives  $-0.2099$  au. The difference is probably due mainly to the different choice of atomic basis functions. The basis set used in ref. [4] is very extensive, consisting of 11s, 7p, 4d and 1f functions for oxygen, 5s and 1p functions for hydrogen. The improvement is reflected rather evenly in all pair energies (cf., table 3), and is probably due to a better description of the angular correlation effects, which was achieved through the use of many basis functions of d- and f-types. It is, however, gratifying to see that such a large fraction of the valence shell correlation energy ( $\approx 75\%$ ) can be obtained even with a basis set of a rather moderate size, a set which can also be used for studies of larger

systems and for the calculation of potential energy surfaces. The situation can be expected to improve further when optimal orbital exponents for basis functions describing angular correlation effects become available.

Work along these lines is presently being done at this Institute.

### Acknowledgements

The author wishes to thank Drs. J. Almlöf and P. Siegbahn for many valuable discussions. The testing and debugging of the program was partly made in Strasbourg, on a grant from Ministère des Affaires Etrangères de France. The hospitality and support of Professor Alain Veillard at Institut de Chimie in Strasbourg is gratefully acknowledged. Part of this work has been sponsored by a grant from the Swedish Natural Science Research Council and the Swedish Board for Technical Development.

### References

- [1] C.F. Bender and E.R. Davidson, *J. Phys. Chem.* 70 (1966) 2675.
- [2] R.J. Buenker and S.D. Peyerimhoff, *Theoret. Chim. Acta* 12 (1968) 183.
- [3] R.P. Hosteny, R.R. Gilman, T.H. Dunning Jr., A. Pipano and I. Shavitt, *Chem. Phys. Letters* 7 (1970) 325.
- [4] W. Meyer, *Intern. J. Quantum Chem.* 5 (1971) 341.
- [5] A. Veillard and E. Clementi, *Theoret. Chim. Acta* 7 (1967) 133.
- [6] R.K. Nesbet, *Phys. Rev.* 155 (1967) 51.
- [7] K.J. Miller and K. Ruedenberg, *J. Chem. Phys.* 48 (1968) 3418.
- [8] W. Kutzelnigg, *Theoret. Chim. Acta* 1 (1963) 327, 343.
- [9] R.K. Nesbet, *J. Chem. Phys.* 43 (1965) 311.
- [10] P.-O. Löwdin, *J. Math. Phys.* 6 (1965) 1341.
- [11] B. Roos and P. Siegbahn, to be published.
- [12] G. Dierksen, private communication.
- [13] J. Almlöf and P. Siegbahn, private communication.
- [14] R.E. Kari and I.G. Csizmadia, *Theoret. Chim. Acta* 22 (1971) 1.
- [15] M.A. Robb and I.G. Csizmadia, *Intern. J. Quantum Chem.* 4 (1970) 365.
- [16] H.F. Schaefer III and C.F. Bender, *J. Chem. Phys.* 55 (1971) 1720.
- [17] J. Almlöf, B. Roos and P. Siegbahn, to be published.