

## Thermodynamic quantities from quantum and classical mechanics

Enrico Clementi

IBM Corporation, Center for Scientific and Engineering Computations,  
Department 48B/ MS 428, Kingston, New York 12401, USA

**ABSTRACT** — The theoretical and computational description of a chemical system requires the availability of a variety of techniques, each one optimally adapted to address a very specific level within the very wide spectrum of chemical complexity. We elect to comment on very accurate methods (Hylleraas CI) for up to three electrons in molecules and for accurate methods (CI) for up to 10 electrons in atoms: Next we cross over to large molecules, where the main interest is in the structure and dynamics. Here we use molecular dynamics, either using at each time step many-body *ab initio* derived interaction potentials (for liquid water) or density functionals and quantum-mechanically obtained energies and forces (for a cluster of 60 carbon atoms). In each of these four topics, the data we report is, in part, unpublished research, which, however, is presented in sufficient detail as to give the reader all the needed information to evaluate our conclusions and suggestions in forecasting aspects of computational chemistry eight to ten years from now.

### INTRODUCTION

There is a truly large range of thermodynamic properties one can obtain from laboratory data, covering a broad spectrum of techniques developed in more than one century. However, more and more one can notice an increasing contribution of thermodynamic data derived from computational methods, the latter being developed approximately in the last one to two decades, namely since computations and computer programs become, the former sufficiently fast, the latter sufficiently available to impact the chemical community.

By and large there are two main avenues of computational methods in chemistry yielding thermodynamic data, precisely computational quantum chemistry and computational statistical mechanics. The former includes ground state and excited state properties and corresponding spectra. The basic equation is the Schrödinger equation in its many approximations. The main statistical mechanics approaches are either time independent (Monte Carlo) or time dependent (Molecular Dynamics). The basic equations are either Newton equations or Langevin equations; as it is known, the latter can be considered as modified from the former when one considers an average over a long time, and resolution constrained to a well-defined model.

There are new approaches in today's computational chemistry, taking their place near the traditional ones. Indeed, the traditional techniques for stationary state quantum chemistry and equilibrium molecular dynamics are being expanded with Quantum Monte Carlo,<sup>1</sup> Quantum Molecular Dynamics,<sup>2</sup> Microdynamics<sup>3</sup> but also by research and computer programs on Data Base, Interactive Animation, Artificial Intelligence and Chemical Knowledge Processing.<sup>4</sup>

This very vigorous expansion should, however, not be taken as indication that the problems characteristic in the traditional areas have been overcome. Indeed, it is somewhat nearer to the truth the observation that new areas are perhaps opened to dissipate the energy of the increased manpower, which has been frustrated in attempting to solve "old" problems!

Presently, we are at the end of the IVth computer generation, where we have witnessed not only increased MIPS performance, but especially MFLOPS performance, either because of vector or parallel architectures. From personal computers, to workstations, to mainframes and supercomputers, the advances have been on a very broad front.

The next computer generation is the VIst and, only more slowly, the much heralded Vth generation. In the VIst generation we have already announcements of supercomputers with peak performance of about

20-30 GFLOPS; thus it is expected we shall grow between 100-500 GFLOPS, partly by decreasing the clock speed, mainly because of multi-processing. We already see workstations with the same speed as the old CRAY XMP (single processor) and the disk storage is being addressed by channels working in parallel, with optical fibers and one or more miles in interconnected distances. Slowly, also the Vth computer generation will appear, since artificial intelligence and expert systems are becoming ubiquitous, and this will add pressure to the hardware manufacturers.

In the conclusions of this review paper we shall consider today's computational techniques but assuming availability of systems with up to 500 GFLOPS and capable of retrieving from disk storage hundreds of Giga words at 100-500 Mbytes/sec transmission. These predictions are "reasonable" in a technical viewpoint, and also "reachable to relatively few" in an economical context. Clearly, the main computing will trail behind these limits, but everybody will experience much improvement.

We shall start by considering one of the first techniques introduced in quantum mechanics<sup>6</sup> and later adapted within the configuration interaction (CI) framework: Hylleraas CI. The examples will deal with the simplest chemical systems but at the highest accuracy level.

Next, we shall move away from the goal of 10 cm<sup>-1</sup> (*i.e.* 0.00005 a.u.) accuracy in absolute energy and move to the milli-hartree accuracy, again using CI, but with standard expansions of determinantal functions. The examples analyzed are iso-electronic series for He (1S), Li (2S), Be (1S), and Ne (1S).

We shall not consider the standard and popular methods like Møller Plesset perturbations, already too well known. There is a tendency in today's literature to overclaim the accuracy one has obtained using MPx. Indeed, we note that most authors are comparing, for example, binding energies in a molecule with atomization products, which are far from the limit of the adopted approximation. For example, often a calculation proceeds to post-Hartree-Fock corrections before reaching a near Hartree-Fock value (which is seldom obtained in today's literature). In addition, basis set superposition (BSS) corrections are often neglected and basis sets with nearly zero BSS error are essentially unknown. In this way *ab initio* techniques are used "semi empirically" with calibrated basis sets, which we are becoming accustomed to seeing discarded, because of their unreliability, every four to five years! A different situation occurs in the use of MBPT, where the methodological rigor is clearly superior.<sup>5</sup>

Indeed, we shall be interested in sufficiently large systems where even a MP2 computation would become computationally much too expensive. This is the area of *ab initio* molecular dynamics and quantum molecular dynamics. We shall consider two systems, one composed of 1000 water molecules in a periodic system, *i.e.* liquid water, and the second with 60 carbon atoms in the C<sub>60</sub> cluster. In the conclusions we shall consider other systems of higher complexity, and for which work is in progress in our laboratory.

The four examples we have mentioned have been the subject of recent and mostly unpublished studies from our department, thus we are in the ideal position to report and discuss the corresponding computer time using the same hardware and system software. This will simplify our extrapolations and yield a reasonably reliable forecast for about ten years from now, namely at about the end of the VIst computer generation.

## VERY FEW ELECTRON SYSTEMS

The traditional approach to quantum mechanical descriptions of many-electron systems in atoms, molecules, and solids is to optimize linear combinations of one-electron functions, namely functions of one electron only (with space and spin components). This type of approach leads then to describing correlation in an *implicit* manner, and hence these techniques converge rather slowly to the true solution of the many-electron time-independent Schroedinger equation.

Instead of the implicit approach, one can argue that the most powerful approach should be one where inter-electronic coordinates are built *explicitly* into the wavefunction. Hylleraas was the first to develop this approach and he used it to calculate the energy and wavefunction of the helium atom with great success.<sup>6</sup> James and Coolidge<sup>7</sup> extended this method to the hydrogen molecule and Kolos and Wolniewicz<sup>8</sup> have shown that, to date, it is the most efficient method available for calculating accurate potential energy curves in the hydrogen molecule. The results of the latter authors even challenged the experimental spectroscopic data. The verification of their prediction was hailed as "a great triumph of *ab initio* calculations" by R. S. Mulliken.<sup>9</sup> However, due to the numerical complications introduced by the inclusion of inter-electronic coordinates in the wavefunctions, Hylleraas' approach was limited to simple cases such as the helium atom and the hydrogen molecule.

With the advent of supercomputers, a more general extension of the principle of explicit inclusion of the inter-electronic separation into the wavefunction is possible. This consists of multiplying standard many-electron wavefunctions by powers of the inter-electronic distance  $r_{ij}$ . This approach is completely general, it can be applied to many-center, multi-electron systems of any type. The only limitation, as is

generally the case in computational chemistry, is imposed by computational resources; later, in section 5, we shall return on this point.

The standard CI approach consists of expanding the many-electron wavefunction  $\Psi(1,2, \dots, N)$  in a linear combination of configuration state functions (CSFs)  $\Phi_k$  such that

$$\Psi_{\text{CI}}(1,2, \dots, N) = \sum_{k=1}^{\text{NCSF}} C_k \Phi_k(1,2, \dots, N), \quad (1)$$

where  $\Phi_k$  is a linear combination of totally anti-symmetric determinantal states each consisting of  $N$  independent molecular orbital functions  $\varphi_i$  and NCSF is the total number of CSFs included. The HCI approach modifies Eq. (1) by expanding the wavefunction in powers of the inter-electronic distances  $r_{ij}$  weighted by the original CSFs  $\Phi_k$ , *i.e.*

$$\Psi_{\text{HCI}}(1,2, \dots, N) = \sum_{\nu} \sum_{i < j=1}^N \sum_{k=1}^{\text{NCSF}} C_{\nu, k} r_{ij}^{\nu} \Phi_k(1,2, \dots, N). \quad (2)$$

The effect of the  $r_{ij}$  terms is to explicitly introduce correlation into the wavefunction. This expansion in the inter-electronic distance is generally assumed to be a power series expansion,<sup>10</sup> thereby limiting  $\nu$  to non-negative integers. It has been shown by others<sup>11</sup> that of the possible non-zero  $\nu$  values,  $\nu = 1$  is the most important and our discussion will be limited to the  $\nu = 0$  (normal CI) and  $\nu = 1$  (what we term HCI) terms. It has also been shown repeatedly that inclusion of the HCI terms does speed the convergence of configuration expansions and that the HCI method is a viable alternative when high-accuracy calculations are needed.

If we now assume that the molecular CSFs are constructed of standard cartesian Gaussians, we can use straightforward techniques useful for performing computations.<sup>12-13</sup> This has been done and a computer package has been created that performs HCI calculations for many-center, two- and three-electron molecular systems.<sup>12-14</sup> Expansion to the most general four-electron integrals has not yet been done. This package is called HYCOIN (*Hylleraas Configuration Interaction*)<sup>12</sup> and it has been used successfully to calculate a number of molecular states.<sup>14-20</sup>

The natural place to begin our examination of correlation in small molecular systems is the ground-state of  $\text{H}_2$ , two electrons and two protons. Table 1 summarizes the effect of basis set size for HCI calculations on this state.<sup>15,17,21</sup> The "exact" variational limit for this state was established by Kolos *et al.*,<sup>22</sup> who used a specialized elliptical coordinate basis set with an explicitly correlated wavefunction, at -1.1744757 hartrees. The largest basis set of our calculation, (15s7p2d1f), produces an HCI energy that is less than 0.3  $\text{cm}^{-1}$  above this variational limit. If spectroscopic accuracy is defined as having an error of less than  $\sim 1 \text{ cm}^{-1}$ , Table 1 indicates that it can be achieved with a basis set of size as small as (13s7p2d).

TABLE 1. The effects of the basis size in  $\text{H}_2$  at an inter-nuclear distance of 1.4011 bohrs (total energies are given in hartrees).

Basis Set	SCF	Conventional-CI	Hylleraas-CI
(13s)	-1.128532	-1.154881	-1.168870
(13s7p)/[13s1p]	-1.133561	-1.170495	-1.174334
(13s7p)/[13s2p]	-1.133561	-1.171378	-1.174380
(13s7p1d)/[13s1p1d]	-1.133610	-1.171661	-1.174399
(13s7p1d)/[13s2p1d]	-1.133618	-1.172596	-1.174456
(13s7p1d)	-1.133618	-1.173306	-1.174467
(13s7p2d)	-1.133619	-1.173858	-1.174473
(14s7p2d1f)	-1.133622	-1.173987	-1.174474
(15s7p2d1f)	-1.133622	-1.173987	-1.174475

Compared with conventional CI, the size of the basis set seems to have much less effect on the recovering of the correlation energy. For example, addition of the first contracted p-gaussian lowers the conventional CI energy by 0.016614 hartree, but brings down the HCI energy by only 0.005464 hartree. Similarly, one d-type function added to the (13s7p)/[13s1p] basis set has an effect of 0.00116 hartree in

conventional CI, compared with 0.00065 hartree in HCI. All of these seem to be due to the fact that most of the correlation energy has already been picked up in the s-function basis set, 86% with HCI compared to 52% with conventional CI. Thus the HCI method should be much superior to conventional CI in terms of convergence to a given accuracy in energy. That is indeed the case. For example, the HCI energy obtained with (13s7p) contracted to [13s1p], -1.174334 hartree, is already much better than CI employing 5s3p3d2f Slater functions, -1.174142 hartree.<sup>23</sup> Addition of one d-function lowers our calculated energy to -1.174399 hartree, which is slightly better than a calculation employing uncontracted Gaussians with angular momenta up to g: (12s6p5d4f2g).<sup>24</sup>

The next obvious step towards more complicated molecules is to increase the number of centers from two to three while keeping only two electrons, this means the study of the non-linear  $\text{H}_3^+$  molecule, a problem that cannot be tackled by the less general elliptical coordinate methods. We have recently examined both the equilibrium energy and calculated a number of points on the potential energy surfaces of both the ground  $^1A_1'$  state and first excited  $^3\Sigma_u^+$  state of  $\text{H}_3^+$ .<sup>19, 20</sup> The lowest energy obtained by us for the equilateral triangle geometry of the ground state was -1.3438279 hartree at the internuclear distance of 1.6500 bohrs with the use of a 13s5p3d basis set on each site (a total of 138 basis functions). This energy is compared in Table 2 with other recent *ab initio* calculations on the equilibrium energy. From Table 2 we can see that this energy is significantly lower than the previous best published variational calculation of -1.343500 hartree,<sup>15</sup> and is in excellent agreement with the results of the latest quantum Monte Carlo calculations, -1.34387 $\pm$ .00005 hartree<sup>31</sup> and -1.3433 $\pm$ .0005 hartree.<sup>32</sup> The energy is also below the soon-to-be published result of Alexander *et al.*,<sup>33</sup> who used a Random-Tempered Optimization method with Gaussian-Type Geminals (also a variational calculation with an explicitly correlated wavefunction), and obtained -1.3438220 hartree at an equilibrium separation of 1.6504 bohrs.

TABLE 2. Selected *ab initio* calculations for the ground state equilibrium energy of  $\text{H}_3^+$ .

Authors	Method	$R_=(b)$	E(h)
Salmon, <i>et al.</i> <sup>25</sup> (1973)	CI, 18 SP	1.6500	-1.34335
Mentch, <i>et al.</i> <sup>26</sup> (1981)	Random Walk	1.6500	-1.3439 $\pm$ .0002
Preiskorn, <i>et al.</i> <sup>27</sup> (1984)	SCC, 24 CGLO	1.6504	-1.343422
Burton, <i>et al.</i> <sup>28</sup> (1985)	CI, 108 PNO	1.6525	-1.34272
Meyer, <i>et al.</i> <sup>29</sup> (1986)	CI, 104 CGTO	1.6504	-1.34340
Anderson <sup>30</sup> (1987)	Random Walk	1.6500	-1.34376 $\pm$ .00003
Traynor, <i>et al.</i> <sup>31</sup> (1988)	Random Walk	1.6500	-1.34387 $\pm$ .00005
Urdaneta, <i>et al.</i> <sup>15</sup> (1988)	HCI, 48 CGTO	1.6504	-1.343500
Huang, <i>et al.</i> <sup>32</sup> (1990)	Random Walk	1.6500	-1.3433 $\pm$ .0005
Alexander, <i>et al.</i> <sup>33</sup> (1990)	RTO, 700 GTG	1.6504	-1.3438220
Frye, <i>et al.</i> <sup>19</sup> (1990)	HCI, 138 GTO	1.6499	-1.3438279

Note: SP = Singer Polynomial, CGLO = Contracted Gaussian Lobe Orbitals, PNO = Pseudo Natural Orbitals, CGTO = Contracted Gaussian Type Orbitals, RTO = Random-Tempered Optimization, GTG = Gaussian Type Geminal, GTO = Gaussian Type Orbitals.

From here, we can continue on to  $\text{H}_3$ . Now, with three electrons the complexity of the integral calculations goes up dramatically and, commensurately, so does the need for increased computational resources. In fact, at this juncture, only small basis set calculations have been performed with the HCI method. An example of this is a 3s1p basis set calculation for linear  $\text{H}_3$  with internuclear distances of 1.75 b.<sup>14</sup> The SCF energy for this configuration was -1.5849093 a.u., the CI energy was -1.6217005 a.u., and the HCI energy was -1.6366379 a.u. This result should be compared with the result of B. Liu<sup>23</sup> who obtained an energy of -1.658743 a.u. at the same saddle point with a CI calculation using a Slater-type function basis set containing orbitals up through 5f. A good computation for  $\text{H}_3$  would require the same type of basis set used above for  $\text{H}_2$ . For example, using the MELD program of Davidson<sup>34</sup> with a (15s5p2d)/[11s5p2d] Gaussian basis set, we obtained a saddle point energy of -1.658323 a.u., which is 0.000423 a.u. above the best results of B. Liu. This basis would be a more than sufficient starting point for an HCI calculation on  $\text{H}_3$  to ensure a few  $\text{cm}^{-1}$  accuracy in the total energy. We now take the liberty of extrapolating our CPU timing for this possibility next.

It should be noted that the following timings were found with a very recently implemented code that has not yet had extensive optimization, therefore the timings and, more importantly the extrapolations, should be taken as indicative rather than concrete predictions as effort to-date has concentrated on "correctness" rather than "performance". To begin, a trivial basis set of a single s-type orbital on each center takes 359 seconds on an IBM ES/3090-J/VF to calculate the integrals. A basis set of 1s1p on

each center (12 total orbitals) takes 191 hours on the same system and a basis set of 3s1p requires approximately 1000 hours. Notice that the number of integrals required to be calculated goes as  $N^6$  where  $N$  is the number of basis functions. It is intriguing, and somewhat depressing, to speculate what would be required to perform a  $H_3$  calculation that would be accurate to within  $10\text{ cm}^{-1}$  of the "exact" value, which is not yet known as it is for  $H_2$ .<sup>22</sup> If the lessons learned from  $H_2$  are valid, then an full HCl calculation with a basis set of (13s7p)/[13s1p] should provide approximately  $10\text{ cm}^{-1}$  accuracy. Scaling from 3s1p to 13s7p by  $N^6$  would call for somewhere around 3500 years on a similar system, not counting the six-index transformation which scales as  $N^7$ . This is clearly not yet approachable without fundamental improvements in either, and probably all, theory, implementation, and hardware. Assuming a significant improvement in the numerical analysis and its corresponding code, at best we could reduce the task to 10-20 years. This is not unfeasible if a more efficient numerical integration is found and if integrals smaller than a given threshold are ignored and/or approximated. Indeed the latter is regularly achieved in SCF computations where savings of a factor of  $10^3$  are well-known and documented.<sup>35</sup>

### FEW ELECTRON ATOMS

Let us now consider a more traditional approach, namely a linear combination of Slater determinantal functions, *i.e.*, the Configuration Interaction approach. We shall start with two electron systems, and then expand up to ten electrons. The systems considered are atoms and ions, specifically, the correlation energy was computed for the neutral ground state of atoms and selected ions of the 2, 3, 4 and 10 electrons isoelectronic series.<sup>36</sup>

We used ATOMCI, an atomic Configuration Interaction program based upon the powerful techniques developed in the early seventies by Sasaki,<sup>37</sup> and used by Sasaki and Yoshimine<sup>38,39</sup> to compute correlation energies and electron affinities of the first row atoms. ATOMCI has been recently revised and extensively documented.<sup>40, 41</sup>

We performed single reference single and double CI (SDCI) calculations using Gaussian Type Orbitals of geometrical basis sets (GBS).<sup>42</sup> Using GBS helps to decrease the effort in the orbital exponent optimization; we recall that in the geometrical basis set<sup>42</sup> the orbital exponents  $\alpha_j$  are related by the equality

$$\alpha_j = \alpha_1 C^{j-1} \quad (j = 1, \dots, n) \quad (3)$$

and therefore only two parameters need to be optimized for a given set of  $n$  elements, *i.e.*  $\alpha_1$  and  $C$ . The same set of exponents were used in orbitals with different angular momenta, and our basis sets included up to  $i$  functions.

TABLE 3. Electronic Correlation Energies ( $E_{\text{SDCI}} - E_{\text{HF}}$ ) and Total Energies ( $E_{\text{SDCI}}$ , parenthesis) in a.u. for some members of the 2, 3, 4 and 10 electrons isoelectronic series.

Z		2-1S	3-2S	4-1S	10-1S
2	He	-0.04196 (-2.90363718)			
3	Li	-0.04339 (-7.27980759)	-0.0450 (-7.47777253)		
4	Be	-0.04415 (-13.65544522)	-0.0472 (-14.32456489)	-0.0899 (-14.66295228)	
10	Ne	-0.04551 (-93.90661793)	-0.0509 (-102.68195335)	-0.1758 (-110.28684560)	-0.3706 (-128.91769539)
18	Ar	-0.04594 (-312.90699489)	-0.0519 (-346.49831600)	-0.2740 (-377.80945430)	-0.3925 (-506.36060249)
30	Zn	-0.04616 (-881.40719686)	-0.0523 (-982.22383604)	-0.4168 (-1079.09690384)	-0.4046 (-1552.96299231)
36	Kr	-0.04622 (-1273.65724063)	-0.0526 (-1421.58688169)	-0.4877 (-1564.74096847)	-0.4077 (-2292.30443949)
54	Xe	-0.04630 (-2882.40730818)	-0.0528 (-3225.67480824)	-0.6996 (-3561.67348114)	-0.4124 (-5374.37544393)

We experimented with several basis sets, the best set was used to compute the energies reported in Table 3. In Table 4 each basis set is characterized by the number of primitive and of contracted functions of

$s$ ,  $p$ ,  $d$  through  $h$  and  $i$  type (in all the sets we have used the same number of  $d$ ,  $f$ ,  $g$  and  $h$  functions). The values of  $\alpha_1$  and  $C$  are also given, together with  $j_{\min}$ , the minimum value for the index  $j$ , again for the  $s$ ,  $p$ ,  $d$  through  $h$  and  $i$  symmetry types.

The correlation energies are reported, together with the total SDCI Energies, in Table 3 for selected members of the He ( $1s^2, 1S$ ), Li ( $1s^2 2s^1, 2S$ ), Be ( $1s^2 2s^2, 1S$ ) and Ne ( $1s^2 2s^2 2p^6, 1S$ ) series, up to the Xe ions. Figure 1 summarizes the results for the correlation energy. Due to the limitations of the single reference SDCI approximation, the quality of the results decreases as the number of electrons increases.

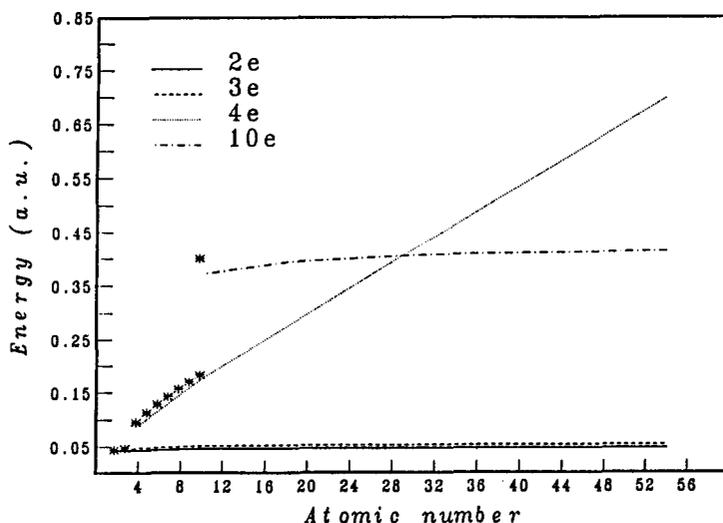


Figure 1. Correlation energies (in a.u.) for the 2, 3, 4 and 10-electrons isoelectronic series. The lines fit the data of Table 4. The stars are semiempirical estimates taken from Refs. 43 and 44.

Thus in Table 3 we report one more figure than physically meaningful (to allow for roundoff) for the correlation energies, but maintain eight decimal figures, namely the numerical accuracy corresponding to the used basis sets, for the total energies. The extension of this study to Multi Reference CI with ATOMCI is currently under way.

TABLE 4. Li( $2S$ ) ground state. Contributions from each symmetry to the total and correlation energies and comparison with the results of ref 50.

Basis	Tot. Energy(a)	Tot. Energy (b)	Cor. Energy(b)
$s$	-7.448667	-7.4486658	-0.0159390
$sp$	-7.473834	-7.4738004	-0.0410736
$spd$	-7.476760	-7.4766756	-0.0439488
$spdf$	-7.477399	-7.4773817	-0.0446550
$spdfg$	-7.477485	-7.4776244	-0.0448976
$spdfgh$		-7.4777252	-0.0449984
$spdfghi$		-7.4777725	-0.0450458

(a): Limits reported for the MCHF calculation of Ref. 50.

(b): This work.

In Fig. 1 the results are interpolated to show the overall behavior of the correlation energy for each isoelectronic series. The values estimated by Clementi<sup>43,44</sup> for some members of each series are also reported for comparison. Clementi's estimates are practically indistinguishable from our calculated values for the 2 and 3-electron series on the scale of the figure, and thus only two reference points (for the He and Li neutral atoms) were drawn. The estimates<sup>43</sup> for  $Z$  greater than 10 were reported as 10% accurate up to  $Z$  equal to 20, and as order of magnitude of the correlation energy for  $Z$  beyond 20. Notice however that the present results agree very well with the estimates of Ref. 43 except for the 10-electron isoelectronic series, especially for high  $Z$ , where they differ by as much as 30% (-0.4077 vs -0.5643 for Kr + 26). Since it appears unlikely that our approximations might result in such a large error

in the total energies for the higher members of the series, we conclude that the correlation energy for the Ne isoelectronic series for high values of  $Z$  in Ref. 43 was overestimated.

For the 2-electron series the correlation energy is accurate to about the fifth decimal figure, but the accuracy of the non relativistic total energy is 1 part per  $10^6$  a.u. for He and goes up to 1 part per  $10^9$  for  $\text{Xe}^{+52}$ . As known, the most accurate computations on the two electron series are those by Pekeris,<sup>45</sup> <sup>46</sup> later often reproduced, He being a test case for many models.<sup>47</sup>

TABLE 5. CPU time (seconds, s, or hours, h, on an IBM 3090 V400) and storage (Mbytes) requirements for the atomic CI calculations on the neutral members of the isoelectronic series here considered, and estimates for some larger systems.

System	GTO's(a)	CSF's(b)	CPU time	Elaps. time (c)	2-el int. (d)	Virt. mem.	Disc mem.
He	126	1241	1300s	5760s	21.0	200	370
Li	107	2526	855s	4320s	11.5	125	230
Be	100	3037	1050s	7800s	8.3	100	160
Ne	100	10573	4350s	19.6h	8.3	220	410
(Ar <sup>(e)</sup> )	124	47000	21h(f)	590h (g)	19.0	700 (i)	1200(i)
(Kr <sup>(f)</sup> )	147	215000	420h(f)	23000h (g)	39.0	3700 (i)	7000(i)

(a): Number of basis functions.

(b): Number of configuration state functions.

(c): Total elapsed time in seconds (s) or hours (h).

(d): Number of computed and stored two-electron integrals (millions).

(e): Rough estimates for an hypothetical SDCl calculation involving 26  $s$ , 20  $p$ , 17  $d,f,g,h$  and 10  $i$  GTO's.

(f): Estimated by quadratic extrapolation of the CPU data. 99% and 99.9% of the time would be taken by the generation of the hamiltonian matrix elements and by its diagonalization for Ar and Kr respectively.

(g): Estimated by assuming a linear relationship between the ratio Elapsed time vs CPU time and the number of electrons in the system.

(i): Estimated from the number of two electron integrals and by a linear fitting of the data for the storage needed to compute and save on disk the energy matrix.

(j): Rough estimates for an hypothetical SDCl calculation involving 30  $s$ , 24  $p$ , 20  $d,f,g,h$  and 13  $i$  GTO's.

For the 3-electron series the accuracy remains nearly the same as for the 2-electron series, because, as long ago realized, the  $1s$ - $2s$  interpair correlation energy is small<sup>43,48</sup> and thus SDCl yields energies not much different from Full CI. As far as we know, the best results for the Li ground state energy are those by Larsson.<sup>49</sup> Notice that the total energy for the top member of the series, Li, compares very well to the very recent result of a large scale MCHF calculation by Sundholm and Olsen<sup>50</sup> involving full scale CI within up to 85 orbitals of symmetries ranging from  $s$  to  $g$ , for a total of 11514 configuration state functions in  $D_{\infty h}$  symmetry (-7.477773 a.u. vs -7.477485 a.u.). In Table 4 the contribution to the total energy and to the correlation energy from each symmetry block are reported and compared with the analogous limits for the calculation of Ref. 50. Our values appear to be very close to Sundholm and Olsen's limits, and the inclusions of several  $g$ ,  $h$  and  $i$  functions seems to be effective in further improving the final result.

For four electrons the limitations of SDCl start to become apparent and the old estimates of -0.094 a.u.<sup>48,51</sup> for Be is likely nearer to the exact one than the value of -0.0899 a.u. reported here. Bunge's very accurate calculated value for the correlation energy of Be is -0.0943 a.u.<sup>52</sup> Our result improves by including contributions from higher excitations. Indeed for one member of the series,  $\text{Ne}^{+6}$ , we repeated the calculation by adding to the singles and doubles a perturbative selection of triple and quadruple excitations from the HF reference state to the lowest two atomic natural orbitals, ANO's, of the  $s$ ,  $p$  and  $d$  symmetries obtained in the single reference SDCl calculations. The value of the correlation energy changed from -0.1758 a.u. to -0.1789 a.u.

The steep increase of the correlation energy along the isoelectronic series has been analyzed in detail elsewhere.<sup>44,53</sup> Here we observe that the effect is essentially linear in  $Z$  and the 4-electron correlation energy becomes larger than 10-electron correlation energy for  $Z$  around 28 (see Fig. 1). As we have previously stressed,<sup>44,53</sup> this energy should not be considered as the *true* correlation energy for the four electrons in the two pairs  $1s^2$  and  $2s^2$ , but rather an example of a strongly multiconfigurational electronic structure which is very poorly described by the single determinant approximation. In the limit of high values of  $Z$  the  $ns$  and  $np$  orbitals are energy degenerate. Indeed a two determinants MCSCF function was sufficient to bring about a constant value for the 4-electron series,<sup>53</sup> namely -0.052, -0.055, -0.055, -0.056, -0.058, -0.059, -0.062 for Be,  $\text{B}^{+1}$ ,  $\text{C}^{+2}$ ,  $\text{N}^{+3}$ ,  $\text{O}^{+4}$ ,  $\text{F}^{+5}$  and  $\text{Ne}^{+6}$  respectively.

The computed correlation energy for the 10-electron isoelectronic series is about 95% of the estimated total.<sup>48</sup> The previous value by Sasaki and Yoshimine<sup>38</sup> for Ne, -0.3697 a.u., compares nicely with our -0.3706 a.u. and with Veillard and Clementi's estimate of -0.389 a.u.<sup>48</sup> Adding a perturbative selection of triple and quadruple excitations from the HF reference state to the lowest two ANO's of the *s*, *p* and *d* symmetries obtained in the single reference SDCI calculations yielded a correlation energy of -0.3800 a.u. for Ne (almost 98% of Veillard and Clementi estimate) and -0.3941 a.u. for Ar<sup>+</sup><sup>8</sup>.

Let us now comment on the computation time, fast memory and auxiliary storage needed for these computations in order to extrapolate, for example, to 18- (Ar) and 36- (Kr) electron systems. We shall consider S states or even closed shell to make the estimate somewhat simpler. In Table 5 we give the pertinent informations on the CPU time, the total elapsed time and the storage requirements for the atomic CI calculations on the neutral members of the isoelectronic series here taken into account. Both the amount of memory required and the time depend strongly upon the number of the two electron integrals computed and stored and upon the number of configurations. While almost 80% of the time is consumed in computing and transforming the integrals for He, approximately 90% of the time is taken by the evaluation of the energy matrix and by the solution of the diagonalization problem for Ne. Approximately the same percentages apply to the amounts of storage needed to store the integrals and the energy expression. As the number of configurations increases, so does the dimension of the CI matrix and the time to diagonalize it. In Table 5 we report also a rough estimate of the time and storage requirements which can be extrapolated for some larger systems. The single reference SDCI calculations for argon and krypton would take approximately 25 days and 2.5 years respectively, and noticeable amounts of fast and disc memory.

### MOLECULAR CLUSTERS WITH QUANTUM MOLECULAR DYNAMICS

Let us now move to an example of a relatively large molecular system for which we wish to determine the equilibrium geometry; we shall assume no symmetry point group, since the geometry determination is one of the main goals.

The study of carbon clusters has attracted attention recently and in the past.<sup>54</sup> Both experimental and theoretical methods have been employed to explain the unusual predominance of the 60-atom cluster in the mass spectra obtained with the laser vaporization cluster beam technique.<sup>55-57</sup> Experimental evidence has shown that (i) the dominance of C<sub>60</sub> increases with longer clustering times, (ii) both C<sub>60</sub><sup>+</sup> and C<sub>60</sub><sup>-</sup> are dominant clusters, (iii) there exist a special binding site for carbon-metal complexes, C<sub>60</sub>X (X = La, Ca, Sr, Ba), and (iv) the C<sub>60</sub> cluster is not reactive. These results have led researchers to believe that a single structure is responsible for the experimental observations with the most probable candidate the truncated icosahedron, the so-called buckminsterfullerene. Having the structure of the soccerball, this C<sub>60</sub> configuration has 12 pentagonal and 20 hexagonal faces with each atom identically bonded to 3 atoms.

Previous theoretical investigations of the C<sub>60</sub> structure have been limited to Hartree-Fock calculations and semi-empirical methods.<sup>56, 57</sup> Due to the large number of electrons the quantum chemical approaches have not been able to incorporate all degrees of freedom in the search for the low energy structure and since the parameters for the empirical potential are fitted to bulk properties its applicability to small clusters is unknown. A novel approach to first principles simulations, proposed by R. Car and M. Parrinello, allows complete degrees of freedom while deriving the interaction potential directly from the electronic ground state.<sup>58</sup> The method, involving both density functional theory and classical molecular dynamic simulation has been implemented in a study of the structure and dynamics of the C<sub>60</sub> buckminsterfullerene. The method is presently limited to closed shell ground state functions and energies.

The present calculation of C<sub>60</sub>, reported elsewhere in detail,<sup>59</sup> utilized the Bachelet, Hamann, Schluter pseudopotential<sup>60</sup> and planewaves for the wavefunction expansion with an energy cutoff of 35 Ryd. The Perdew-Zunger<sup>61</sup> form of the LDA was used for the exchange-correlation energy. The simulation employed a supercell of 17.5Å<sup>3</sup> and FCC periodic boundary conditions. Approximately 32,000 planewaves were required for each state. The "mass",  $\mu$ , was fixed at 500 a.u. and the time step for integrating the equations of motion ( Eqs.(17) ) was 3 a.u. The simulation required 256MB of memory and 90 seconds per iteration on the IBM-3090/600J. The wall-clock time was reduced by a factor of 5 by running in parallel on 6 processors. To obtain the ground-state structure and achieve equilibration about 300 hours of 3090/600J cpu time were used.

The initial C<sub>60</sub> configuration had the structure of a soccerball with each bond the same length as in graphite. Once a self-consistent solution to the KS equations had been achieved, geometry optimization was performed by solving a set of steepest descent equations. The final ground-state structure has two different bond lengths, a short bond, 1.389Å, on the edge between two neighboring hexagons and a longer bond, 1.448Å on the edge between an adjacent pentagon and hexagon. The previous quantum chemical calculation found similar results with short and long bonds of 1.453Å and 1.369Å, respectively.<sup>56</sup> The radius of the ball from the present calculation, 3.53Å, agrees with the experimental estimate of 3.5Å.<sup>55</sup>

The ground-state structure obtained from the density functional calculation has been used as input into a Hartree-Fock calculation where the bond order (see Table 6) and the SCF energy (Table 7) were determined.<sup>62</sup> In addition to the energy, Table 7 also contains the results for the electron affinity and the ionization potential.

TABLE 6. Bond order for the C<sub>60</sub> cluster from Hartree-Fock calculations.<sup>62</sup>

	Bond Order double	Bond Order single
Basis set A <sup>(a)</sup>	1.4753	1.1649
Basis set B <sup>(b)</sup>	1.4911	1.1878
Benzene (Basis set A <sup>(a)</sup> )	1.4471	

(a): 9/5 single zeta basis set. See Ref. 62  
(b): 9/5 double zeta basis set. See Ref. 62

TABLE 7. Energies, electron affinity, and ionization potential for the C<sub>60</sub> cluster from Hartree-Fock calculations.

	A <sup>(a)</sup>	B <sup>(b)</sup>	HF <sup>(c)</sup>	EXPT
SCF Energy (Kcal/mol)	-37.632	-37.845	-37.819	
Correlation (Kcal/mol)	-13.10918	-13.0900		
Cohesive Energy	124.31	139.74	114	
Electron affinity (eV)	3.06	1.57	0.80	2.6-2.8
Ionization potential (eV) <sup>(d)</sup>	10.80	9.80	8.24	6.42-7.87

(a): 9/5 single zeta basis set. See Ref. 62.  
(b): 9/5 double zeta basis set. See Ref. 62.  
(c): 7/3 double zeta basis set. See Ref. 56.  
(d) Koopman's theorem.

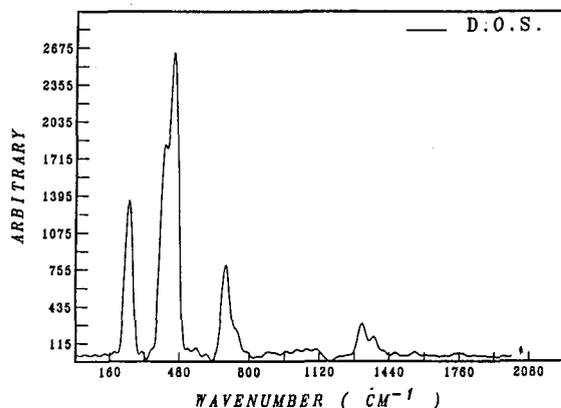


Figure 2. Phonon density of states (DOS) of the buckminsterfullerene.

The dynamics of the C<sub>60</sub> cluster were also investigated. The nuclei were first heated by scaling their velocities to achieve the target temperature. Since heating the atoms this way caused the wavefunction to drift from the Born-Oppenheimer surface, periodic steepest descents quenches for the wavefunction were required to regain the self-consistent solution. Once the velocity scaling ended and a final minimization for the wavefunction performed, the atoms were allowed to evolve in time according to the equations of motions. A total of 1.4 ps (20,000 MD steps) were simulated following the velocity scaling. After a period of equilibration the atomic trajectories were analyzed and the phonon density of states determined. In Fig. 2, the phonon spectrum is presented. The high (1311 cm<sup>-1</sup>) and low (452 cm<sup>-1</sup>) frequency A<sub>g</sub> modes are expected to be Raman active and should be observed in future experiments.<sup>63</sup>

So far the CP method has been primarily limited to semi-conductors in various states, bulk, liquid and clusters. In order to extend the method to more complex systems involving localized electronic density there is a need to include mixed basis sets in the expansion for the wavefunction. Work is in progress in our laboratory along this line.

Next we shall move to another example of molecular dynamics simulation, this time with hydrogen atoms and hydrogen bonds, where it appears that the quantum molecular dynamics approach by Car-Parrinello might not yield accurate results.

### MOLECULAR DYNAMICS OF LIQUID WATER WITH AN AB INITIO POTENTIAL

Recently we have, once more, re-visited<sup>64,65</sup> the task of formulating a new *ab initio* potential for the gas, liquid and crystal phases of water, which improves the MCY (Matsuoka Clementi Yoshimine) two-body potential<sup>66</sup> by including a polarizable water model in place of the computationally expensive three-<sup>67</sup> and four-body<sup>68</sup> corrections. The previous preliminary results<sup>64</sup> (limited to a truncated expression of the model) are now replaced with MD simulations -below summarized- where the full polarization model<sup>64</sup> is tested. The new potential is designated as NCC, a short form for Nieser-Corongiu-Clementi. It consists of two parts

$$V_{NCC} = \sum_{i,j < i} [ V_{\text{two-body}(i,j)} ] + V_{\text{pol}} \quad (4)$$

where the pairwise additive part of the potential is essentially the MCY<sup>66</sup> (refitted and with a few new terms).<sup>64</sup>

The second part of  $V_{NCC}$  is a polarization term. Previous computations, where three<sup>67,69-71</sup> and four-body<sup>68</sup> corrections were considered, have shown that the many-body corrections are necessary for accurate quantitative predictions. However, the computational time increases dramatically when the two-body MCY potential is extended with three- and four-body corrections, term by term. Therefore we require a computationally less expensive algorithm which takes advantage of the "large memory" not available on previous computers.

In the NCC potential we chose an explicit representation of the polarization effects<sup>72,(GC10)</sup> by inducing dipole moments on every interacting molecule. The permanent dipole moments of the water molecules are represented by three point charges per molecule. The polarization on one molecule is hence primarily due to the global point charge distribution of the surrounding matter. These induced dipole moments in turn cause polarization on the other molecules and thus this effect must also be included in the polarization potential. The induced polarization is, as usual, taken as a linear response to the electric field since the field is not very large

$$\mu_{i\lambda}^{ind} = \tilde{\alpha}_{i\lambda} E_{i\lambda}^{tot} \quad (5)$$

where  $\mu_{i\lambda}^{ind}$  is the  $\lambda$ -th induced moment on molecule  $i$ . The polarizability  $\tilde{\alpha}_{i\lambda}$  is assumed to be a static property of molecule  $i$ .

In the spirit of deriving a nonempirical potential to be used for liquid water simulations, all parameters of the NCC potential have been fitted to *ab initio* calculated data. We refer the interested reader to Refs. 64,65 where the *ab initio* calculations of interaction energies for various geometries of 250 water trimer and 350 water dimer configurations are discussed (a list of the interaction energies and geometries are given in Ref. 64b). We use the *ab initio* interaction energies of trimers of water molecules to fit the many-body parameters, namely in the parameterization of the locations of the induced dipole moments and the point charges, as well as the polarizability and the value  $q$  of the point charge. This set of electrostatic parameters was then used -unmodified- in the fit of the two-body potential to *ab initio* calculated interaction energies of water dimers.

One important characterization of the water dimer potential surface is its absolute minimum, *i.e.* the most stable water dimer configuration. This configuration is predicted as an open form with a nearly linear hydrogen bond. The binding energy is computed as 5.18 kcal/mole at an intermolecular O-O distance of 2.97 Å. The potential around this configuration is very flat. In comparing this data with experimental values, care has to be taken since the available experimental results span a rather broad range, particularly with respect to the binding energy.<sup>73</sup> Recent results<sup>74,75</sup> seem to settle around 5.3 kcal/mole; possibly  $5.1 \pm 0.3$  kcal/mole is an acceptable estimate.

The reliability of the NCC potential is assessed by simulating water properties in the gas phase (monomer and dimer), in the liquid phase (static, dynamic and collective properties), as well as in the crystal phase (1h phase of ice), thus covering a very extended range of properties seldom reliably simulated by the same model, without *ad hoc* reparametrization.

In Fig. 3 we report a comparison of the X-ray and neutron beam scattering intensities obtained with the NCC potential and by experiments<sup>76,77</sup> (top insets). We recall that a correct simulation of the second peak in the liquid water oxygen-oxygen pair correlation is a necessary requirement to obtain reliable

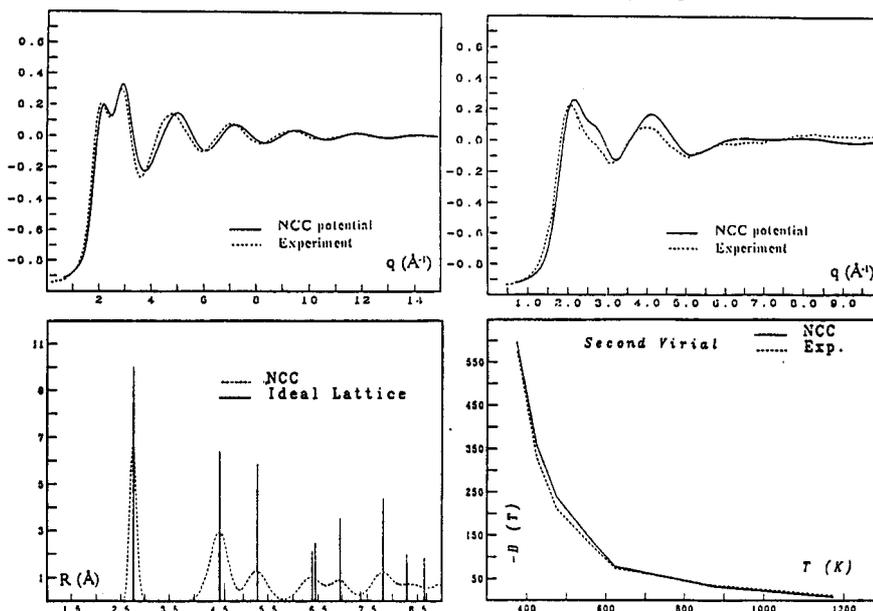


Figure 3. The X-ray (top left), neutron (top right) scattering intensities, pair correlation function  $g_{O-O}$  of ice (bottom left), and second virial coefficient of steam (bottom right) from the NCC potential and experimental data.

X-ray and neutron beam intensities. To this figure, we add a comparison of the oxygen-oxygen pair correlation function of the ideal lattice of Ih ice<sup>78</sup> and the one obtained with a 20 picosecond MD simulation for a periodic system of 192 water molecules at 50 K (left bottom inset). Finally, in Fig. 3 we include the computed and experimental<sup>79</sup> second virial coefficient of steam (bottom right inset). We note, that the second virial coefficient is simulated without accounting for the water vibrational modes<sup>80</sup> (the present simulation assumes water as a rigid molecules).

Fig. 4 shows the good agreement between the NCC pair correlation functions and those derived from experiments by Narten<sup>76</sup> and Soper<sup>81</sup> (left insets). We have also added the pair correlation functions by Cieplak, Kollman and Lybrand,<sup>82</sup> CKL, (right insets) who use a polarization potential similar to the one of Ref. 64, however, with empirical parametrization. The *ab initio* potential appears to be more realistic than the empirical one.

Below we report simulated quantities obtained with NCC. These are the evaporation energy (-10.65 kcal/mole without vibrational correction vs -9.98 kcal/mole from experiments<sup>83</sup>), the average dipole moment in liquid water at room temperature (2.8 D vs an inferred experimental value<sup>84</sup> of 2.45 D), the dipole moment of one water molecule (1.85 D vs 1.83 D from experiments<sup>84</sup>), the heat capacity at constant volume (17.6 cal/(mole K) vs 17.9 cal/(mole K) from experiments<sup>85</sup>), the diffusion coefficient ( $2.5 \times 10^{-5}$  cm<sup>2</sup>/sec vs  $2.4 \times 10^{-5}$  cm<sup>2</sup>/sec from experiments<sup>86</sup>), the <sup>1</sup>H-NMR relaxation time  $\tau_2$  ( $1.8 \pm 0.2$  ps vs 2.0 ps from experiments<sup>87</sup>), the low frequency sound mode (1280 m/sec vs  $1390 \pm 100$  from experiments<sup>88</sup>), the high frequency sound mode ( $3200$  m/sec vs  $3310 \pm 250$  from experiments<sup>89</sup>). The computed dimer energy is -5.18 kcal/mole compared with an experimental value<sup>73-75</sup> of  $-5.4 \pm 0.2$  kcal/mole; despite the good agreement, we are of the opinion that these values are somewhat overestimated. The computed second virial coefficient values in cm<sup>3</sup>/mole are -596 (at T=373 K), -239 (at T=479 K), -127 (at T=573 K) and -77 (at T=673 K), to be compared with the experimental values<sup>79</sup> of -580, -212, -117, and -73, respectively. From analysis of the rotational autocorrelation function, we obtain an IR librational band with a peak at about 700 cm<sup>-1</sup> to be compared with the experimental<sup>90</sup> band also at 700 cm<sup>-1</sup>. The time of flight and the density of states, computed from the coherent and incoherent structure factors, respectively, are in qualitative agreement with available low resolution experiments.<sup>91,92</sup> The computed pressure is -1100 atm. The somewhat short distance position, relative to experiments,

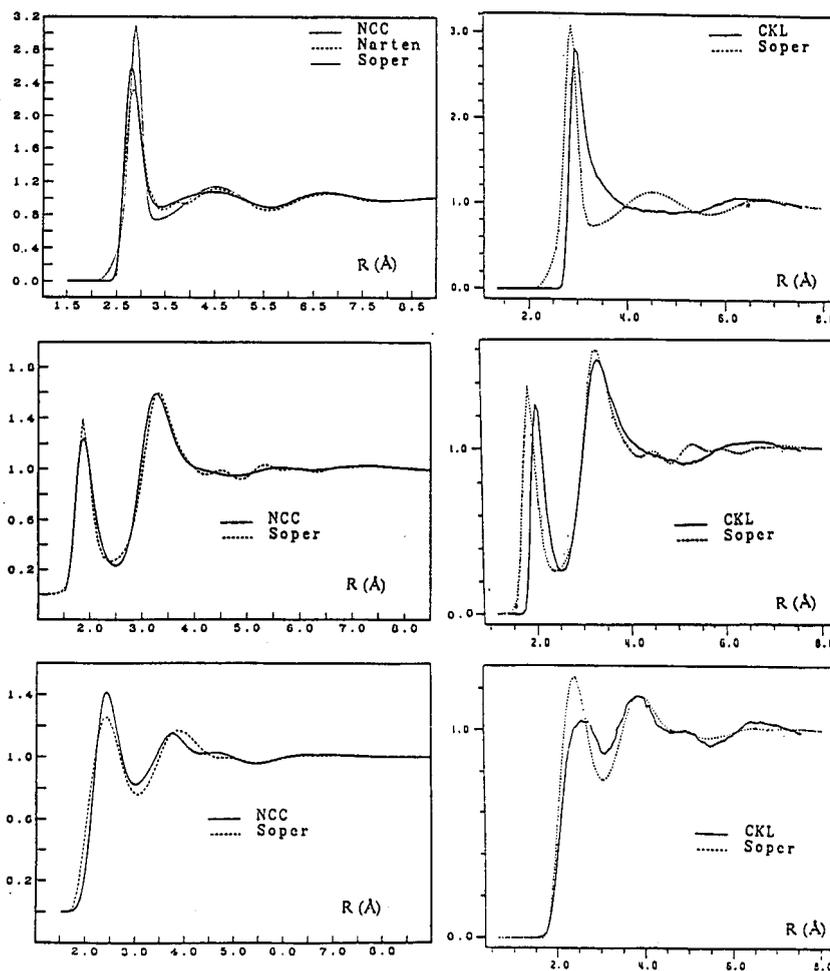


Figure 4. In the right insets, from top to bottom, the pair correlation functions  $g_{O-O}(r)$ ,  $g_{O-H}(r)$  and  $g_{H-H}(r)$  obtained from the NCC potential and experimental data. In the left insets, the same quantities from the CKL potential.

of the first peak for the  $g_{oo}(r)$ , the possibly a bit overestimated binding energy of the dimer, and the error in the pressure can be "corrected" by forcing less binding in the potential; indeed, by decreasing one of the parameters by only 1.2% one can obtain a nearly exact position for the first peak, a pressure around 300 atm and an evaporation energy of 9.92 kcal/mole. Preliminary results<sup>93</sup> from a simulation with a flexible water model potential predict an evaporation energy around 10.3 kcal/mole and a pressure of about -1800 atm. Other data are not too dissimilar from those of Ref. 94, but a much longer simulation is needed to confirm the above data. This concludes the simulations carried out to test the NCC potential; we note that the above list is, to our knowledge, the most comprehensive one in literature to test the reliability of a potential -with the same model and parameters- in simulations of water in the gas, liquid, and crystal phases.

The MD simulations for the liquid water have been obtained from a periodic system of 512 water molecules, at 310 K, using a sixth order Gear-predictor corrector algorithm, Ewald-sums, a cutoff radius of 12.4 Å, a time step of 0.5 femtoseconds, and a simulation length of 32 ps after equilibration. The MD computer program, KGNNCC, for the NCC model, is made available to the interested reader, as part of the MOTEC-90 initiative.<sup>95</sup> The computer time for this simulation is 53 hours for equilibration and 34 hours for the collection of statistics on an IBM ES/3090-600J running in parallel. The entire calculation takes about 40MBytes of main memory. The simulation with the flexible model takes 4 times longer, due to the smaller time step. Notice that by increasing the sample from 512 to 1000 molecules, the execution time increases by a factor of about 8.

## CONCLUSIONS

All the computations above neglect relativistic effects. Whereas this might not be too inconsistent for a computation on  $H_2$  or  $H_3$  aiming at 1 to 10  $cm^{-1}$  accuracy, it clearly becomes a more and more questionable approximation when one deals with molecules with atoms past Neon and aims at millihartree accuracy while neglecting relativistic effects. It has been noted long ago<sup>96</sup> that relativistic effects scale as the second to third power of the atomic number ( $Z^2$  and  $Z^3$ ) while correlation effects scale with the number of electrons ( $Z$  for neutral atoms) and are taken over by the former (the relativistic ones) at about  $Z=12$ . This "caveat" will likely be acknowledged more and more in the coming years. If we compare an Hartree-Fock computation with a Fock-Dirac computation for a given molecule and with the "same" basis set we note that the latter will require about 5 to 10 times more CPU time than the former. There are three main reasons for this.<sup>97</sup> Firstly, the kinetic balance<sup>98</sup> brings about an effective increase in the starting basis set; then the SCF part is more complex in the Fock-Dirac formalism, and finally if one starts with cartesian Gaussian functions there is the need of a transformation to spherical functions. In conclusion, presently relativistic computations are problematic mainly because we are still searching for proper methods, but once this problem is solved, these will turn out to be five to ten times more expensive than equivalent non relativistic computations, a factor which is offset by the performance gain of the Vth generation relative to the IVth.

Let us now consider the  $H_3$  surface with a 1.0 to 10.0  $cm^{-1}$  accuracy and with the HCI technique. From the discussion in Section 2, it appears that - assuming one has improved the numerical analyses and the code - a reasonably accurate surface can be obtained with a 50-100 GFLOPS system. Keeping in mind the recent performance announcements particularly from the Japanese computer industry, there is reason to assume that this problem will have an accurate HCI solution before the end of this century, namely from the VIth generation computers. However one could consider alternatives, particularly Quantum Monte Carlo techniques<sup>1</sup> or geminal techniques.<sup>33</sup>

Essentially the same conclusion holds for atomic computations with the CI technique: systems with a 50 GFLOPS performance could do the job. There is, however, an apparent need to expand the ATOMCI code either in the direction of Multi Reference CI or of MCSCF followed by extended CI. Present efforts by Prof. Sasaki and coworkers to parallelize the code are another most welcome step. However one could also look at alternatives like "direct" techniques,<sup>99</sup> thus using many more configurations and at the same time circumventing the I/O bottleneck. But it is unlikely that we will see many computations with 30-40 electrons and with an *absolute* accuracy of 0.001 a.u. in this century.

Let us now move to large systems, for example liquid water, to be simulated by a sample of 1000 molecules of water using molecular dynamics. This larger than the usual number of water molecules has been chosen mainly to ensure proper treatment of collective effects, like the dielectric constant and sound wave simulations, where there are large fluctuations. Can we simulate this system using quantum mechanics at each time step for energy and forces? Alternatively stated, can we use the Car-Parrinello technique?<sup>58</sup> Scaling from 32000 plane waves needed for 60 carbon atoms to the 3000 atoms of the 1000 water molecules, we estimate the need of about  $1.5 \times 10^6$  plane waves. The corresponding CPU time on an IBM 3090J is conservatively about 90 hours per time step with a storage requirement of 20 Gbytes. This is contrasted to about 1.4 minutes per time step using *ab initio* potentials of NCC type and a flexible water. While a single Car-Parrinello time step might be reduced to less than one hour on a large VIth generation computer, the  $10^5$  required time steps seem to make this task unfeasible within this century. Notice that the number of plane waves could be considerably smaller by using a mixed basis, namely plane waves and Gaussian orbitals localized on the nuclei. Work is in progress in our laboratory along this line.

We would like to forecast that direct use of quantum mechanics will be more and more diffuse in molecular dynamics computations. Considering the analysis made above however, it seems that one will have

to merge energies and forces quantum mechanically obtained with energies and forces derived from *ab initio* potentials. This will be particularly important for the study of the active site in protein and enzymes. Indeed, only in the neighborhood of the active site one will follow the reaction at the quantum mechanical level while all the remaining interactions would be treated at the classical level, using *ab initio* derived potentials. A precursor of this approach is a study of the proton transfer in papaine.<sup>100</sup>

The common denominator in our conclusions is that if we rely mainly on the increased performance offered by the computer industry, then computational chemistry will only inch forward at best. The main avenue lies in alternative methods and new techniques. Thus Quantum Monte Carlo, *ab initio* derived polarization potentials, Car-Parrinello and its extension (as outlined above), MBPT, geminals and direct techniques seem to represent the future of quantum chemistry. At the other side of the spectrum, semi-empirical methods, including Density Functional added to the Hartree-Fock, should be also considered. Finally, techniques designed to move away from basis set expansion, toward numerical methods, should be examined with rigor, both to escape the BSS error and the  $N^4$  dependency.

**Acknowledgements** It is my pleasure to thank Drs. G. Corongiu, B. Feuston, D. Frye and A. Rizzo for their permission to use unpublished data.

## REFERENCES

1. M. H. Kalos *Phys. Rev.* **128**, 1791 (1962); D. M. Ceperley, B.J. Alder, *Science* **231**, 555 (1986). J. Anderson, *J. Chem. Phys.* **85**, 2839, (1987) and references therein.
2. M. Parrinello, *Modern Techniques in Computational Chemistry: MOTECC-90*, E. Clementi, Ed., ESCOM, Leiden, pp 731-744 (1990).
3. D. K. Bhattacharya, E. Clementi, *Modern Techniques in Computational Chemistry: MOTECC-90*, E. Clementi, Ed., ESCOM, Leiden, pp 889-918 (1990).
4. A. M. De Callatay, *Natural and Artificial Intelligence*, North-Holland, Amsterdam, (1986).
5. R. J. Bartlett, *Ann. Rev. Phys. Chem.* **32**, 359 (1981); M. Urban, I. Cernusak, V. Kelló, J. Noga, *Methods in Computational Chemistry*, S. Wilson, Ed., Plenum, New York, (1987).
6. E. A. Hylleraas, *Z. Physik* **54**, 347 (1929).
7. H. M. James, A. S. Coolidge, *J. Chem. Phys.* **1**, 825 (1933).
8. W. Kolos, L. Wolniewicz, *J. Chem. Phys.* **41**, 3663 (1964); *ibid.* **43**, 2429 (1965); *Phys. Rev. Lett.* **20**, 243 (1968).
9. R. S. Mulliken, W. C. Ermler, *Diatomc Molecules: Results of Ab Initio Calculations*, Academic, New York, (1977).
10. H. A. Bethe, E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms*, Plenum/Rosetta, New York, (1977).
11. W. Kutzelnigg, *Theor. Chim. Acta* **68**, 445 (1985).
12. D. Frye, G. C. Lie, E. Clementi, *HYCOIN: Documentation and User's Guide*; IBM Technical Report KGN-180, Kingston, (1989).
13. D. Frye, A. Preiskorn, G. C. Lie, E. Clementi, *Modern Techniques in Computational Chemistry: MOTECC-90*, E. Clementi, Ed., ESCOM, Leiden, pp 235-275 (1990).
14. D. Frye, A. Preiskorn, E. Clementi, *J. Comp. Chem.*, submitted.
15. C. Urdaneta, A. Largo-Cabrerizo, J. Lievin, G. C. Lie, E. Clementi, *J. Chem. Phys.* **88**, 2091 (1988).
16. D. Frye, G. C. Lie, E. Clementi, *J. Chem. Phys.* **91**, 2366 (1989).
17. D. Frye, G. C. Lie, E. Clementi, *J. Chem. Phys.* **91**, 2369 (1989).
18. A. Preiskorn, G. C. Lie, D. Frye, E. Clementi, *J. Chem. Phys.* **92**, 4941 (1990).
19. D. Frye, A. Preiskorn, G. C. Lie, E. Clementi, *J. Chem. Phys.* **92**, 4948 (1990).
20. A. Preiskorn, D. Frye, E. Clementi, in preparation.
21. D. Frye, G. C. Lie, E. Clementi, *HCI Calculations on the Potential Energy Curve for the  $X^1\Sigma_g^+$  State of  $\text{H}_2$* ; IBM Technical Report KGN-179; Kingston, (1989).
22. W. Kolos, K. Szalewicz, H. Monkhorst, *J. Chem. Phys.* **84**, 3259 (1986).
23. B. Liu, *J. Chem. Phys.* **80**, 581 (1984).
24. D. Feller, private communication.
25. L. Salmon, R. D. Poshusta, *J. Chem. Phys.* **59**, 3497 (1973).
26. F. Mentch, J. B. Anderson, *J. Chem. Phys.* **74**, 6307 (1981).
27. A. Preiskorn, W. Woznicki, *Mol. Phys.* **52**, 1291 (1984).
28. P. G. Burton, E. Von N.-Felsobuki, G. Doherty, M. Hamilton, *Mol. Phys.* **55**, 527 (1985).
29. W. Meyer, P. Botschwina, P. Burton, *J. Chem. Phys.* **84**, 891 (1986).
30. J. B. Anderson, *J. Chem. Phys.* **86**, 2839 (1987).
31. C. A. Traynor, J. B. Anderson, *Chem. Phys. Lett.* **147**, 389 (1988).
32. S. Huang, Z. Sun, W. A. Lester, Jr., *J. Chem. Phys.* **92**, 597 (1990).
33. S. A. Alexander, H. J. Monkhorst, R. Roeland, K. Szalewicz, to be published.
34. E. Davidson, *Modern Techniques in Computational Chemistry: MOTECC-90*, E. Clementi, Ed., ESCOM, Leiden, pp 553-592 (1990).
35. E. Clementi, G. Corongiu, S. Chakravorty, *Modern Techniques in Computational Chemistry: MOTECC-90*, E. Clementi, Ed., ESCOM, Leiden, p 346 (1990).
36. A. Rizzo, E. Clementi, M. Sekiya, *Chem. Phys. Lett.*, (in press).
37. F. Sasaki, *Int. J. Quantum Chem.*, **8**, 605 (1974).
38. F. Sasaki, M. Yoshimine, *Phys. Rev. A*, **9**, 17 (1974).
39. F. Sasaki, M. Yoshimine, *Phys. Rev. A*, **9**, 26 (1974).
40. F. Sasaki, M. Sekiya, T. Noro, K. Ohtsuki, T. Osanai, *Modern Techniques in Computational Chemistry: MOTECC-90*, E. Clementi, Ed., ESCOM Science Publisher, Leiden, p 181 (1990).
41. M. Sekiya, T. Noro, K. Ohtsuki, F. Sasaki, A. Rizzo, E. Clementi, *MOTECC-90: Input/Output Documentation*; a special publication by the Department of Scientific Engineering Computations, IBM Kingston, (1990).

42. E. Clementi, G. Corongiu, *Chem. Phys. Lett.*, **90**, 359 (1982).
43. E. Clementi, *J. Chem. Phys.*, **38**, 2248 (1963).
44. E. Clementi, A. Veillard, *J. Chem. Phys.*, **44**, 3050 (1966).
45. C. L. Pekeris, *Phys. Rev.*, **112**, 1649 (1958).
46. C. L. Pekeris, *Phys. Rev.*, **115**, 1216 (1959).
47. E. R. Davidson, *Int. J. Quantum Chem.*, **37**, 811 (1990).
48. A. Veillard, E. Clementi, *J. Chem. Phys.*, **49**, 2415 (1968).
49. S. Larsson, *Phys. Rev.*, **169**, 49 (1968).
50. D. Sundholm, J. Olsen, *Phys. Rev.*, **42**, 2614 (1990).
51. D. Feller, C. M. Boyle, E. R. Davidson, *J. Chem. Phys.*, **86**, 3424 (1987).
52. C. F. Bunge, *Phys. Rev. A*, **14**, 1965 (1976).
53. H. Hartmann, E. Clementi, *Phys. Rev.*, **133**, 1295 (1964).
54. K. S. Pitzer, E. Clementi, *J. Am. Chem. Soc.*, **81**, 4477 (1959).
55. J. R. Health, S. C. O'Brien, R. F. Curl, H. W. Kroto, R. E. Smalley, *Com. on Cond. Matt. Phys.* **13**, 119 (1987).
56. H. P. Luthi, J. Almlof, *Chem. Phys. Lett.* **135**, 357 (1987).
57. M. D. Newton, R. E. Stanton, *J. Am. Chem. Soc.* **108**, 2469 (1986).
58. R. Car, M. Parrinello, *Phys. Rev. Lett.* **55**, 2471 (1985).
59. B. P. Feuston, E. Clementi, W. Andreoni, M. Parrinello, *to be published*.
60. G. B. Bachelet, D. R. Hamann, M. Schuller, *Phys. Rev. B* **23**, 5048 (1982).
61. J. P. Perdew, A. Zunger, *Phys. Rev. B* **23**, 5048 (1981).
62. G. Corongiu, E. Clementi, *to be published*.
63. E. Brendals, B. N. Cyvin, J. Brunvoli, S. J. Cyvin, *Spectro. Lett.* **21**, 313 (1988).
64. a) U. Niesar, G. Corongiu, M.-J. Huang, M. Dupuis, E. Clementi, *Int. J. Quant. Chem. Symp.* **23** 421 (1989); b) U. Niesar, G. Corongiu, M.-J. Huang, M. Dupuis, E. Clementi, IBM Technical Report KGN-191: Kingston, (1989).
65. U. Niesar, G. Corongiu, E. Clementi, G. R. Kneller, D. K. Bhattacharya, *J. Phys. Chem.* **94**, 7949 (1990).
66. O. Matsuoka, E. Clementi, M. Yoshimine, *J. Chem. Phys.* **64**, 1351 (1976).
67. E. Clementi, G. Corongiu, *Int. J. Quant. Chem. Symp.* **10**, 31 (1983).
68. J. H. Detrich, G. Corongiu, E. Clementi, *Chem. Phys. Lett.* **112**, 426 (1984).
69. M. Wojcik, E. Clementi, *J. Chem. Phys.* **84**, 5970 (1986).
70. M. Wojcik, E. Clementi, *J. Chem. Phys.* **85**, 3544 (1986).
71. M. Wojcik, E. Clementi, *J. Chem. Phys.* **85**, 6085 (1986).
72. J. O. Hirschfelder, C. F. Curtiss, R. B. Bird, *Molecular Theory of Gases and Liquids*, John Wiley, New York, (1954).
73. P. Obza, R. Zahradnik, *Chem. Rev.* **88**, 871 (1988).
74. T. R. Dyke, K. M. Mack, J. S. Muenter, *J. Chem. Phys.* **66**, 498 (1977).
75. L. A. Curtiss, D. J. Frurip, M. L. Blander, *J. Chem. Phys.* **64**, 1120 (1982).
76. A. H. Narten, H. A. Levy, *J. Chem. Phys.* **55**, 2263 (1971).
77. W. E. Thiessen, A. H. Narten, *J. Chem. Phys.* **77**, 2656 (1982).
78. P. W. Deutsch, B. N. Hale, R. C. Ward, D. A. Reago, Jr., *J. Chem. Phys.* **78**, 5103 (1983).
79. G. S. Kell, G. E. McLaurin, E. Whalley, *J. Chem. Phys.* **64**, 5308 (1968).
80. K. Refson, G. C. Lie, E. Clementi, *J. Chem. Phys.* **87**, 3634 (1987).
81. A. K. Soper, R. N. Silver, *Phys. Rev. Lett.* **49**, 471 (1982); A. K. Soper, M. G. Phillips, *Chem. Phys.* **107**, 47 (1986).
82. P. Cieplak, P. Kollman, T. Lybrand, *J. Chem. Phys.* **92**, 6755 (1990).
83. N. E. Dosey, *Properties of Ordinary Water Substances*, Reinhold, New York, (1940).
84. D. Eisenberg, W. Kauzmann, *The structure and properties of water*, Clarendon, Oxford, (1969).
85. G. S. Kell, *Water: A Comprehensive Treatise, Vol. 1*, F. Franks, Ed., Plenum, New York, p 363 (1971).
86. K. Krynicki, C. D. Green, D. W. Sawyer, *Discuss. Faraday Soc.* **66**, 199 (1978).
87. R. G. Gordon, *Adv. Mag. Resonance* **3**, 1 (1968).
88. J. Teixeira, J. Leblond, *J. Phys. Lett.* **39**, 83 (1978); see also A. J. Matheson, *Molecular Acoustics*, Wiley, New York, (1971).
89. J. Teixeira, M. C. Bellissent-Funel, S. H. Chen, B. Dorner, *Phys. Rev. Lett.* **54**, 2681 (1985).
90. G. E. Walrafen, *Water: A Comprehensive Treatise, Vol. 1*, F. Franks, Ed., Plenum, New York, p. 151 (1971).
91. S. H. Chen, J. Teixeira, *Adv. Chem. Phys.* **64**, 1 (1988).
92. D. I. Page, *Water: A Comprehensive Treatise, Vol. 1*, F. Franks, Ed., Plenum, New York, p. 333 (1971).
93. G. Corongiu, E. Clementi, unpublished data.
94. G. C. Lie, E. Clementi, *Phys. Rev. A* **33**, 2679 (1986).
95. E. Clementi, Ed., *Modern Techniques in Computational Chemistry, MOTECC-90*, ESCOM, Leiden, (1990).
96. E. Clementi, *J. Mol. Spec.* **12**, 18 (1964).
97. A. Mohanti, private communication.
98. A. Mohanti, E. Clementi *J. Chem. Phys.* **93**, 1829 (1990).
99. B. Roos, P. E. Siegbahn, *Methods of Electronic Structure Theory*, H. F. Schaefer III, W. H. Miller, Eds., Plenum: New York, (1977); I. Shavitt, *ibid.*, (1977); E. Davidson, *J. Comput. Phys.* **17**, 87 (1975).
100. G. Bolis, M. Ragazzi, D. Salvaderi, D. R. Ferro, E. Clementi, *Int. J. Quantum Chem.* **14**, 815 (1978).