

## Møller–Plesset third order calculations with large basis sets

Ian L. Alberts and Nicholas C. Handy

Citation: *J. Chem. Phys.* **89**, 2107 (1988); doi: 10.1063/1.455107

View online: <http://dx.doi.org/10.1063/1.455107>

View Table of Contents: <http://jcp.aip.org/resource/1/JCPSA6/v89/i4>

Published by the AIP Publishing LLC.

---

### Additional information on J. Chem. Phys.

Journal Homepage: <http://jcp.aip.org/>

Journal Information: [http://jcp.aip.org/about/about\\_the\\_journal](http://jcp.aip.org/about/about_the_journal)

Top downloads: [http://jcp.aip.org/features/most\\_downloaded](http://jcp.aip.org/features/most_downloaded)

Information for Authors: <http://jcp.aip.org/authors>

## ADVERTISEMENT



Explore the **Most Cited**  
Collection in Applied Physics

AIP  
Publishing

# Møller–Plesset third order calculations with large basis sets

Ian L. Alberts and Nicholas C. Handy

University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, United Kingdom

(Received 19 April 1988; accepted 28 April 1988)

The value of the Møller–Plesset third order calculations is examined. An efficient method for the evaluation of the gradient of the MP3 energy is reported, and it has been programmed for both restricted and unrestricted Hartree–Fock wave functions. Large basis set calculations (TZ2P or better) are reported for the optimization of geometries and the determination of harmonic frequencies (which are obtained by finite differences of analytic gradients). The molecules selected are  $\text{NH}_2$ ,  $\text{PH}_2$ ,  $\text{AsH}_2$ ,  $\text{H}_2\text{O}$ ,  $\text{NH}_3$ ,  $\text{H}_2\text{CO}$ ,  $\text{HCN}$ , and  $\text{C}_2\text{H}_2$ . For the closed shell systems, the RMP3 predictions for bond lengths are inferior ( $\approx 0.006$  Å) to RMP2 predictions ( $\approx 0.003$  Å) for single bonds, and for multiple bonds the RMP3 bond lengths are too short by approximately the same amount ( $\approx 0.01$  Å) that RMP2 are too long. For the open shell systems, the UMP3 geometrical parameters show only a marginal improvement over UMP2, except for  $\text{PH}_2$  where the bond length error is reduced to  $0.003$  Å. The results for harmonic frequencies show a similar comparison between MP2 and MP3. On the basis of this experience, it appears that large scale calculations at the MP3 level are not recommended; MP2 calculations with a large basis set are much cheaper and provide results with a similar, if not superior, accuracy.

## I. INTRODUCTION

The accurate prediction of the properties of small and medium size molecules is now a goal within the grasp of the quantum chemist. He has available a range of methods, of a varying level of sophistication. In this paper we shall look at the perturbation approach for the inclusion of electron correlation, commonly called the Møller–Plesset (MP) method.<sup>1</sup>

The MP approach has been popularized by its availability in the GAUSSIAN series of programs, specifically at second, third, and fourth order,<sup>2</sup> MP2, MP3, and MP4. For closed shell systems, one proceeds from the restricted Hartree–Fock approximation, and thus the series RMP2, RMP3, and RMP4 is obtained. For open shell systems, the unrestricted Hartree–Fock<sup>3</sup> approximation must be used, leading to the series UMP2, UMP3, and UMP4. For a review of the results which may be obtained with those methods, see Ref. 2.

Recently, it has been shown that at the RMP2 level, substantially improved accuracy is achieved if basis sets larger than the recommended 6-31G\* (or 6-31G\*\*) are used.<sup>4</sup> If one performs calculations with a TZ2P basis set, then there are sufficient calculations to show that for single bonds an accuracy of  $\approx 0.003$  Å may be achieved, and for multiple bonds  $\approx 0.01$  Å. This represents an improvement by a factor of 2 or 3 over 6-31G\* basis calculations. With this RMP2 method, automatic optimization of structures through analytic gradients is possible at little extra cost over the energy calculation. Simandiras *et al.*<sup>4</sup> have also shown that RMP2 analytic second derivatives may be evaluated efficiently; these are now available in the quantum chemistry package CADPAC.<sup>5</sup> The results of calculations by this approach have shown that it is possible to achieve a 2% average accuracy in predicted harmonic frequencies, when a TZ2P (+f) basis set is used. This is a substantial improvement over the

$\approx 10\%$  accuracy which is obtained by the RHF method and brings quantum chemistry into the realm of useful predictive powers, without the need for any scaling procedures.<sup>6</sup>

MP2 is the simplest scheme for the inclusion of electron correlation. No iterative procedure is necessary. Its success outlined above is rather remarkable. The question must then be asked whether this accuracy can be maintained as the level of sophistication is improved. The next level of sophistication is MP3 which is an  $n^6$  procedure. It will be cheaper than any of the other improved levels of sophistication, all of which require iterations; we may mention CISD,<sup>7</sup> CEPA- $n$ ,<sup>8</sup> CPF,<sup>9</sup> and CCSD,<sup>10</sup> which are single-reference based and limited to single and double excitations. With the exception of CISD, all these methods are size consistent<sup>11</sup> which is now recognized to be more important than obeying the variational upper bound principle.

The purpose here is an in-depth examination of MP3. We shall investigate both RMP3 and UMP3. Before commencing, we recall some evidence from the convergence of the MP $n$  series.<sup>12</sup> The convergence of the RMP $n$  series was erratic, especially at distorted geometries. In our calculations, we find that for  $\text{H}_2\text{O}$ , with a TZ2P basis set, the energy lowering from RMP2, RMP3, and RMP4, at equilibrium geometry, is 0.265, 0.003, and 0.018 hartree. In other words RMP3 does not appear to give the improvement that RMP2 and RMP4 achieve. At  $1.5r_e$  with  $C_{2v}$  symmetry the MP3 contribution was found to raise the energy.<sup>12</sup> Of course, we note that MP2 and MP3 deal only with double excitations, whereas MP4 introduces single, triple, and quadruple excitations. On the other hand, the convergence of the UMP $n$  series is much more smooth, although at distorted geometries the series is slowly convergent, predominantly it appears due to spin contamination. For  $\text{NH}_2$ , at  $r_e$  with the TZ2P basis set, we find that the energy lowering from the UMP2, UMP3, and UMP4 calculations is 0.196, 0.016, and 0.011 hartree. We thus anticipate for the later results in this

paper that the UMP3 and RMP3 calculations exhibit a different behavior.

Before we can proceed with a discussion of our results, we shall outline the method we have used to evaluate MP3 gradients for the optimization of geometries. There have been two earlier papers on the evaluation of these gradients; the first by Gauss and Cremer,<sup>13</sup> which when their Appendix is included, gives an efficient way to proceed. The second paper by Bartlett and co-workers,<sup>14</sup> published almost simultaneously, outlines the same procedure, and presents some DZ basis set calculations on H<sub>2</sub>O. There is no evidence in these papers to suggest that efficient algorithms have been written.

For a method to be viable and successful, it must be capable of being used with large basis sets (i.e., more than 100 basis functions). Even on today's computers, that means the programs must operate "out-of-core." The best algorithms will take advantage of available vector or array facilities. For these reasons we present in the next section our approach for the evaluation of UMP3 (and RMP3) gradients, and we discuss some of the complexities which were met in the efficient evaluation.

## II. THE EVALUATION OF MP3 GRADIENTS

The notation we use is the same as that presented in our earlier papers on the derivation of MP2 analytic second derivatives<sup>15</sup> and MP2 formulas for infrared intensities.<sup>16</sup> Thus a perturbed orbital is represented, for example, through

$$\phi_i^{(x)} = \sum_p U_{pi}^x \phi_p + \phi_i^x, \quad (1)$$

where  $\phi_i^x$  denotes the evaluation of  $\phi_i$  with differentiated basis functions, and  $U_{pi}^x$  are solutions of the coupled perturbed Hartree-Fock (CPHF) equations.<sup>17</sup> The orthonormality constraint yields

$$U_{pq}^x + U_{qp}^x + S_{pq}^x = 0. \quad (2)$$

We shall derive the theory for the UMP3 gradient, superscripts of  $\alpha$  or  $\beta$  will identify the spin of the orbitals being discussed,  $i, j, \dots$  denote occupied orbitals,  $a, b, \dots$  denote virtual orbitals, and  $p, q, \dots$  refer to all orbitals.

In the UHF framework, the coupled equations (CPUHF) take the form

$$\begin{aligned} (\epsilon_a^\alpha - \epsilon_i^\alpha) U_{ai}^{\alpha x} + \sum_{bj} A_{aibj}^{\alpha\alpha} U_{bj}^{\alpha x} + \sum_{bj} A_{aibj}^{\alpha\beta} U_{bj}^{\beta x} \\ = -\epsilon_{ai}^{\alpha x} + S_{ai}^{\alpha x} \epsilon_i^\alpha + \frac{1}{2} \sum_{kj} A_{aikj}^{\alpha\alpha} S_{kj}^{\alpha x} + \frac{1}{2} \sum_{kj} A_{aikj}^{\alpha\beta} S_{kj}^{\beta x}, \end{aligned} \quad (3)$$

where

$$\begin{aligned} A_{pqrs}^{\alpha\alpha} &= 2(pq|rs)^{\alpha\alpha} - (pr|qs)^{\alpha\alpha} - (ps|qr)^{\alpha\alpha}, \\ A_{pqrs}^{\alpha\beta} &= 2(pq|rs)^{\alpha\beta}, \end{aligned} \quad (4)$$

and

$$\epsilon_{pq}^\alpha = h_{pq}^\alpha + \sum_j [(pq|jj)^{\alpha\alpha} - (pj|qj)^{\alpha\alpha}] + \sum_j (pq|jj)^{\alpha\beta}. \quad (5)$$

Equation (3), together with its  $\beta$  analogs, may be summarized as

ized as

$$\begin{pmatrix} \mathbf{H}^{\alpha\alpha} & \mathbf{H}^{\alpha\beta} \\ \mathbf{H}^{\beta\alpha} & \mathbf{H}^{\beta\beta} \end{pmatrix} \begin{pmatrix} \mathbf{U}^{\alpha x} \\ \mathbf{U}^{\beta x} \end{pmatrix} = \begin{pmatrix} \mathbf{b}^{\alpha x} \\ \mathbf{b}^{\beta x} \end{pmatrix} \quad (6)$$

and they may be solved for the independent occupied-virtual block of the  $\mathbf{U}^x$  matrix. The dependent members (occupied-occupied or virtual-virtual) may be determined through

$$\begin{aligned} U_{pq}^{\alpha x} &= \frac{1}{(\epsilon_p^\alpha - \epsilon_q^\alpha)} \left( -\epsilon_{pq}^{\alpha x} + S_{pq}^{\alpha x} \epsilon_q^\alpha + \sum_{bj} A_{pqbj}^{\alpha\alpha} U_{bj}^{\alpha x} \right. \\ &\quad + \sum_{bj} A_{pqbj}^{\alpha\beta} U_{bj}^{\beta x} - \frac{1}{2} \sum_{kj} A_{pqkj}^{\alpha\alpha} S_{kj}^{\alpha x} \\ &\quad \left. - \frac{1}{2} \sum_{kj} A_{pqkj}^{\alpha\beta} S_{kj}^{\beta x} \right) \end{aligned} \quad (7)$$

but this is not satisfactory in the case of nearly degenerate orbitals ( $\epsilon_p = \epsilon_q$ ). We use the alternative scheme

$$U_{pq}^x = U_{qp}^x = -\frac{1}{2} S_{pq}^x \quad (8)$$

but we must then take care that the matrix  $\epsilon$  is not diagonal at the perturbed geometry; i.e.,  $\epsilon_{pq}^{(x)} \neq 0$ .

In a spin-orbital formalism, and a nondiagonal representation of the Fock matrix (i.e., not canonical orbitals) the MP3 energy may be written

$$\begin{aligned} E_3 &= - \sum_{nmpq} \langle 0 | H | D_n \rangle (H_0)_{nm}^{-1} \\ &\quad \times \langle D_m | H | D_q \rangle (H_0)_{qp}^{-1} \langle D_p | H | 0 \rangle \\ &= -\mathbf{I} \cdot \mathbf{H}_0^{-1} \cdot \mathbf{H} \cdot \mathbf{H}_0^{-1} \cdot \mathbf{I}, \end{aligned} \quad (9)$$

where

$$(H_0)_{nm}^{-1} = \langle D | H_0 - E_0 | D \rangle_{nm}^{-1}, \quad (10)$$

$D_m, D_n$  denote double replacement determinants.

Direct differentiation of Eq. (9) yields

$$\begin{aligned} \frac{dE_3}{dx} = E_3^{(x)} &= -2\mathbf{I}^{(x)} \cdot \mathbf{H}_0^{-1} \cdot \mathbf{H} \cdot \mathbf{H}_0^{-1} \cdot \mathbf{I} \\ &\quad - \mathbf{I} \cdot \mathbf{H}_0^{-1} \cdot \mathbf{H}^{(x)} \cdot \mathbf{H}_0^{-1} \cdot \mathbf{I} \\ &\quad + 2\mathbf{I} \cdot \mathbf{H}_0^{-1} \cdot \mathbf{H}_0^{(x)} \cdot \mathbf{H}_0^{-1} \cdot \mathbf{H} \cdot \mathbf{H}_0^{-1} \cdot \mathbf{I}. \end{aligned} \quad (11)$$

We can now expand Eq. (11) to yield the following expression:

$$\begin{aligned} E_3^{(x)} &= -2 \sum_{ijab} (ia|jb)^{(x)} p_{ij}^{ab} - \sum_{abcd} (ac|bd)^{(x)} t_{ij}^{ab} t_{ij}^{cd} \\ &\quad - \sum_{klij} (ki|lj)^{(x)} t_{ij}^{ab} t_{kl}^{ab} \\ &\quad - \sum_{ikac} t_{ij}^{ab} t_{kj}^{cb} [(ia|kc)^{(x)} - (ac|ki)^{(x)}] \\ &\quad + \sum_{ac} Y_{ac} \epsilon_{ac}^{(x)} - \sum_{ik} Y_{ik} \epsilon_{ik}^{(x)}, \end{aligned} \quad (12)$$

where

$$p_{ij}^{ab} = 2r_{ij}^{ab}/D_{ij}^{ab}, \quad (13)$$

$$r_{ij}^{ab} = \frac{1}{8} \sum_{cd} t_{ij}^{cd} (ab || cd) + \frac{1}{8} \sum_{kl} t_{kl}^{ab} (kl || ij) + \frac{1}{4} \sum_{kc} [t_{ik}^{ac} (bk || jc) + t_{ik}^{cb} (ak || jc) + t_{kj}^{ab} (bk || ic) + t_{kj}^{cd} (ak || ic)], \quad (14)$$

$$Y_{ac} = 4 \sum_{ijb} r_{ij}^{cb} t_{ij}^{ab} / D_{ij}^{cb}, \quad (15)$$

$$Y_{ik} = 4 \sum_{jab} r_{kj}^{ab} t_{ij}^{ab} / D_{kj}^{ab},$$

and

$$t_{ij}^{ab} = [(ia | jb) - (ib | ja)] / D_{ij}^{ab}, \quad (16)$$

$$D_{ij}^{ab} = \epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j.$$

On substitution of the explicit form for the differentiated orbitals, one obtains

$$E_3^{(x)} = \sum_{ai} U_{ai}^x L_{ai}^{\text{MP3}} + W^x, \quad (17)$$

where  $W^x$  involves derivative one and two electron integrals and the derivative overlap matrix. The "Lagrangian"  $L$  is given by

$$L_{ai}^{\text{MP3}} = \sum_{bc} Y_{bc} A_{bcai} - \sum_{jk} Y_{jk} A_{jkai} - 4 \sum_{jb} \left[ \sum_c p_{ij}^{cb} (ac | jb) + \sum_k p_{kj}^{ab} (ki | jb) \right] + 4 \sum_{cbd} f_{ab}^{cd} (ic | bd) - 4 \sum_{jkl} f_{ij}^{kl} (ka | lj) + 2 \sum_{kc} \left[ \sum_j f_{jk}^{ac} \{ (ij | kc) - (ic | kj) \} - \sum_b f_{ik}^{bc} \{ (ba | kc) - (bc | ka) \} \right], \quad (18)$$

where

$$f_{ab}^{cd} = \frac{1}{4} \sum_{kl} t_{kl}^{ab} t_{kl}^{cd}, \quad f_{ij}^{kl} = \frac{1}{4} \sum_{cd} t_{ij}^{cd} t_{kl}^{cd}, \quad (19)$$

$$f_{jk}^{ac} = \sum_{dl} t_{jl}^{ad} t_{kl}^{cd}.$$

The evaluation of  $L$  can be arranged in terms of matrix multiplication operations, hence the code can be vectorized, making the evaluation highly efficient. This is the reason for a full 4-index transformation, so that  $r_{ij}^{ab}$  could be efficiently processed. Within the unrestricted framework, the spin-orbital formula (12) can be separated into its space-spin components, using

$$(t_{ij}^{ab})^{\alpha\alpha} = \frac{[(ia | jb)^{\alpha\alpha} - (ib | ja)^{\alpha\alpha}]}{\epsilon_a^\alpha + \epsilon_b^\alpha - \epsilon_i^\alpha - \epsilon_j^\alpha}, \quad (20)$$

$$(t_{ij}^{ab})^{\alpha\beta} = \frac{(ia | jb)^{\alpha\beta}}{\epsilon_a^\alpha + \epsilon_b^\beta - \epsilon_i^\alpha - \epsilon_j^\beta}$$

and similarly for the  $\beta$  terms.

The result is a formula

$$E_3^{(x)} = \sum_{ai}^{\alpha} L_{ai}^{\alpha} U_{ai}^{\alpha x} + \sum_{ai}^{\beta} L_{ai}^{\beta} U_{ai}^{\beta x} + W^x, \quad (21)$$

where  $L_{ai}^{\alpha}$  and  $L_{ai}^{\beta}$  are the  $\alpha$  and  $\beta$  forms of the Lagrangian matrix (18).

The final form for the MP3 gradient is obtained from Eq. (21) through application of the idea of Handy and Schaefer,<sup>18</sup> viz.,

$$L^T U^x = L^T (H^{-1} \cdot b^x) = [(H^{-1})^T \cdot L]^T \cdot b^x = Z^T \cdot b^x, \quad (22)$$

where

$$H^T \cdot Z = L. \quad (23)$$

In unrestricted terms this becomes

$$\begin{pmatrix} H^{\alpha\alpha} & H^{\alpha\beta} \\ H^{\beta\alpha} & H^{\beta\beta} \end{pmatrix} \begin{pmatrix} Z^{\alpha} \\ Z^{\beta} \end{pmatrix} = \begin{pmatrix} L^{\alpha} \\ L^{\beta} \end{pmatrix}. \quad (24)$$

This only requires the solution of one set of equations to obtain  $Z$ , instead of the  $3N$  sets, since all the matrices involved are independent of the nuclear perturbations. The UMP3 gradient can therefore be rewritten as

$$E_3^{(x)} = \sum_{ai}^{\alpha} Z_{ai}^{\alpha} b_{ai}^{\alpha x} + \sum_{ai}^{\beta} Z_{ai}^{\beta} b_{ai}^{\beta x} + W^x. \quad (25)$$

The importance of this is that the derivative integrals involved with the  $b^x$  matrix do not have to be stored or transformed. After collecting terms in Eq. (25), the gradient expression can be written in terms of one and two particle coefficient matrices ("effective density matrices") running over the molecular orbitals. In terms of spin orbitals this yields

$$E_3^{(x)} = \sum_{pq} g_{pq} S_{pq}^x + \sum_{pq} \gamma_{pq} h_{pq}^x + \sum_{pqrs} \Gamma_{pqrs} (pq | rs)^x. \quad (26)$$

The one and two particle coefficient matrices can be defined from

$$\begin{aligned} & \sum_{bc} Y_{bc} h_{bc}^x - \sum_{jk} Y_{jk} h_{jk}^x - \sum_{ai} Z_{ai} h_{ai}^x \\ & \equiv \sum_{pq} \gamma_{pq} h_{pq}^x - 2 \sum_{ijab} p_{ij}^{ab} (ia | jb)^x - \sum_{abcd} f_{ab}^{cd} (ac | bd)^x \\ & \quad - \sum_{ijkl} f_{ij}^{kl} (ki | lj)^x - \sum_{ikac} f_{ik}^{ac} [(ia | kc)^x - (ac | ki)^x] \\ & \quad + \sum_{pq} \gamma_{pq} \sum_j [(pq | jj)^x - (pj | qj)^x] \\ & \equiv \sum_{pqrs} \Gamma_{pqrs} (pq | rs)^x, \end{aligned} \quad (27)$$

$$\begin{aligned}
& 2 \sum_{ijab} \left[ \sum_k p_{ij}^{ab}(ka|jb) S_{ki}^x + \sum_c p_{ij}^{ab}(ic|jb) S_{ac}^x + 2 \sum_k p_{kj}^{ab}(ki|jb) S_{ai}^x \right] + 2 \sum_{abcd} \left[ \sum_e f_{ab}^{cd}(ac|bd) S_{ea}^x + 2 \sum_i f_{ab}^{cd}(ic|bd) S_{ai}^x \right] \\
& + 2 \sum_{ijkl} \sum_m f_{ij}^{kl}(km|lj) S_{mi}^x + \sum_{ikac} \left[ \sum_b f_{ik}^{ac}\{(bi|kc) - (bc|ki)\} S_{ba}^x + \sum_j f_{ik}^{ac}\{(ja|kc) - (ac|kj)\} S_{ji}^x \right. \\
& \left. + 2 \sum_j f_{jk}^{ac}\{(ij|kc) - (ic|kj)\} S_{ai}^x \right] - \frac{1}{2} \sum_{pqkj} \gamma_{pq} A_{pqkj} S_{kj}^x - \frac{1}{2} \sum_{pq} \gamma_{pq} (\epsilon_q + \epsilon_p) S_{pq}^x \equiv \sum_{pq} g_{pq} S_{pq}^x. \quad (29)
\end{aligned}$$

Backtransforming the density matrices to the AO basis yields the final form for the UMP3 gradient:

$$E_3^{(x)} = \sum_{\mu\nu} g_{\mu\nu} S_{\mu\nu}^x + \sum_{\mu\nu} \gamma_{\mu\nu} h_{\mu\nu}^x + \sum_{\mu\nu\rho\sigma} \Gamma_{\mu\nu\rho\sigma} (\mu\nu|\rho\sigma)^x. \quad (30)$$

To summarize our procedure for the evaluation of a UMP3 gradient:

- (1) UHF-SCF calculation.
- (2) Full 4-index transformation of the integrals into the MO basis.
- (3) Construction  $L^\alpha$ ,  $L^\beta$  and the solution of the single set of equations to give  $Z^\alpha$ ,  $Z^\beta$ .
- (4) Construction of the effective density matrices  $g$ ,  $\gamma$ , and  $\Gamma$  (both  $\alpha$  and  $\beta$  forms) in the MO basis.
- (5) These matrices are back transformed to the AO basis;  $\alpha$  and  $\beta$  dependence has now disappeared. They are added to the equivalent matrices from the UHF and UMP2 gradient theory.
- (6) The derivative AO integrals are evaluated, multiplied by the AO density matrices, then summed to give the final result.

As described above, we believe there are advantages in doing a full 4-index transformation, because this enables the  $r$  matrix to be defined, and hence allows the vectorized construction of  $L$ . The integrals are sorted onto external files such that for a particular  $p$  and  $q$ , all integrals of the form  $J_{(rs)\alpha}^{(pq)\alpha} = (pq|rs)^{\alpha\alpha}$ ,  $K_{(rs)\alpha}^{(pq)\alpha} = (pr|qs)^{\alpha\alpha}$ ,  $J_{(rs)\beta}^{(pq)\alpha} = (pq|rs)^{\alpha\beta}$ , and the equivalent  $\beta$  lists, are available.

The results of doing this are:

- (a) It facilitates the use of matrix multiplication on these integral blocks to perform the intermediate summations in the formation of  $L$ .
- (b) It avoids the problem of storing some large vector in core, for example, the complete  $t_{ij}^{ab}$  vector, because it can be formed as required. This obviously becomes significant for large calculations when memory requirements need to be kept to a minimum.

The UMP2 and UMP3 gradient programs use the derivative integrals and general transformation and symmetry facilities of the Cambridge Analytic Derivatives Package, CADPAC. The timings for our current program for the gradients of UHF, UMP2, UMP3 are in the ratio 1:2.5:6.0.

### III. APPLICATIONS

We shall discuss calculations at the RMP2 and RMP3 level on  $H_2O$ ,  $NH_3$ ,  $H_2CO$ , and  $HCN$ , and at the UMP2 and UMP3 level on  $NH_2$ ,  $PH_2$ , and  $AsH_2$ . For these doublet open shell systems  $\langle \hat{S}^2 \rangle_0 = 0.76$ , so spin contamination is

not important. For the RMP2 calculations, the harmonic frequencies are obtained analytically; for all other (RMP3, UMP2, UMP3) calculations, these frequencies are obtained from finite difference calculations using the gradient codes.

The basis sets used for the first row atoms and H are by now fairly standard. The DZP basis involves the 4s2p for first row atoms, or 2s for H, Dunning<sup>19</sup> contraction of the 9s5p or 4s Huzinaga<sup>20</sup> primitive sets, with one set of polarization functions. The TZ2P basis involves the 5s4p for first row atoms, or 3s for H, Dunning contractions of the 10s6p or 5s Huzinaga primitive sets. The larger basis sets are constructed from the 8s6p contraction of van Duijneveldt's 13s8p primitive set,<sup>21</sup> and the 6s contraction of the 10s primitives for H. The polarization function exponents are for H: 1p, 1.0; 2p, 1.5, 0.5; 3p, 1.8, 0.6, 0.2. For C 1d 0.8; 2d, 1.2, 0.4; 3d, 1.8, 0.6, 0.2; 1f, 1.0. For N 1d 0.8; 2d, 1.35, 0.45; 3d, 2.4, 0.8, 0.266; 1f, 1.0. For O 1d 0.9; 2d, 1.35, 0.45; 3d, 2.7, 0.9, 0.3; 1f 1.0.

For phosphorus, the DZP basis involves the Dunning and Hay<sup>22</sup> 6s4p contraction of the 11s7p Huzinaga<sup>23</sup> primitive set. The TZ2P basis involves the 9s6p contraction of the 12s9p McClean and Chandler<sup>24</sup> primitive set. The 11s8p contraction is built from the 9s6p contraction by adding more diffuse  $s$  and  $p$  functions formed by following the approximate geometrical progression of the exponents. The polarization function exponents are 1d, 0.5; 2d, 0.75, 0.25; 3d, 0.85, 0.3, 0.1; 1f, 0.6.

The arsenic calculations use the DZP basis which involves the 10s6p2d contraction of a Dunning 13s9p5d primitive set. The TZ2P basis involves the 13s9p3d contraction of a Dunning 20s14p6d primitive set.<sup>25</sup> The exponents of the  $f$  polarization functions are 1f, 0.5; 2f, 0.75, 0.25.

#### A. $NH_2$

Previous studies of  $NH_2$  using unrestricted techniques have generally not used large basis sets. For example, Pople and co-workers<sup>2</sup> have published optimized geometries of  $NH_2$  at UHF, UMP2, UMP3, and UMP4 levels of theory using the 6-31G\* basis set and also at UMP4 with a 6-311G\*\* basis set. Their results are in reasonable agreement with experiment, but the studies presented here show the effect of increasing the basis set size and also calculate the corresponding harmonic vibrational frequencies.

The molecular properties of  $NH_2$  at various levels of theory using a selection of basis sets are given in Table I. The RHF and UHF results using a large 8s6p4d 1f/6s3p basis set (close to the HF basis set limit) are very similar and display the typical errors for SCF optimized geometries. The NH bond length,  $r_{NH}$ , is 0.016 Å too short and the angle consid-

TABLE I. SCF, UMP2, and UMP3 optimized geometries and harmonic frequencies for NH<sub>2</sub>.

Method	Basis set	$r_{\text{NH}}$	$\theta_{\text{HNH}}$	$\omega_1$	$\omega_2$	$\omega_3$
RHF	8s6p4d 1f/6s3p	1.008	104.8	3620	1643	3708
UHF	8s6p4d 1f/6s3p	1.009	104.8	3605	1639	3693
UMP2	6-31G**	1.029	103.3	...	...	...
	DZP	1.026	103.3	3492	1569	3620
	5s4p2d/3s2p	1.020	103.3	3465	1555	3571
	8s6p3d/6s2p	1.019	103.3	3454	1555	3558
	8s6p4d 1f/6s3p	1.020	103.4	3455	1544	3563
UMP3	6-31G**	1.031	103.2	...	...	...
	DZP	1.027	103.2	3477	1576	3592
	5s4p2d/3s2p	1.020	103.2	3453	1565	3547
	8s6p3d/6s2p	1.019	103.2	3443	1561	3535
	8s6p4d 1f/6s3p	1.020	103.4	3447	1556	3542
UMP4	6-31G**	1.034	102.9	...	...	...
	6-311G**	1.030	101.6	...	...	...
CISD	DZP <sup>b</sup>	1.029	103.1	...	...	...
Expt. <sup>c</sup>		1.024	103.3	...	...	...

<sup>a</sup> From Ref. 2.<sup>b</sup> From Ref. 26.<sup>c</sup> From Ref. 34.

erably too large by 1.5° compared to the experimental values. At the UMP2 level, the 6-31G\* geometries are much improved over UHF (the bond length is now in error by + 0.005 Å), with the DZP basis the bond length is reduced to an error of + 0.002 Å, and with the large basis, 8s6p4d 1f/6s3p, the bond is 0.004 Å too short, with bond angle being correct to within 0.1°. At the UMP3 level the geometries are very similar to the equivalent UMP2 values with the same basis set. It is also noted that the CISD geometry<sup>26</sup> with a DZP basis set is similar to the UMP3 DZP geometry. The UMP4 6-31G\* geometry<sup>2</sup> is different to that at UMP2 6-31G\* in such a way that if this difference is carried over to large basis sets the UMP4 geometry will almost be coincident with experiment.

Experimental harmonic frequencies are not available, however from previous calculations, SCF harmonic frequencies close to the basis set limit are typically 10% greater than the "true" harmonic values. It can be seen from Table I that the UMP2 DZP frequencies are significantly lower than the UHF values calculated close to the basis set limit. The UMP2 values reduce as the basis set increases and it is expected that large basis set values will be good estimates of the harmonic frequencies. We base this on our experience with RMP2.<sup>4</sup> The UMP3 harmonic frequencies are similar to the UMP2 values using the same basis set. There is a small reduction in the symmetric and asymmetric stretching modes  $\omega_1$  and  $\omega_3$  at UMP3 compared to UMP2, while the  $\omega_2$  bending frequency is increased by an even smaller magnitude. For example, employing an 8s6p3d/6s2p quality basis set,  $\omega_1$  and  $\omega_3$  are reduced by 11 and 23 cm<sup>-1</sup>, respectively, and  $\omega_2$  is increased by 6 cm<sup>-1</sup>. For the case of NH<sub>2</sub> the results predicted at the UMP3 level of theory do not display any significant differences compared to those predicted by UMP2.

TABLE II. SCF, UMP2, and UMP3 optimized geometries and harmonic frequencies for PH<sub>2</sub>.

Method	Basis set	$r_{\text{PH}}$	$\theta_{\text{HPH}}$	$\omega_1$	$\omega_2$	$\omega_3$
RHF	11s8p3d 1f/6s3p	1.409	93.7	2517	1223	2517
UHF	11s8p3d 1f/6s3p	1.409	93.7	2511	1220	2513
UMP2	DZP	1.411	92.3	2502	1188	2513
	9s6p2d/3s2p	1.412	92.1	2480	1150	2489
	11s8p2d/6s2p	1.411	92.0	2471	1142	2480
	11s8p3d/6s3p	1.411	92.0	2461	1154	2468
	11s8p3d 1f/6s3p	1.412	91.9	2459	1151	2468
UMP3	DZP	1.413	92.2	2488	1194	2497
	9s6p2d/3s2p	1.414	92.0	2455	1146	2463
	11s8p2d/6s2p	1.414	91.9	...	...	...
	11s8p3d/6s3p	1.414	91.8	...	...	...
	11s8p3d 1f/6s3p	1.415	91.9	...	...	...
Expt. <sup>a</sup>		1.418	91.7	...	...	...

<sup>a</sup> From Ref. 35.

## B. PH<sub>2</sub>

Previous theoretical studies of PH<sub>2</sub><sup>27</sup> have not used large basis sets. The molecular properties of PH<sub>2</sub> at various levels of theory using a selection of basis sets are given in Table II. The large basis set UHF optimized geometry predicts the equilibrium  $r_{\text{PH}}$  to be 0.009 Å too short and  $\theta_{\text{HPH}}$  2.0° too large compared to experiment. At the UMP2 level the values are improved at the DZP level compared to UHF, and using the large 11s8p3d 1f/6s3p basis set  $r_{\text{PH}}$  is too short by 0.006 Å and  $\theta_{\text{HPH}}$  is too large by 0.2°. The UMP3 geometries improve upon the UMP2 geometries using the same basis set. We see that for the largest basis set, 11s8p3d 1f/6s3p, the  $r_{\text{PH}}$  bond length is only 0.003 Å too short.

For the harmonic frequency predictions, we observe that the trends are rather similar to those for NH<sub>2</sub>, there being no significant change between the UMP2 and UMP3 values. We suggest that our best values, from the 9s6p2d/3s2p basis, of 2455, 1146, and 2463 cm<sup>-1</sup> will be very close to the exact harmonic frequencies.

## C. AsH<sub>2</sub>

The molecular properties of AsH<sub>2</sub> are shown in Table III. The equilibrium structures calculated at the RHF and UHF levels are similar for the three basis sets employed. At the UMP2 level, the bond length decreases from 1.520 Å using a DZ basis set to 1.492 Å with the TZ2P quality basis set, the equilibrium bond angle similarly closing up from 92.3° at DZ to 90.6° at TZ2P. The unusual feature of the optimized geometries is that the inclusion of electron correlation has the effect of reducing the bond length. CISD calculations<sup>28</sup> also show the same effect. At the TZ2P level there is a reduction of 0.016 Å from UHF to UMP2. The bond angle also closes by 2.2°, however this follows the trend of NH<sub>2</sub> and PH<sub>2</sub> where there was a 1.5° reduction. For NH<sub>2</sub> there was a bond length increase of 0.01 Å and for PH<sub>2</sub> of 0.003 Å in line with UHF bond length predictions being too short and UMP2 correcting for this. However, in the case of AsH<sub>2</sub> the involvement of a full complement of *d* electrons

TABLE III. Optimized geometries and harmonic frequencies for AsH<sub>3</sub>.

Method	Basis set	$r_{\text{AsH}}$	$\theta_{\text{HAsH}}$	$\omega_1$	$\omega_2$	$\omega_3$
RHF	DZ	1.511	93.1	2297	1126	2291
	DZP	1.503	93.3	2349	1129	2340
	TZ2P	1.508	92.9	2344	1115	2340
UHF	DZ	1.511	93.0	2291	1126	2284
	DZP	1.503	93.2	2348	1131	2338
	TZ2P	1.508	92.8	2346	1125	2341
UMP2	DZ	1.520	92.3	2218	1048	2210
	DZP	1.486	92.2	2371	1068	2371
	TZ2P	1.492	90.6	2293	1145	2367
CISD	DZP	1.488	92.3	2368	1084	2365

has a significant effect. These *d* orbitals cause the bond angle to be closer to 90° as predicted by UMP2 TZ2P.

The harmonic frequencies of AsH<sub>3</sub> show a rather erratic behavior. There is certainly evidence in Table III, that with a TZ2P basis set the RHF and UHF calculations have converged to  $\omega_1$  (symmetric stretch)  $\approx 2340$ ,  $\omega_2$  (bend)  $\approx 1120$ , and  $\omega_3$  (asymmetric stretch)  $\approx 2345$  cm<sup>-1</sup>. Introduction of electron correlation has a varied effect on these frequencies, dependent on the basis set, e.g.,  $\omega_2$  changes by -63 cm<sup>-1</sup> with DZP and by +20 cm<sup>-1</sup> with TZ2P set. Our only option is to suggest that our TZ2P values of 2293, 1145, and 2367 cm<sup>-1</sup> are nearest the truth, but in this instance we do not believe these values have converged with respect to either basis set or method. However, we report these calculations because there appears to be no other calculations on AsH<sub>3</sub> in the literature.

#### D. H<sub>2</sub>O

H<sub>2</sub>O has been studied extensively experimentally and theoretically. Experimental geometries and harmonic frequencies are well known and believed to be highly accurate. A complete list of theoretical H<sub>2</sub>O studies will not be given, but in view of the current investigations reference is given to results predicted by Møller-Plesset perturbation theory<sup>2,29,30</sup> and CI<sup>31</sup> methods.

The molecular properties of H<sub>2</sub>O at various levels of theory using a series of basis sets are given in Table IV. The RMP2 geometrical parameters are well known and have been shown to be highly accurate.<sup>30</sup> At the RMP3 level values predicted via use of a DZP basis set show an equilibrium bond length  $r_{\text{OH}}$  that is closer to experiment than the corresponding RMP2 DZP bond length, and a bond angle  $\theta_{\text{HOH}}$  that is slightly inferior. Increasing the basis set size has the effect of reducing the RMP3 bond length and in the limit of saturating the *s*, *p*, and *d* shells and incorporating an *f* function the RMP3 equilibrium bond length of 0.954 Å is too short by 0.004 Å and the bond angle of 104.9° is too large by 0.4°. This is in considerably poorer agreement with experiment than the RMP2 limiting values of 0.959 Å and 104.4°. It should be noted that the CISD DZP optimized geometry as predicted by Lee *et al.*<sup>31</sup> is similar to the RMP3 DZP geometry. They also performed a CISDTQ DZP calculation resulting in a predicted geometry that is different to the

TABLE IV. Optimized geometries and harmonic frequencies for H<sub>2</sub>O.

Method	Basis set	$r_{\text{OH}}$	$\theta_{\text{HOH}}$	$\omega_1$	$\omega_2$	$\omega_3$
RHF	8s6p4d 2f/6s3p2d <sup>a</sup>	0.940	106.3	4130	1746	4231
RMP2	6-31G <sup>ab</sup>	0.969	104.0	3772	1737	3916
	DZP	0.962	104.5	3913	1665	4059
	5s4p2d/3s2p	0.958	104.5	3859	1641	3985
	8s6p3d/6s2p	0.957	104.5	3842	1628	3968
RMP3	8s6p4d 1f/6s3p	0.959	104.4	3839	1629	3966
	6-31G <sup>ab</sup>	0.967	104.3	...	...	...
	DZP	0.959	104.8	3966	1688	4089
	5s4p2d/3s2p	0.954	104.9	3930	1673	4032
RMP4	8s6p3d/6s2p	0.952	104.9	3918	1662	4019
	8s6p4d 1f/6s3p	0.954	104.9	3916	1664	4017
	6-31G <sup>ab</sup>	0.970	103.9	...	...	...
RMP4	6-311G <sup>ab</sup>	0.959	102.4	...	...	...
RMP4	39-STO <sup>c</sup>	0.960	104.4	3823	1674	3936
CISD	DZP <sup>d</sup>	0.959	104.8	3959	1690	4082
CISD	TZ2P <sup>d</sup>	0.961	104.6	3901	1668	4006
CISDTQ	DZP <sup>d</sup>	0.964	104.4	3887	1674	4017
Expt. <sup>e</sup>		0.958	104.5	3832	1649	3942

<sup>a</sup> From Ref. 37.

<sup>b</sup> From Ref. 2.

<sup>c</sup> From Ref. 29.

<sup>d</sup> From Ref. 31.

<sup>e</sup> From Ref. 36.

RMP3 geometry; the bond length is 0.005 Å longer and the angle is reduced by 0.4°. Although the DZP basis set does not contain sufficient functions to allow definite conclusions to be made, this does suggest that higher order excitations are important. Indeed Pople's RMP4 6-31G\* geometry<sup>2</sup> is different from the RMP2 6-31G\* geometry such that if this is carried over to large basis sets the RMP4 optimized geometry would be almost coincident with experiment. Unfortunately, there are not enough large basis set MP4 calculations in the literature to be certain that this is valid. However the RMP4 calculation performed by Bartlett *et al.*<sup>29</sup> employing a 39 Slater type orbital basis gives an  $r_{\text{OH}}$  of 0.960 Å and a  $\theta_e$  of 104.4° which is a significant improvement on any of the RMP3 results.

Turning now to the harmonic frequencies the RMP2 values using the largest basis set are in error compared to experiment by +7 (0.2%), -20 (1.2%), and +24 cm<sup>-1</sup> (0.6%) for the symmetric stretch  $\omega_1$ , bend  $\omega_2$ , and asymmetric stretch  $\omega_3$ . Pulay suggests that the small errors in the stretching frequencies are a reflection of the highly accurate geometry.<sup>32</sup> The RMP3 harmonic frequencies are higher and in greater error compared to experiment than the RMP2 frequencies with the same basis set. For example using the 8s6p3d/6s3p basis set the RMP3 frequencies are in error by 86 (2.2%), 13 (0.8%), and 77 cm<sup>-1</sup> (2.0%) compared to RMP2 errors of 10 (+0.3%), -19 (-1.2%), and 26 cm<sup>-1</sup> (+0.7%). It could be argued that this reflects the poorer MP3 geometry. The CISD DZP frequencies are similar to the RMP3 DZP values and the CISDTQ frequencies at the DZP level are a significant improvement on either of these, suggesting once again the importance of higher order excitations. Bartlett's full MP4<sup>29</sup> calculation using the 39

TABLE V. Optimized geometries for NH<sub>3</sub>.

Method	Basis set	$r_{\text{NH}}$	$\theta_{\text{HNH}}$
RHF	8s6p4d 1f/6s3p1d <sup>a</sup>	0.998	108.1
RMP2	6-31G <sup>ab</sup>	1.017	106.3
	DZP	1.014	106.7
	5s4p2d/3s2p	1.009	107.2
	8s6p3d/6s2p	1.008	107.1
	13s8p2d/8s2p1d	1.007	107.3
RMP3	6-31G <sup>ac</sup>	1.017	106.2
	DZP	1.013	106.6
	5s4p2d/3s2p	1.006	107.1
	8s6p3d/6s2p	1.006	107.0
RMP4	6-31G <sup>ac</sup>	1.021	105.8
Expt. <sup>d</sup>		1.012	106.7

<sup>a</sup>From Ref. 37.<sup>b</sup>From Ref. 30.<sup>c</sup>From Ref. 2.<sup>d</sup>From Ref. 38.

Slater type orbital basis yields the most accurate harmonic frequencies in error compared to experiment by  $-9$  ( $-0.2\%$ ),  $+25$  ( $+1.5\%$ ), and  $-6$   $\text{cm}^{-1}$  ( $+0.2\%$ ). This investigation suggests that RMP3 yields poorer results than RMP2 and for highly accurate results there is a need to include the effects of higher order excitations, for example in a full RMP4 calculation. Of course, this conclusion is not definitive since it is based on limited MP4 data.

### E. NH<sub>3</sub>

The geometrical predictions for NH<sub>3</sub> in Table V show a very similar pattern to those for H<sub>2</sub>O. On going from RMP2 to RMP3 with the larger basis sets, the bond length is reduced by  $\approx 0.003$  Å, so that it is in error by  $-0.006$  Å. The RMP4 results indicate that if it was possible to perform RMP4 with the large basis sets, the predictions would have high accuracy.

### F. H<sub>2</sub>CO

This molecule is interesting in so far as there is a considerable change in the predicted CO bend length when com-

TABLE VI. Optimized geometries for H<sub>2</sub>CO.

Method	Basis set	$r_{\text{CO}}$	$r_{\text{CH}}$	$\theta_{\text{HCH}}$
RHF	5s4p2d/3s2p <sup>a</sup>	1.178	1.092	116.1
RMP2	6-31G <sup>ab</sup>	1.221	1.104	115.6
	DZP	1.223	1.102	116.2
	5s4p2d/3s2p	1.211	1.097	116.5
	8s6p3d/6s2p	1.210	1.096	116.6
	6-31G <sup>ab</sup>	1.210	1.104	116.0
RMP3	DZP	1.211	1.101	116.4
	5s4p2d/3s2p	1.198	1.097	116.5
Expt. <sup>c</sup>		1.203	1.099	116.5

<sup>a</sup>From Ref. 30.<sup>b</sup>From Ref. 2.<sup>c</sup>From Ref. 39.

TABLE VII. Optimized geometries and harmonic frequencies for HCN.

Method	Basis set	$r_{\text{CN}}$	$r_{\text{CH}}$	$\omega_1$	$\omega_2$	$\omega_3$
RHF	5s4p2d/3s2p <sup>a</sup>	1.124	1.057	3600	869	2408
RMP2	6-31G <sup>ab</sup>	1.177	1.070	3517	703	2038
	DZP	1.184	1.069	3506	719	1999
	5s4p2d/3s2p	1.164	1.064	3461	668	2033
	8s6p3d/6s2p	1.162	1.063	3461	642	2030
	5s4p2d 1f/3s2p	1.164	1.065	3473	738	2043
RMP3	6-31G <sup>ab</sup>	1.158	1.067	...	...	...
	DZP	1.163	1.067	3539	750	2236
	5s4p2d/3s2p	1.144	1.063	3488	694	2260
	8s6p3d/6s2p	1.143	1.062	...	...	...
	5s4p2d 1f/3s2p	1.144	1.062	...	...	...
RMP4	DZP	1.182	1.071	...	...	...
	TZ2P	1.165	1.067	...	...	...
CISD	DZP	1.165	1.068	3541	755	2215
Expt. <sup>c</sup>		1.153	1.065	3442	727	2129

<sup>a</sup>From Ref. 40.<sup>b</sup>From Ref. 2.<sup>c</sup>Geometry from Ref. 41 and frequencies from Ref. 42.

paring RMP2 and RMP3. The best RMP2 values for  $r_e$  are  $\approx 0.008$  Å too long, but our equivalent RMP3 calculations in Table VI indicate a bond length which is  $\approx 0.005$  Å too short. This change is similar to that observed with the 6-31G<sup>\*</sup> basis set.

### G. HCN

This behavior for multiple bonded systems is again reflected for HCN. We see in Table VII that the CN bond is  $\approx 0.01$  Å too long for RMP2 with large basis sets, but for RMP3 it is an equivalent amount too short. We have also performed RMP4 calculations, using the GAUSSIAN 82<sup>33</sup> program. The RMP4 predicted value of  $r_e$  is close to the RMP2 value using the TZ2P basis set and is  $0.012$  Å longer than the experimental value! These numbers display an erratic convergence of the bond length in HCN as the level of perturbation expansion increases from RMP2 through RMP3 to RMP4. This is reflected in the CN stretching frequency predictions: with a TZ2P basis the RMP2, RMP3 and experimental values are 2033, 2260, and  $2129$   $\text{cm}^{-1}$ . The behavior for  $\omega_1$  (CH) and  $\omega_3$  (bend) are similar to those found and discussed above for H<sub>2</sub>O.

### H. C<sub>2</sub>H<sub>2</sub>

This is a well-studied molecule, and only a selection of results are included in Table VIII. The CC bond length is reduced, as now expected, by  $0.012$  Å on going from RMP2 to RMP3 with the TZ2P basis set, with the result that the RMP3 value is  $0.007$  Å too short. It is noted that more sophisticated methods, such as CPF, give a much better prediction.

It has recently been recognized that there is an anomalous behavior of the  $\omega_4$  ( $\pi_g$ ) vibration, with respect to basis set.<sup>4</sup> It is worthwhile commenting on the changes in the frequencies as one proceeds from RMP2 to RMP3. The greatest change, as expected, is in the  $\omega_2$  (CC) vibration, from



TABLE VIII. Optimized geometries and harmonic frequencies for C<sub>2</sub>H<sub>2</sub>.

Method	Basis set	$r_{CC}$	$r_{CH}$	$\omega_1$	$\omega_2$	$\omega_3$	$\omega_4$	$\omega_5$
RHF	8s6p3d 2f/6s3p1d <sup>a</sup>	1.179	1.059	3665	2210	3556	814	868
RMP2	6-31G* <sup>b</sup>	1.218	1.066	3593	2006	3516	444	783
	DZP	1.226	1.067	3555	1956	3470	560	732
	5s4p2d/3s2p	1.208	1.060	3528	1984	3438	523	723
	8s6p3d 1f(=0.8)/6s3p	1.207	1.061	3537	1982	3447	588	746
RMP3	6-31G* <sup>b</sup>	1.206	1.066	...	...	...	...	...
	DZP	1.214	1.066	3578	2074	3482	605	751
	5s4p2d/3s2p	1.196	1.060	3547	2096	3446	566	745
CPF	DZP <sup>a</sup>	1.222	1.069	3533	2006	3443	547	719
	5s4p2d/3s2p	1.204	1.063	3498	2028	3404	473	706
	5s4p2d 1f/3s2p	1.203	1.063	3522	2038	3429	606	754
Expt. <sup>c</sup>		1.203	1.062	3497	2011	3415	624	747

<sup>a</sup> From Ref. 4.<sup>b</sup> From Ref. 2.<sup>c</sup> From Ref. 43.

1984 to 2096 cm<sup>-1</sup>, with the TZ2P basis, compared to the experimental value of 2011 cm<sup>-1</sup>. The next greatest change is in the  $\omega_4$  ( $\pi_g$ ) vibration, from 523 to 566 cm<sup>-1</sup>, compared to the experimental value of 624 cm<sup>-1</sup>. However, we see that the corresponding value with the CPF method is 473 cm<sup>-1</sup>.<sup>4</sup> Inclusion of an  $f$  function in the basis set has been shown to being the RMP2 and CPF values of  $\omega_4$  into good agreement with experiment,<sup>4</sup> suggesting that the corresponding RMP3 value would then be too high. Other frequency changes on proceeding from RMP2 to RMP3 are not significant.

#### IV. DISCUSSION

All these calculations using the RMP2, UMP2, RMP3, and UMP3 methods with TZ2P quality basis sets allow us to make some tentative conclusions on the value of such calculations.

(i) Our study on the open shell system NH<sub>2</sub> shows very little change in either geometrical parameters or frequencies on proceeding from UMP2 to UMP3. For PH<sub>2</sub> the bond length prediction is improved by  $\approx 0.003$  Å at UMP3, with a consequential change in the frequencies of  $\approx 25$  cm<sup>-1</sup>. Our conclusion has to be that some improvement is achieved with these UMP3 calculations, but it is questionable whether the level of increased accuracy is in proportion to the cost of such calculations. UMP2 with a TZ2P basis seems a reliable method (provided there is no spin contamination!).

(ii) Our study on H<sub>2</sub>O reflects one of the major difficulties of MP calculations. Our earlier comments have discussed the strange convergence characteristics of the energy series MP2, MP3, MP4 for H<sub>2</sub>O, and it appears that this behavior is present for geometrical parameters and molecular frequencies. The TZ2P, RMP3, bond length is in error by  $-0.004$  Å, compared to the  $+0.001$  Å RMP2 error. The frequency errors for RMP3 are  $+98$ ,  $+24$ ,  $+90$  cm<sup>-1</sup> compared to the RMP2 errors of  $+27$ ,  $-8$ ,  $+43$  cm<sup>-1</sup> with the TZ2P basis. Clearly such calculations at the RMP3 level are not as successful as those at the RMP2 level. These comments are supported by the calculations on NH<sub>3</sub>.

(iii) There is a considerable difference in the bond length predictions, for RMP2 and RMP3, for multiple bonds. With the TZ2P basis set the respective errors in the predictions are CO (in H<sub>2</sub>CO):  $+0.008$ ,  $-0.005$  Å; CN (in HCN);  $+0.011$ ;  $-0.009$  Å; CC (in C<sub>2</sub>H<sub>2</sub>):  $+0.005$ ,  $-0.007$  Å. In other words, there is no improvement in the predictive value on proceeding from RMP2 to RMP3, if large basis sets are used. Note that this is in contrast to the general conclusion of Pople and co-workers,<sup>2</sup> who observe that with a 6-31G\* basis, for most molecules MP3 bond lengths are closer to the experimental values than are those at either Hartree-Fock or MP2.

So we are forced from the evidence of these calculations to conclude that the predictive powers of the MP3 approach are no better, and often inferior to those of MP2, if large basis sets are used. Also, MP3 is more expensive than MP2 and it has proved possible to program analytic MP2 second derivatives, whereas analytic MP3 second derivatives would be very difficult to program, and the results here definitely show that it is not worthwhile.

The remaining part of the discussion must be one of philosophy. All our numerical evidence in earlier publications<sup>12</sup> is that the MP method is convergent, that is the energy series MP2, MP3, MP4 . . . ultimately converges to the exact eigenvalue of the Schrödinger equation. We must therefore expect properties to converge, both with respect to MP $n$  level and basis set. The RMP $n$  energy series does not converge smoothly, especially at distorted geometries. The UMP $n$  series appears to converge smoothly, but often slowly because of spin contamination. We must therefore not be surprised if there is an erratic convergence of other properties, especially for RMP $n$ .

We believe that it is better to select a method, or level of calculation, and then use as large a basis set as possible with that method. Such an approach appears to remove one uncertainty in the results obtained. The other approach, which is to use a more sophisticated method, with a consequently smaller basis set, leaves a greater uncertainty, especially when it is realized that more sophisticated methods neces-

sarily need larger basis sets for meaningful results. For all these reasons, we continue to believe that the RMP2 and UMP2 methods have much to offer: the method is simple, large basis sets can be used, and the predictive powers on the strength of many reported calculations are very good, although this may be fortuitous. It is admittedly slightly uncomfortable that RMP3 and UMP3 do not provide better predictions, but at the end of the day, quantum chemistry is designed to serve the chemical community, and if there is a good *ab initio* method which needs no scaling or other factors and which is successful for very many systems, then it must be of value.

## ACKNOWLEDGMENTS

I. A. thanks Dr. R. D. Amos and Dr. J. E. Rice for all their help with the interface of these programs to the Cambridge Analytic Derivatives Package, for which they now form a part.

- <sup>1</sup>C. Møller and M. S. Plesset, *Phys. Rev.* **46**, 618 (1934).
- <sup>2</sup>W. J. Hehre, L. Radom, P. v. R. Schleyer, and J. A. Pople, *Ab Initio Molecular Orbital Theory* (Wiley-Interscience, New York, 1986).
- <sup>3</sup>J. A. Pople and R. K. Nesbet, *J. Chem. Phys.* **22**, 571 (1954).
- <sup>4</sup>E. D. Simandiras, J. E. Rice, T. J. Lee, R. D. Amos, and N. C. Handy, *J. Chem. Phys.* **88**, 3187 (1988).
- <sup>5</sup>R. D. Amos and J. E. Rice, Cambridge Analytic Derivatives Package, version 4.0, 1987.
- <sup>6</sup>P. Pulay, G. Fogarasi, G. Pongor, J. E. Boggs, and A. Vargha, *J. Am. Chem. Soc.* **105**, 7037 (1983); H. Sellers, P. Pulay, and J. E. Boggs, *ibid.* **107**, 6487 (1985).
- <sup>7</sup>I. Shavitt, in *Methods of Electronic Structure Theory*, edited by H. F. Schaefer (Plenum, New York, 1977); V. R. Saunders and J. H. van Lenthe, *Mol. Phys.* **48**, 923 (1983).
- <sup>8</sup>W. Meyer, *J. Chem. Phys.* **58**, 1017 (1973); *Int. J. Quantum Chem. Symp.* **5**, 341 (1971); R. Ahlrichs, in *Methods in Computational Molecular Physics*, edited by G. H. F. Diercksen and S. Wilson (Reidel, Dordrecht, 1983).
- <sup>9</sup>R. Ahlrichs, P. Scharf, and C. Erhardt, *J. Chem. Phys.* **82**, 890 (1985); R. Ahlrichs and P. Scharf, *Adv. Chem. Phys.* **67**, 501 (1987).
- <sup>10</sup>R. J. Bartlett and G. D. Purvis, *Int. J. Quantum Chem.* **14**, 561 (1978).
- <sup>11</sup>J. A. Pople, J. S. Binkley, and R. Seeger, *Int. J. Quantum Chem. Symp.* **10**, 1 (1976).
- <sup>12</sup>P. J. Knowles, N. C. Handy, and K. Somasundram, *Theor. Chim. Acta* **68**, 87 (1985).
- <sup>13</sup>J. Gauss and D. Cremer, *Chem. Phys. Lett.* **138**, 131 (1987).
- <sup>14</sup>E. A. Salter, G. W. Trucks, G. Fitzgerald, and R. J. Bartlett, *Chem. Phys. Lett.* **141**, 61 (1987).
- <sup>15</sup>N. C. Handy, R. D. Amos, J. F. Gaw, J. E. Rice, and E. D. Simandiras, *Chem. Phys. Lett.* **120**, 151 (1985).
- <sup>16</sup>E. D. Simandiras, R. D. Amos, and N. C. Handy, *Chem. Phys.* **114**, 9 (1987).
- <sup>17</sup>J. Gerratt and I. M. Mills, *J. Chem. Phys.* **49**, 1719 (1968).
- <sup>18</sup>N. C. Handy and H. F. Schaefer, *J. Chem. Phys.* **81**, 5031 (1984).
- <sup>19</sup>T. H. Dunning, *J. Chem. Phys.* **53**, 2823 (1970); **55**, 716 (1971).
- <sup>20</sup>S. Huzinaga, *J. Chem. Phys.* **42**, 1293 (1965).
- <sup>21</sup>F. B. Van Duijneveldt, IBM publication RJ945 (1971).
- <sup>22</sup>T. H. Dunning and P. J. Hay, *Applications of Electronic Structure Theory*, edited by H. F. Schaefer (Plenum, New York, 1977).
- <sup>23</sup>S. Huzinaga, Approximate Atomic Wavefunctions, Department of Chemistry Technical Report, University of Alberta, Edmonton, Alberta, Canada, 1971.
- <sup>24</sup>A. D. Mclean and G. S. Chandler, *J. Chem. Phys.* **72**, 5639 (1980).
- <sup>25</sup>T. H. Dunning (private communication).
- <sup>26</sup>S. Bell and H. F. Schaefer, *J. Chem. Phys.* **67**, 5173 (1977).
- <sup>27</sup>A. Hinchcliffe and D. G. Bounds, *J. Mol. Struct.* **54**, 231 (1979); J. R. Ball and C. Thomson, *Int. J. Quantum Chem.* **14**, 39 (1978).
- <sup>28</sup>J. E. Rice (private communication).
- <sup>29</sup>R. J. Bartlett, I. Shavitt, and G. D. Purvis, *J. Chem. Phys.* **71**, 281 (1979).
- <sup>30</sup>E. D. Simandiras, N. C. Handy, and R. D. Amos, *Chem. Phys. Lett.* **133**, 324 (1987).
- <sup>31</sup>T. J. Lee, R. B. Remington, Y. Yamaguchi, and H. F. Schaefer, *J. Chem. Phys.* (in press).
- <sup>32</sup>G. Fogarasi and P. Pulay, *Annu. Rev. Phys. Chem.* **35**, 191 (1984).
- <sup>33</sup>J. S. Binkley, M. J. Frisch, D. J. DeFrees, K. Raghavachari, R. A. Whiteside, H. B. Schlegel, E. M. Fluder, and J. A. Pople, Department of Chemistry, Carnegie-Mellon University, Pittsburgh, PA.
- <sup>34</sup>G. Herzberg, *Molecular Structure and Molecular Spectra* (Van Nostrand, Amsterdam, 1966).
- <sup>35</sup>J. M. Berthou, B. Pascat, and H. Guenebaut, *Can. J. Phys.* **50**, 2265 (1972).
- <sup>36</sup>A. R. Hoy, I. M. Mills, and G. Strey, *Mol. Phys.* **24**, 1265 (1972).
- <sup>37</sup>R. D. Amos, *J. Chem. Soc. Faraday Trans. 2* **83**, 1595 (1987).
- <sup>38</sup>W. S. Benedict and E. K. Plyer, *Can. J. Phys.* **79**, 65 (1983).
- <sup>39</sup>K. Yamada, T. Nakagawa, K. Kuchitsu, and Y. Morino, *J. Mol. Spectrosc.* **38**, 70 (1971).
- <sup>40</sup>R. D. Amos, J. F. Gaw, N. C. Handy, E. D. Simandiras, and K. Somasundram, *Theor. Chim. Acta* **71**, 41 (1987).
- <sup>41</sup>G. Winnewisser, A. G. Maki, and D. R. Johnson, *J. Mol. Spectrosc.* **39**, 149 (1971).
- <sup>42</sup>G. Strey and I. M. Mills, *Mol. Phys.* **26**, 129 (1973).
- <sup>43</sup>G. Strey and I. M. Mills, *J. Mol. Spectrosc.* **59**, 103 (1976).