A systematic determination of extended atomic orbital basis sets and application to molecular SCF and MCSCF calculations

by

David Francis Feller

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Department: Chemistry
Major: Physical Chemistry

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1979

TABLE OF CONTENTS

I. INTRODUCTION

Theoretical chemistry seeks as one of its primary goals the ability to compute atomic and molecular properties with sufficient accuracy to be of value in a coordinated exploration of chemical phenomena with the experimentalist. For practical purposes these properties may be divided into two broad categories: those that involve energy or, more likely, energy differences (such as reaction surfaces, excitation energies or ionization potentials) and those that do not (such as dipole and quadrupole moments or electric field components). This work deals primarily with the former.

Discrepancies between the true value of a physical observable and the value predicted by an ab initio calculation result from the need to invoke four main approximations which are necessitated by the intractability of the exact molecular Schroedinger equation: (i) the Born-Oppenheimer approximation which assumes the separability of the electronic and nuclear motions; (ii) finite basis set expansion lengths for the atomic or molecular orbitals; (iii) the single configuration, independent particle approximation associated with algorithms like the Hartree-Fock (HF) self-consistent-field (SCF) technique; (iv) neglect of relativistic effects. The first and last of these are excellent approximations for compounds composed of elements from the first and second rows of the periodic table. The third approximation and various methods

for relaxing this restriction have been well-documented. The essential decision to be made with regard to the third item is whether or not the problem demands the use of more sophisticated techniques than HF-SCF with their substantial increase in computational requirements. On the other hand, the adverse effects of basis set truncations in accordance with the second approximation is often not adequately considered. Ideally one hopes to be able to choose a basis with the minimum flexibility needed for the problem under consideration so that the question one is seeking to resolve can be answered in the most economical way. Such a choice must draw on the body of information concerning basis sets which has been accumulating since the appearance in 1951 of C. C. J. Roothaan's [1] work on the formulation of the Hartree-Fock equations using finite expansions of SCF orbitals in terms of analytical basis functions. The present work deals with the problems associated with determining such finite basis sets for adequate approximations in molecules.

In the first section of this paper we shall deal with the problem of effectively determining the exponential parameters and expansion lengths needed for obtaining any desired deviation from the infinite basis set limit in atomic or molecular calculations. A convenient contracted orbital scheme and polarization expansion lengths will be discussed. In the second section we shall look at a new integral approximation which can facilitate the use of large, flexible basis sets by

employing a small substitute basis to generate the numerous multicenter, electron-repulsion integrals. The final section deals with the three lowest spin states of the HON to HNO isomerization process. The effect of basis set improvement will be studied as it applies to both correlated and uncorrelated wavefunctions. Since the metastable conformation, HON, has never been observed experimentally we shall further investigate a possible bi-molecular, least motion mechanism for converting the metastable form to the stable form.

We hope that this study will further elucidate the relationship between some physical properties and basis sets, thus providing additional guidelines for the individual performing quantum chemical calculations.

II. AN OPTIMAL EVEN-TEMPERED APPROACH TO THE COMPLETE ORBITAL BASIS

A. Justification for Even-tempered Basis Sets

If optimal performance is to be achieved with gaussian basis sets, which are necessarily restricted in size due to computational limitations, then a better understanding of the relationship between such sets and a complete set is needed. This understanding will make possible a more accurate assessment of convergence rates for various expectation values derived from variational calculations.

About 15 years ago Schwartz [2] wrote, "The first essential in talking of convergence rates is to have an orderly plan of procedure. That is, one must choose a set of basis functions to be used and then gradually add more and more of these terms to the variational calculation in some systematic manner." The even-tempered basis introduced by Ruedenberg et al. [3] is ideally suited for this purpose. Even-tempered (ET) gaussian primitives are defined in terms of two optimizable parameters per symmetry by

$$\chi(k\ell m) = N(\zeta_k)\exp(-\zeta_k r^2) r^\ell S_\ell^m(\theta,\phi) \qquad k = 1,2,... \tag{2.1a}$$

$$\zeta_k = \alpha\beta^k \tag{2.1b}$$

While this restriction of the orbital exponents to a geometric sequence results in a small loss of variational

freedom when compared to an independently optimized set, this loss must be seen in perspective. For example, the first row elements might require an additional s-type primitive or two to obtain groundstate energies within a few millihartrees of Huzinaga's [4] completely optimized exponent set for an (8s, 4p) basis, but the difference between the best energy obtainable with this set and the infinite basis set limit is several times larger. Table 1 shows the actual total energies for the (8s,4p) ET and independently optimized atomic basis sets on carbon. All energies in this table and in the remainder of the paper are in atomic units (1 a.u. = 1 hartree = 627.5 Kcal/mole).

Table 1. Total energies for the carbon atom with ET and independently optimized gaussian bases

Basis	Energy	ΔE_∞	ΔE_H
Huzinaga (8s,4p)	-37.6798	0.0088	0.0
ET (8s,4p)	-37.6681	0.0205	0.0117
ET (9s,4p)	-37.6768	0.0118	0.0030
ET (23s,11p)	-37.6886	0.0000	-0.0088

[a] ΔE_∞ is the difference between the energy for this basis and the HF limit energy.

[b] ΔE_H is the difference between the energy of this basis and the Huzinaga (8s,4p) basis result.

For this particular basis the ET choice of exponents has resulted in a 3/1 reduction in the number of nonlinear param-

eters, the independent exponents, which would have to be varied in the optimization procedure. In larger basis sets the savings are even greater.

In this section we shall see to what extent the ET gaussian basis can be viewed as a finite grid for an approximate numerical integration of an exact integral representation of the radial portion of an atomic orbital. Next, the optimal atomic ET parameters for hydrogen, carbon, oxygen, sulfur and selenium as a function of basis set size will be investigated and shown to suggest a simple rule for generating (α,β) for arbitrary expansion lengths. Finally, the optimal molecular ET parameters of several molecules are given as a function of basis set size and a suitable general procedure for arriving at useful molecular basis sets is discussed.

B. Even-tempered Gaussian and Exponential Expansions and Gaussian Integral Transforms for Atomic Orbitals

In order to exhibit the relationship between exact HF-AO's and their approximations in terms of even-tempered gaussian AO's, we note that any atomic orbital of symmetry (ℓ,m) can be expressed through an integral transform over gaussian radial functions in the following manner where $S_\ell^m(\theta,\phi)$ is a normal-

$$\phi_{\ell m}(r,\theta,\phi) = S_\ell^m(\theta,\phi) r^\ell \int_0^\infty d\zeta e^{-\zeta r^2} f_\ell(\zeta) \qquad (2.2)$$

ized spherical harmonic. In the present context is is convenient to write this representation in the form

$$\phi_{\ell m}(r,\theta,\phi) = \int_{-\infty}^{\infty} d(\ell n\zeta)\; g_{\ell m}(\zeta;r,\theta,\phi)\, a_{\ell}(\zeta) \qquad (2.3)$$

where

$$g_{\ell m}(\zeta;r,\theta,\phi) = N_{\ell}\; e^{-\zeta r^2}\; r^{\ell} S_{\ell}^{m}(\theta,\phi) \qquad (2.4a)$$

with $$N_{\ell} = \{\zeta^{2\ell+3}\; 2^{4\ell+7}/\pi[(2\ell+1)!!]^2\}^{1/4} \qquad (2.4b)$$

is a normalized gaussian-type primitive AO. If we consider for $\phi(\ell,m)$ exponential-type AO's of the form

$$\chi(\xi;r,\theta,\phi) = M_{\ell}\; e^{-\xi r}\; r^{\ell} S^{m}(\theta,\phi) \qquad (2.5a)$$

with $$M_{\ell} = \{(2\xi)^{2\ell+3}/(2\ell+2)!\}^{1/2} \qquad (2.5b)$$

then the transform function a_{ℓ} in Equation (2.3) becomes

$$a_{\ell}^{ex}(\xi,\zeta) = [2^{\ell+1}/(\ell+1)!\,\sqrt{\pi/2}\,]^{1/2}(\xi^2/4\zeta)^{(2\ell+5)/4}\; e^{-(\xi^2/4\zeta)} \qquad (2.6)$$

This is a generalization of a formula given by Kikuchi [5] for simple exponentials. Bishop and Somorjai [6] as well as Taylor [7] have also examined transforms of radial functions.

It has been shown by Raffenetti [8] that any HF-SCF AO can be efficiently expanded in terms of even-tempered exponential-type AO's of the form (2.5a),

$$\phi_{\ell m} = \sum_{\nu} b_{\nu} \chi_{\ell m}(\xi_{\nu}), \quad \xi_{\nu} = \hat{\alpha}\hat{\beta}^{\nu} \qquad (2.7)$$

Combining such an expansion with the integral transforms for $\chi_{\ell m}$, we find for $\phi_{\ell m}$ the transform function

$$a_{\ell m}(\xi,\zeta) = \sum_{\nu} b_{\nu}\, a_{\ell}^{ex}(\xi_{\nu},\zeta) \qquad (2.8)$$

where a_{ℓ}^{ex} is given by Equation (2.6). Examples of such HF-AO transforms are shown in Figures 1a, 1b and 1c. Figure 1a corresponds to the (1s) orbital, Figure 1b to the (2s) orbital and Figure 1c to the (2p) orbital of the HF wavefunction of the carbon groundstate. The values for the b_{ν} and ξ_{ν} are taken from Raffenetti's (6s,4p) even-tempered exponential expansion. Since this is an extremely accurate wavefunction (triple zeta in s and double zeta in p) the curves in these three figures can be considered as very close to the gaussian transforms of the exact carbon HF-SCF orbitals.

Approximation of the integral transform (2.3) by means of a numerical integration implies replacement of the integral by a sum over a number of grid points ζ_k (k=1,2,3...). Since it is apparent that the intervals $(\zeta_{k+1}-\zeta_k)$ should increase as ζ_k becomes larger and larger, one reasonable choice of gridpoints is given by the even-tempered exponents introduced in Equation (2.1b), namely $\zeta_k = \alpha\beta^k$. This choice leads to a set of <u>equi-distant</u> gridpoints when $\ln(\zeta)$ is chosen as the integration variable, as has been implied in Equation (2.3). Since the distance between neighboring gridpoints $(\ln\zeta_k)$ is $\ln(\beta)$, the

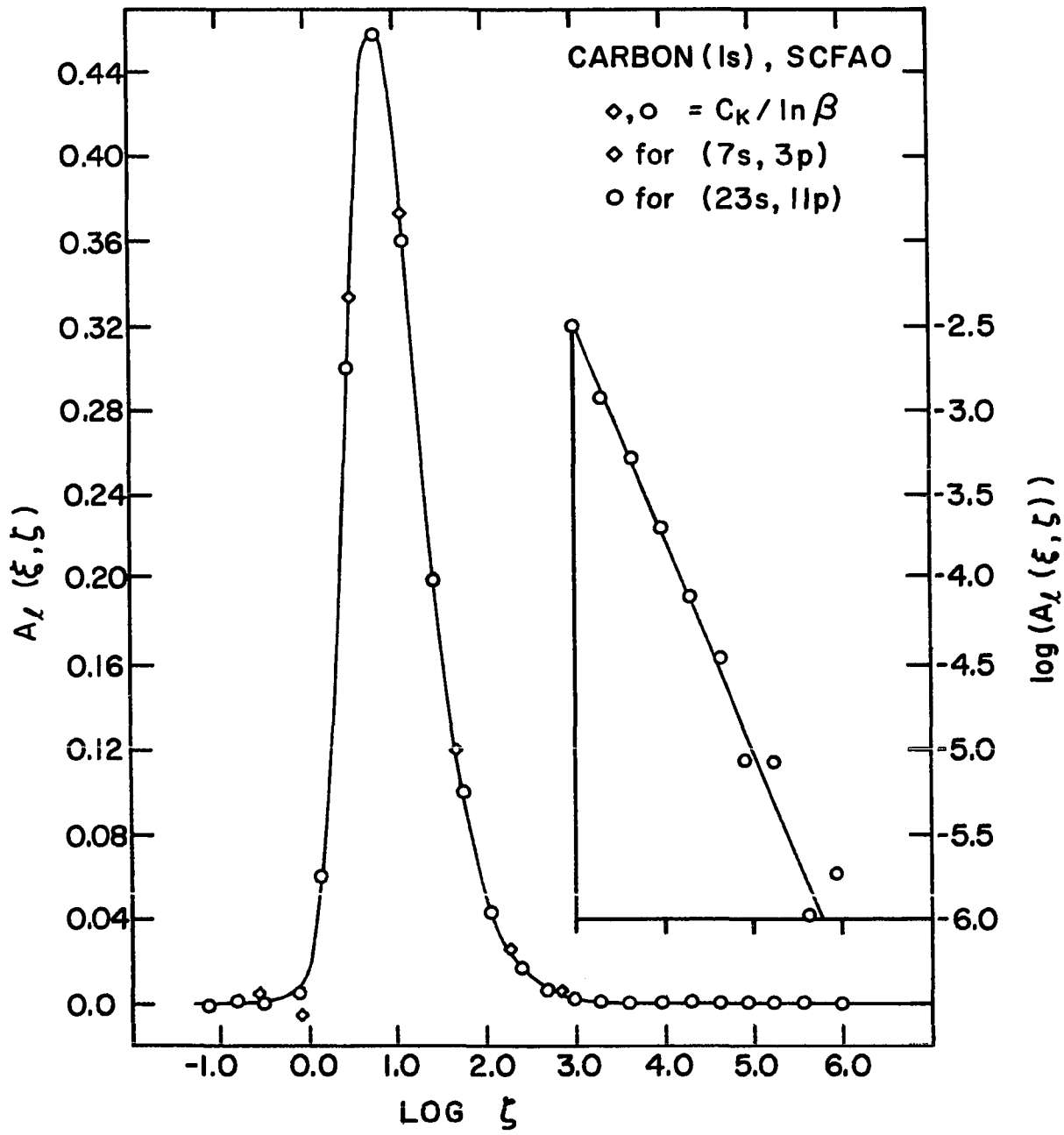


Figure 1a. The integral transform for the carbon(^{3}P) 1s atomic orbital

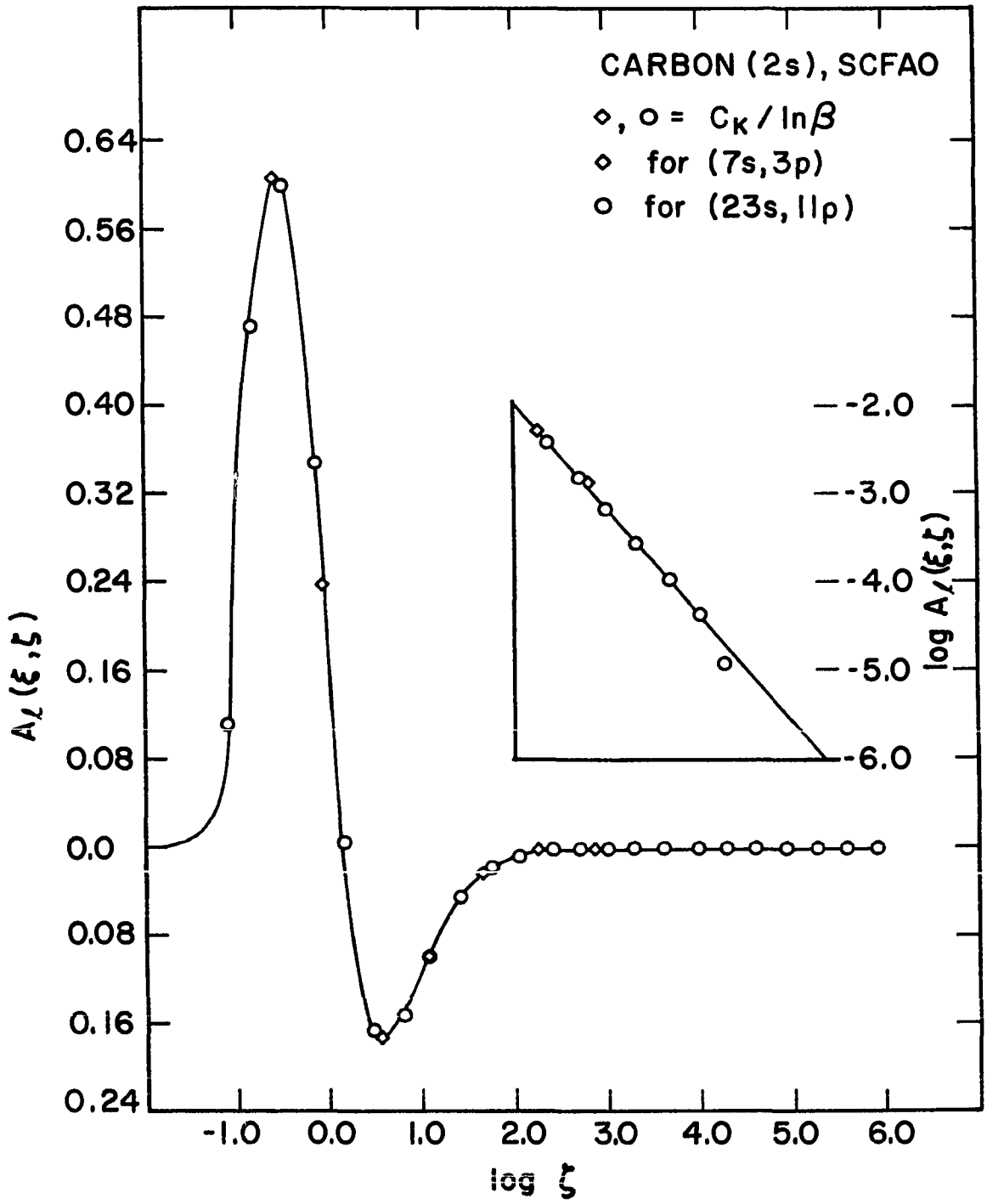


Figure 1b. The integral transform for the carbon(^{3}P) 2s atomic orbital

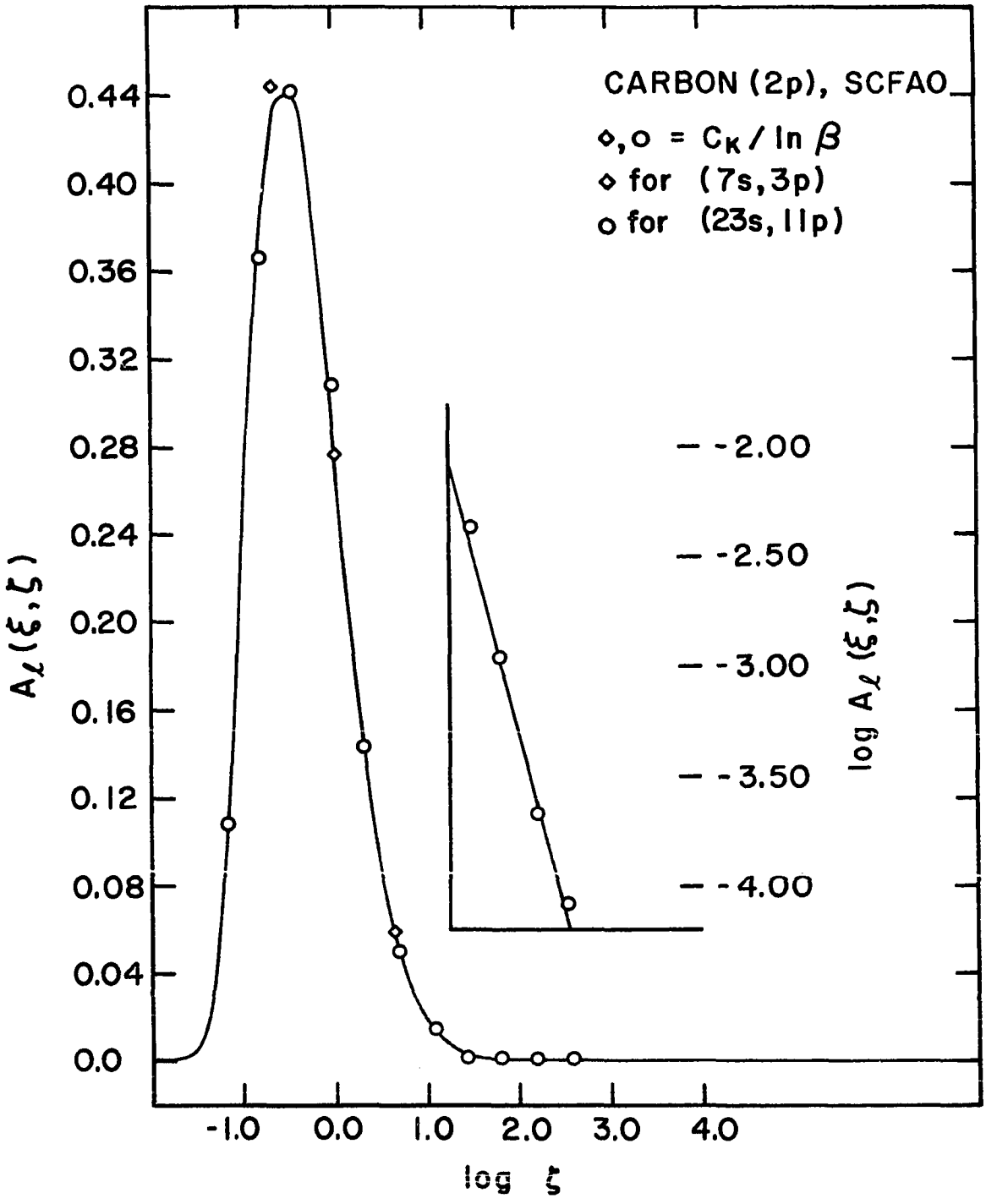


Figure 1c. The integral transform for the carbon(^{3}P) 2p atomic orbital

even-tempered grid approximation to the integral transform (2.3) thus becomes

$$\phi_{\ell m} \approx \sum_k g_{\ell m}(\zeta_k) a_\ell(\zeta_k) \ell n\beta \qquad (2.9)$$

This type of approximation for HF-AO's in terms of $g_{\ell m}$ can be compared to those expansions that result from direct HF-SCF calculations based on expansions of the SCF-AO's in terms of even-tempered gaussian primitives AO's, viz.

$$\tilde{\phi}_{\ell m} = \sum_k g_{\ell m}(\zeta_k)\ c_k^\ell,\ \zeta_k = \alpha_\ell \beta_\ell{}^k \qquad (2.10)$$

where c_k^ℓ as well as the α_ℓ and β_ℓ^k are variationally determined. In view of Equations (2.8), (2.9) and (2.10) one would expect relations like

$$c_k^\ell \approx a(\alpha_\ell \beta_\ell^k) \ell n\beta_\ell \approx \ell n\beta_\ell \sum_\nu b_\nu^\ell\ a_\ell^{ex}(\xi_\nu,\ \zeta_k = \alpha_\ell \beta_\ell^k), \qquad (2.11)$$

in which b_ν^ℓ, ξ_ν come from HF calculations in terms of exponentials, whereas c_k^ℓ, α_ℓ, β_ℓ come from HF calculations in terms of gaussians. The a_ℓ^{ex} are given by Equation (2.6).

As a first example we consider the hydrogen (1s) function $(\xi^3/\pi)^{1/2}\ e^{-\xi r}$. In this case Equation (2.11) simplifies to

$$c_k \approx (8/\pi)^{1/4}\ \ell n\beta (4\alpha\beta^k)^{-5/4}\ e^{-1/4\alpha\beta^k} \qquad (2.12)$$

If the c_k,α,β are determined from a six term even-tempered gaussian expansion (corresponding to an error in the total energy of 0.2 millihartrees), then the agreement of the left and right hand side of (2.12) is better than two significant figures. For an eight term expansion (corresponding to an energy error of 0.01 millihartrees) the agreement is better than four significant figures.

Next we consider the carbon ground state HF-AO's whose integral transforms were shown in Figures 1a-1c. Specifically we choose two sets of parameters with one set resulting from an SCF calculation using a "small" basis of even-tempered gaussian primitives and the second set resulting from an SCF calculation with a "large" basis of even-tempered primitives, the former being a (7s,3p) basis, the latter a (23s,11p) basis. In order to test the degree of validity of the identity (2.11) we simply plot for the three atomic orbitals the values of the quantities $(c_k^\ell/\ln\beta)$ for the appropriate abscissa values of $\zeta_k=\alpha\beta^k$ on the curve for $a_\ell(\zeta_k)$. The (7s,3p) values are entered as diamonds, the (23s,11p) values as circles. It is apparent that not only is the agreement perfect for the large basis, but it is also very good for the small basis. This agreement between the direct variational coefficients c_k^ℓ and the exact transform functions shows that the variational representation in terms of even-tempered gaussian primitives approaches the exact SCF solution in a systematic manner. The integral transform acts as it were a "slidewire" with the coefficients for

finite expansions behaving like beads on the wire. The grid-points of a particular expansion merely determine the positions of the beads on the wire and these positions can be altered without departing from the wire. We have noticed similar behavior for expansions of molecular orbitals.

Another interesting aspect of the discussed results is that they establish a clear relation between the expansion of an atomic orbital in terms of gaussian primitives and its expansion in terms of exponential-type primitives. Equation (2.11) shows how to obtain the coefficients of the even-tempered gaussian expansion when the gaussian and exponential exponents and the exponential coefficients are known. However, it is also possible to invert the process. Since gaussian expansions are always substantially longer than exponential-type expansions of equal quality, it is apparent that the number of coefficients c_k^ℓ for which (2.11) applies is larger than the number of terms in the summation over v. If it is more than twice as large, then there are sufficient equations available to determine the values of the parameters for the exponential expansion when the gaussian parameters c_k^ℓ, α, β, are known. This can be accomplished by a (partly linear, partly nonlinear) least-squares calculation based on minimizing the quantity

$$\sum_k \left[c_k^\ell - f^\ell(b_1^\ell, b_2^\ell, \ldots; \xi_1, \xi_2, \ldots)\right]^2$$

with respect to the parameters b_v^ℓ and ξ_v. Here the f_k^ℓ are the functions defined by the righthand side of equation (2.11) and the values of c_k^ℓ, α, β are supplied by the gaussian expansion. It is thus possible to deduce the complete even-tempered exponential-type expansion from sufficiently large even-tempered gaussian expansions! This procedure works quite well, as the results shown in Table 2 for the carbon 1s and 2s orbitals attest. As for going from the exponential-type to

Table 2. Comparison between the fitting coefficients derived from an ET gaussian basis and the SCF orbital coefficients with an ET exponential basis for carbon[a]

	Fitting Coefficients From a (19s,9p) ETG Basis With a = 0.702 and b = 1.666					
	C(1)	C(2)	C(3)	C(4)	C(5)	C(6)
1s	-0.003	0.010	-0.002	0.922	0.080	0.002
2s	-1.250	-0.060	0.524	0.118	-0.004	0.007
	SCF Orbital Coefficients From a (6s,4p) ETE Basis With a' = 0.705 and b' = 1.667					
	C(1)	C(2)	C(3)	C(4)	C(5)	C(6)
1s	-0.000	0.002	0.016	0.913	0.077	0.001
2s	-1.252	-0.062	0.545	0.092	-0.011	0.001

[a]Both basis sets gave an energy of -37.68859 E_h. The parameters a and b were determined by nonlinear optimization, while a' and b' are the SCF optimized even-tempered exponential values of Raffenetti.

the gaussian expansion, it was possible to predict the coefficients of a 9 term gaussian expansion of the 1s and 2s orbitals of carbon with sufficient accuracy that an energy only 0.009 E_h above the SCF coefficients was produced without actually performing the SCF calculation with gaussians.

Although the shape of the gaussian transform is highly independent of the basis set size, as already shown in Figures 1a-1c, it is also nearly independent of the atomic number. Of course the size of the atomic orbitals decreases as the nuclear charge is increased, resulting in a shift of the transform function to higher log ζ values for larger Z. However, by shifting the transforms of two different elements so that they are superimposed the similarity of the two sets is apparent. In Figures 1d-1f the gaussian transforms for the atomic orbitals of fluorine, shown as solid black dots, are shifted into alignment with the carbon transforms. The magnitude of the shift is approximately 0.4 for both s and p symmetries. The fluorine basis was (22s,11p). Only the data points near the maximum and minimum were plotted for the 1s and 2s. Most surprising is the excellent agreement for the 2p AO in spite of the additional three electrons occupying that orbital in fluorine.

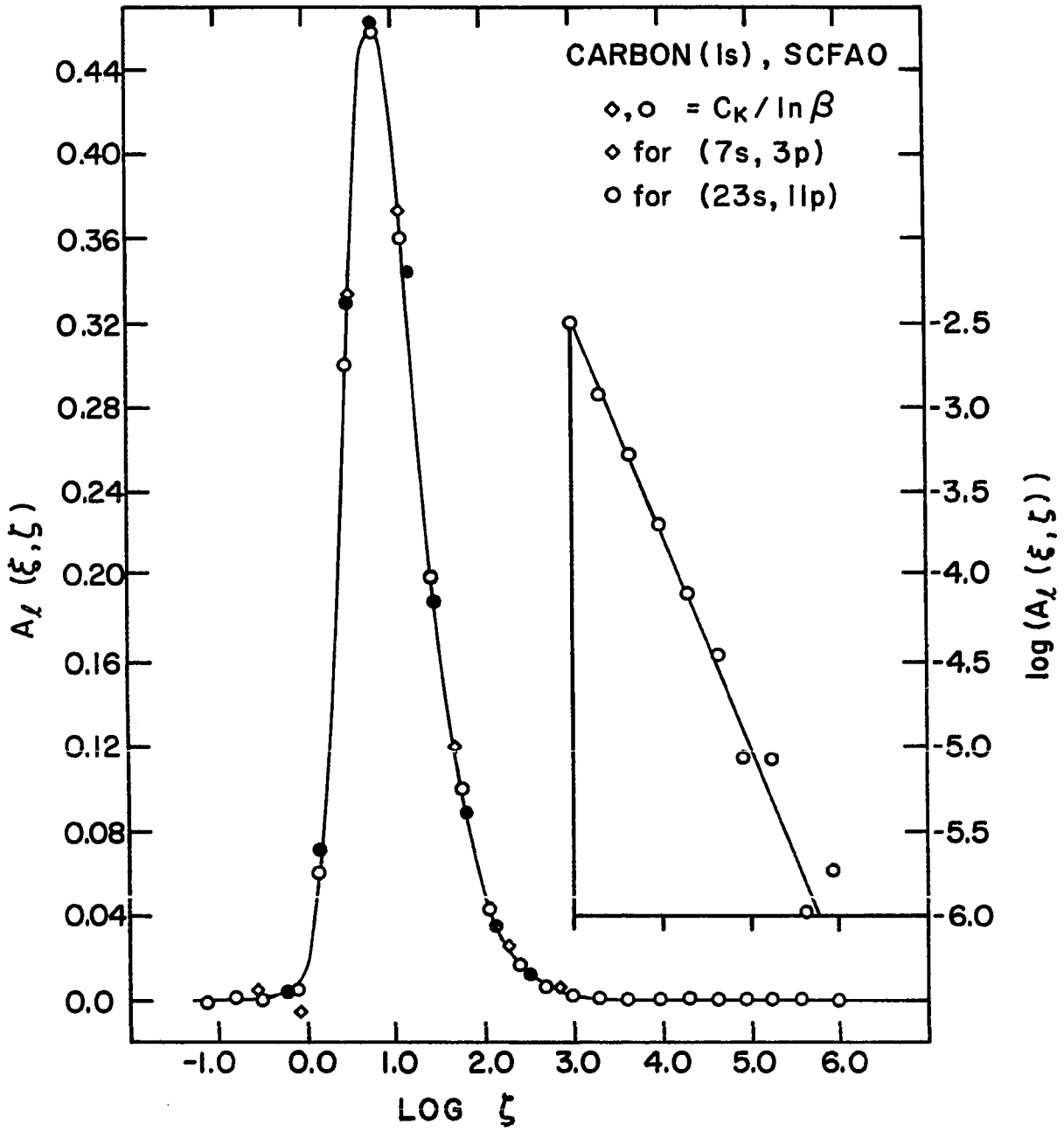


Figure 1d. Superposition of the integral transforms for the carbon and fluorine 1s atomic orbital

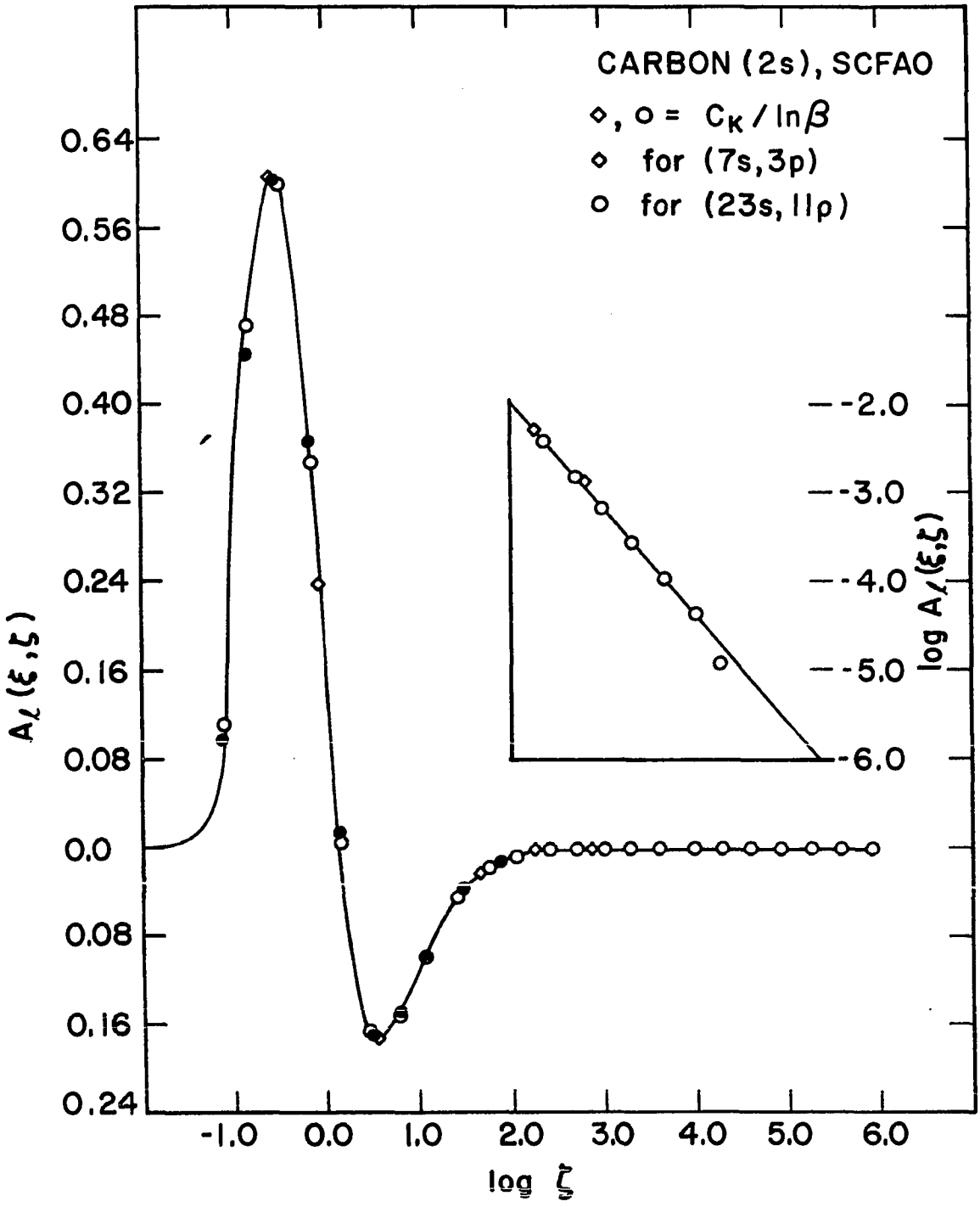


Figure 1e. Superposition of the integral transforms for the carbon and fluorine 2s atomic orbital

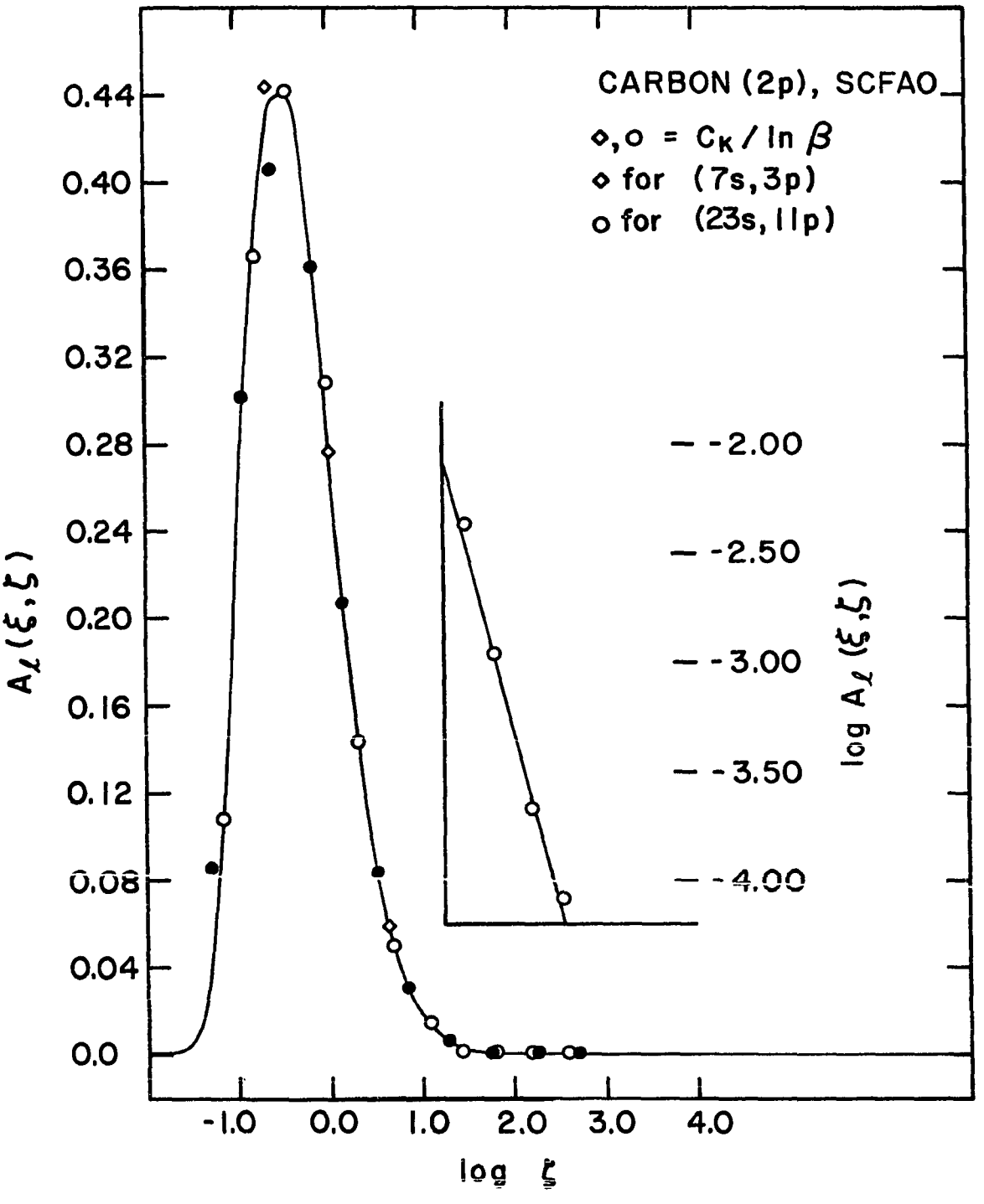


Figure 1f. Superposition of the integral transforms for the carbon and fluorine 2p atomic orbital

C. Regularities in the Optimal Atomic ETG Exponential Parameters

How do the optimal atomic (α,β) depend on the expansion length used in an atomic SCF calculation? In order to answer this question, several basis sets ranging in size up to 16 s-type and 7 p-type gaussian primitives for the first row elements and up to 9 p-type primitives for sulfur were optimized by varying (α,β) until the lowest energy for the appropriate groundstate was obtained. For the s symmetry optimizations, four p-primitives were used for the 2p AO in C and O, while six p-primitives were used in S and Se. For the p symmetry seven s-primitives were used for the 1s and 2s AO's in C and O, while eight s-primitives were used in S and Se. It was established that the optimization of one symmetry is highly independent of the number of primitives in the other symmetries so long as the other symmetry is not overly truncated. Table 3 gives the optimal values found.

In Figures 2a-2c the $\ln(\ln(\beta))$ for some of the optimal parameters listed in Table 3 is seen to be linearly related to the $\ln(N-1)$ where N is the number of primitives of that symmetry in the basis set. This linear dependence is suggested by the behavior of largest exponent, i. e. $\lim \ln(\alpha\beta^N) = \infty$ whereas individually $\lim \ln(\alpha) = -\infty$ and $\lim \ln(\beta) = 0$. Thus

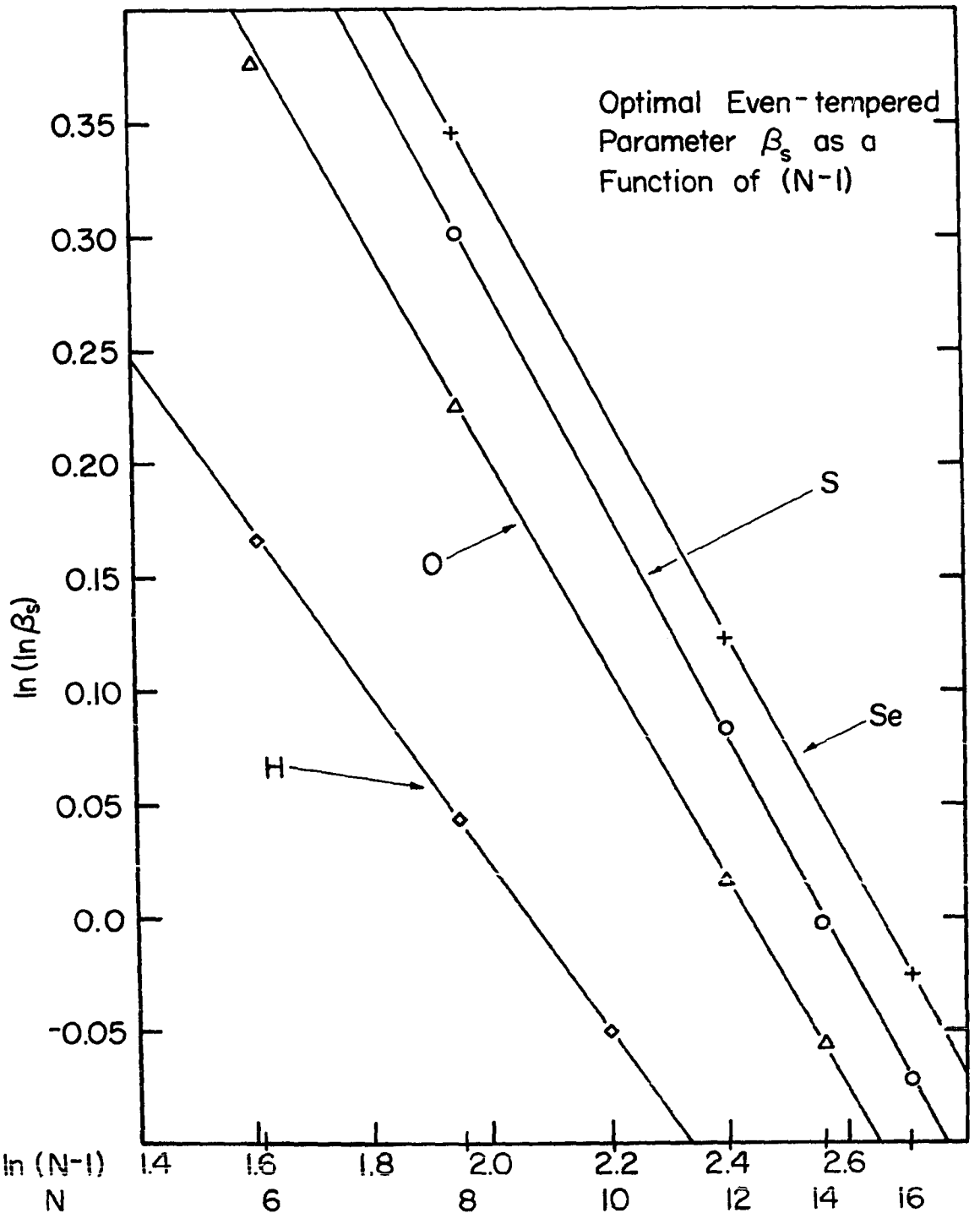


Figure 2a. Optimal atomic β_s as a function of the number of primitives in the basis set

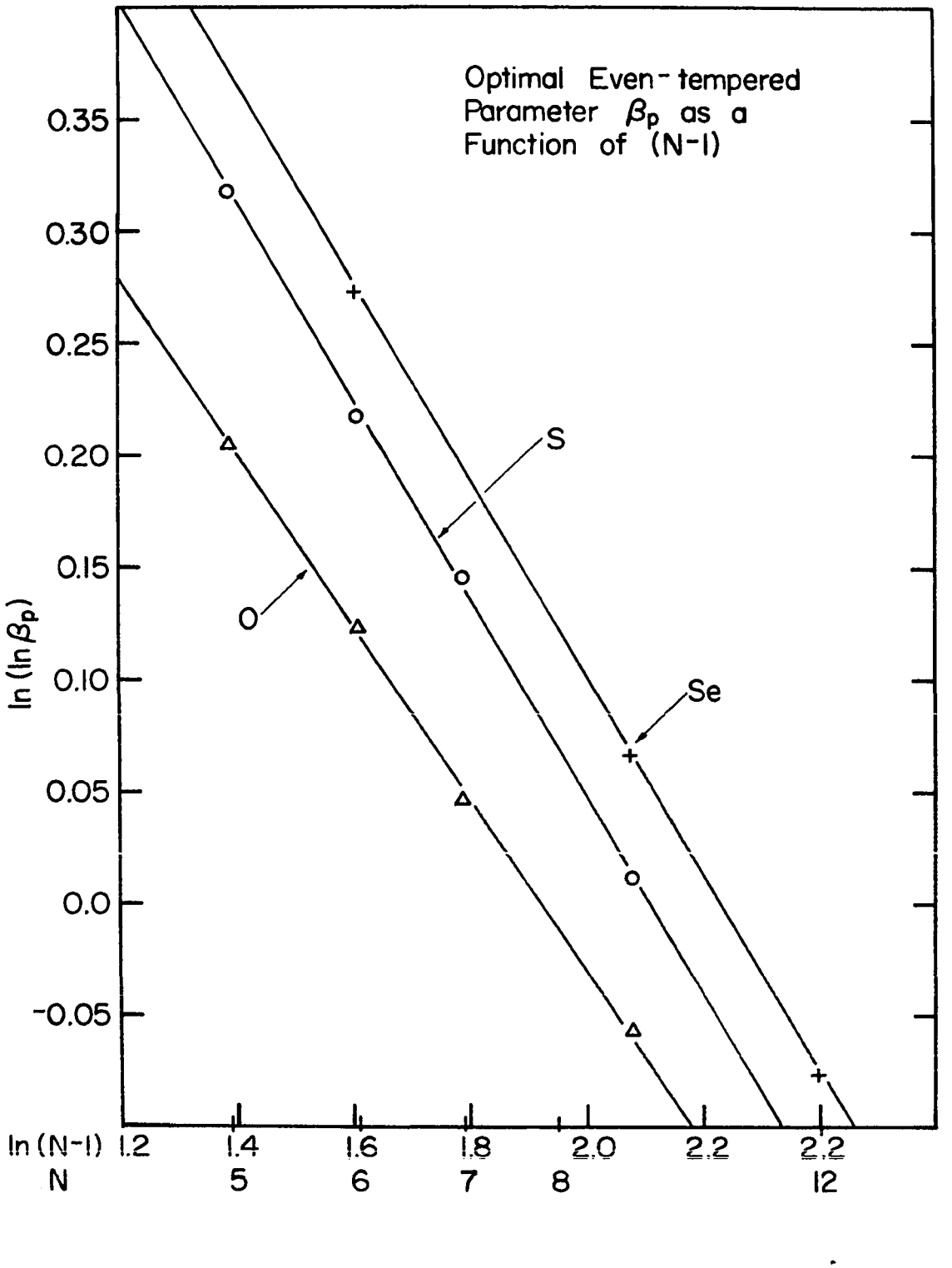


Figure 2b. Optimal atomic β_p as a function of the number of primitives in the basis set

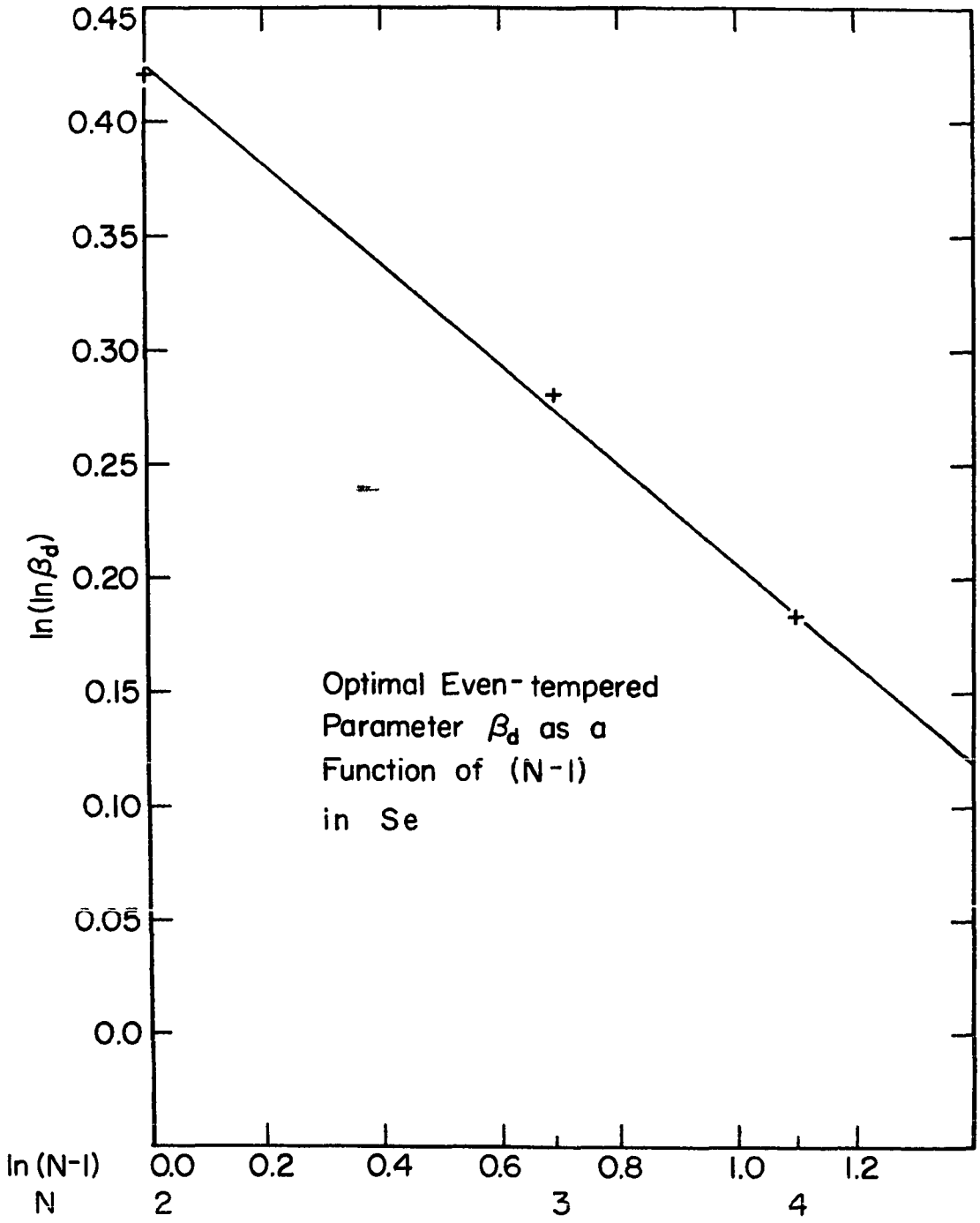


Figure 2c. Optimal atomic β_d as a function of the number of primitives in the basis set

Table 3. Optimal ET parameters for ground state carbon, oxygen, sulfur and selenium

	Carbon		Oxygen		Sulfur		Selenium	
Ns	Alpha	Beta	Alpha	Beta	Alpha	Beta	Alpha	Beta
5	0.05813	5.0784	---	---	---	---	---	---
6	0.06304	4.3406	0.11979	4.2931	---	---	---	---
7	0.05701	3.7973	0.10763	3.7710	---	---	---	---
8	0.05090	3.5089	0.09469	3.4995	0.07515	3.8708	0.10891	4.1055
9	0.05019	3.2937	0.09357	3.2790	0.07917	3.5992	---	---
10	0.04989	3.1112	0.09311	3.0895	0.07982	3.3199	---	---
11	0.04798	3.9032	0.08874	2.8978	0.06601	3.0920	---	---
12	0.04495	3.7647	0.07982	2.7649	0.06442	2.9667	0.08420	3.1013
13	---	---	---	---	0.06330	2.8338	---	---
14	0.04339	3.5790	0.08032	2.5709	0.06056	2.7079	---	---
15	---	---	---	---	0.05901	2.6143	---	---
16	---	---	---	---	0.05829	2.5385	0.06852	2.6510

	Carbon		Oxygen		Sulfur		Selenium	
Np	Alpha	Beta	Alpha	Beta	Alpha	Beta	Alpha	Beta
3	0.04550	4.4504	0.08159	4.5997	---	---	---	---
4	0.04168	3.7920	0.07205	3.8890	---	---	---	---
5	0.03806	3.3503	0.06523	3.4182	0.07202	3.9473	---	---
6	0.03523	3.0382	0.05859	3.1034	0.06446	3.4728	0.12810	3.7250
7	0.03229	2.8151	0.05477	2.8553	0.05263	3.1757	---	---
8	---	---	---	---	0.04859	2.9505	---	---
9	---	---	---	---	0.04488	2.7523	0.06941	2.9142
12	---	---	---	---	---	---	0.04422	2.5211

	Selenium	
Nd	Alpha	Beta
2	0.68011	4.5820
3	0.57640	3.7667
4	0.44612	3.3172

$\ln\beta$ must tend to zero less strongly than N tends to infinity. This would suggest a function of the form

$$\ln(\ln(\beta)) = k*\ln(N-1) + l \qquad (2.13)$$

Figure 3 shows the approximately linear dependence of $\ln(\alpha)$ on $\ln(\beta)$. For the larger bases this is seen to give a very close fit to an equation of the form

$$\ln(\alpha) = m*\ln(\beta) + n \qquad (2.14)$$

The values for the constants appearing in Equations (2.13) and (2.14) which are obtained by least squares fitting the data for H, C, O, S and Se are listed in Table 4. It would seem a

Table 4. Constants in the straight line approximations for the ET parameters

Atomic	s-symmetry				p-symmetry			
Number	k	l	m	n	k	l	m	n
1	-0.369	0.763	0.467	-3.983				
6	-0.465	1.084	0.704	-3.810	-0.362	0.689	0.708	-4.163
8	-0.443	1.084	0.769	-3.269	-0.362	0.685	0.820	-3.760
16	-0.487	1.278	0.711	-3.509	-0.440	0.929	1.505	-4.640
34	-0.485	1.290	--	--	-0.443	0.988	--	--

difficult task to a priori predict the parameters of equations (2.13) and (2.14) from simple analytical reasoning. The slopes of $\ln(\ln(\beta))$ as a function of $\ln(N-1)$ can be found by a r-weighted least-squares fitting of accurate exponential basis set atomic calculations [9], but the intercept of this line and

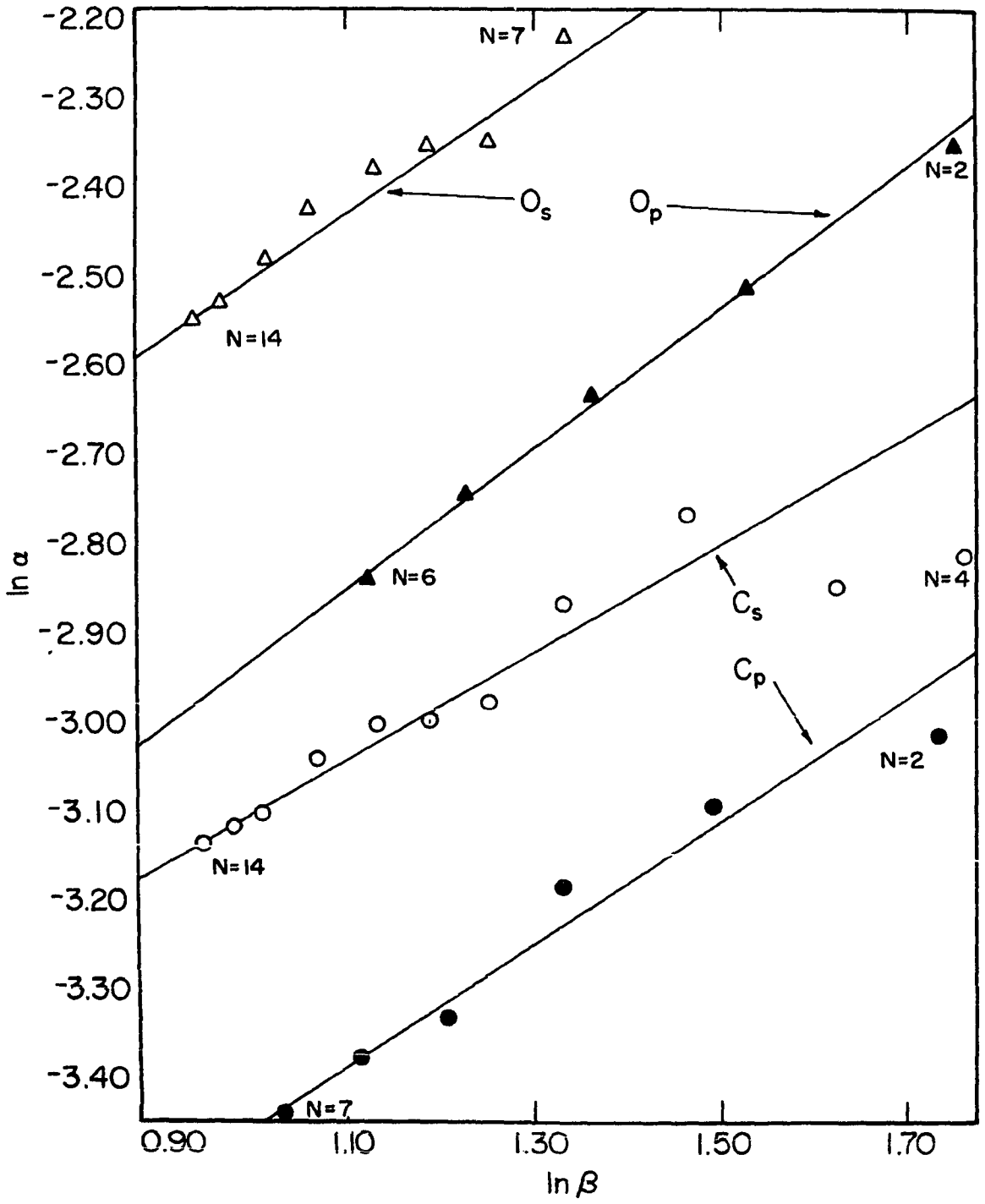


Figure 3. Optimal ETG atomic parameters: $\ell n\alpha_\ell$ as a function of β_ℓ

both the slope and intercept of the $\ln(\alpha)$ vs. $\ln(\beta)$ line cannot accurately be determined in this manner.

Energy losses encountered when optimal (α,β) values are replaced with those values predicted from Equations (2.13) and (2.14) vary with basis sets and are listed in Table 5. By comparing the values in this table with the optimal ET exponential results of Raffenetti for the same atoms it can be seen that the ratio of the number of gaussians to exponentials required to achieve the same total energy is approximately 3:1 for the first row, 2.4:1 for the second row and 2:1 for the third row of the periodic table, reflecting the fact that the build-in advantage of the latter in describing the cusp is becoming less important to the total energy.

Rapid convergence of the two columns results from the decreasing deviation of the optimal (α,β) points from the linearly interpolated values and from the simultaneous increase in the flatness of the energy surface as a function of the ET parameters. Additional optimization of the (23s,11p) basis for carbon confirmed the validity of the linear points at large N values.

Not only is the set of optimal ET parameters a smooth function of the number of gaussians used for the expansion, but also it behaves smoothly in going across the periodic table. In Figure 4 the double logarithm of the optimal beta values for a (7s,4p) ETG basis, as determined by Raffenetti [10], are plotted against the logarithm of the atomic numbers.

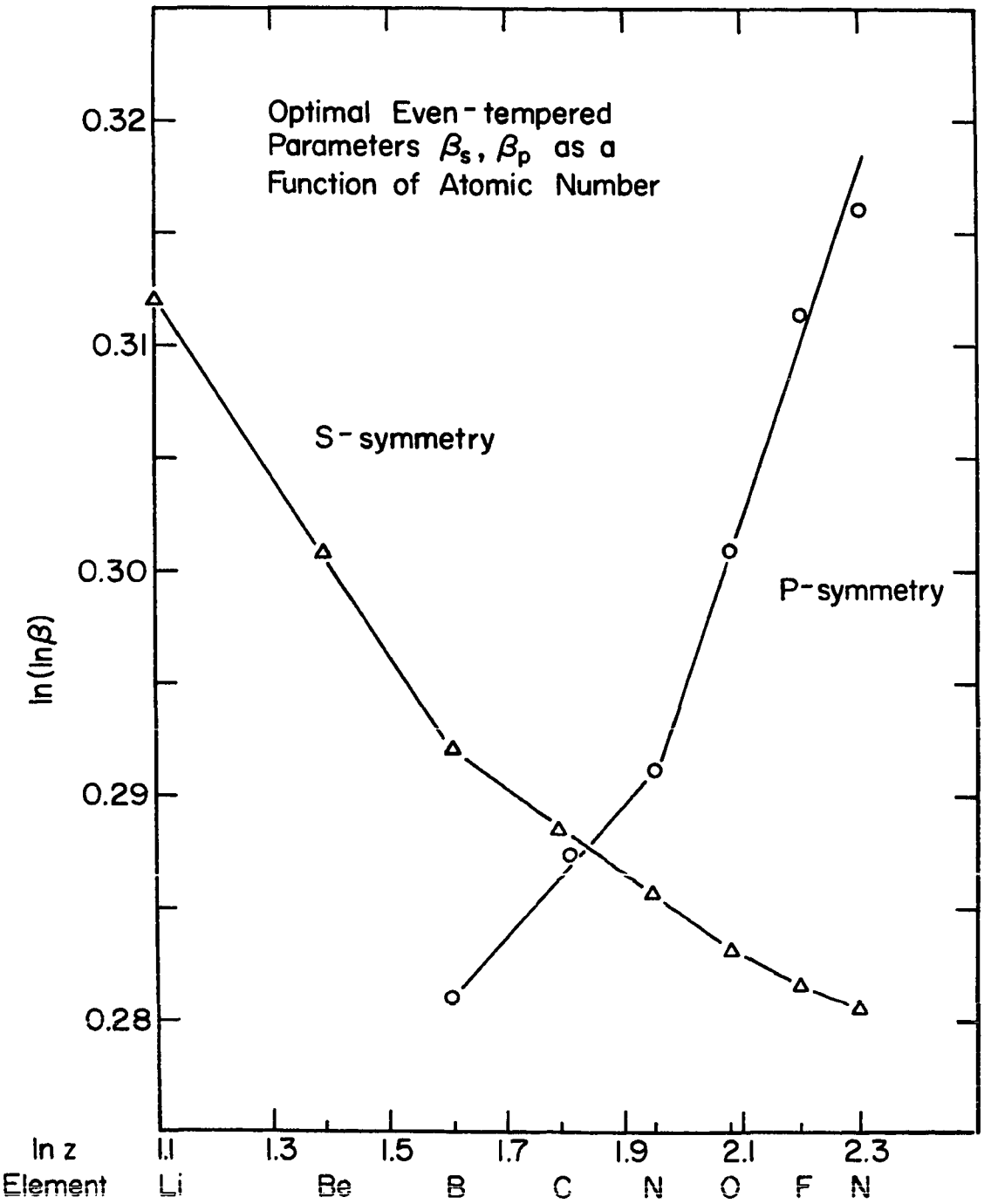


Figure 4. Dependence of the ET β_ℓ for a (7s,4p) basis on the atomic number

Table 5. Optimal and "straight line" atomic energies

Basis	Carbon		Oxygen	
	optimal	str. line	optimal	str. line
(7s,3p)	-37.630142	-37.628856	-74.343343	-74.330671
(9s,4p)	-37.676799	-37.676754	-74.716530	-74.715517
(11s,5p)	-37.685532	-37.685492	-74.791194	-74.791182
(13s,6p)	-37.687815	-37.687811	-74.804559	-74.804558
(15s,7p)	-37.688380	-37.688380	-74.808117	-74.808117
(17s,8p)		-37.688541		-74.808117
(19s,9p)		-37.688592		-74.809266
(21s,10p)		-37.688610		-74.809350
(23s,11p)	-37.688614	-37.688614		-74.809381

Basis	Sulfur	
	optimal	str. line
(10s,6p)	-397.26057	-397.25823
(12s,7p)	-397.43035	-397.43023
(14s,8p)	-397.47966	-397.47966
(16s,9p)		-397.49538
(18s,10p)		-397.50121
(20s,11p)		-397.50352
(22s,12p)		-397.50431
(24s,13p)		-397.50462

Basis	Selenium	
	optimal	str. line
(8s,6p,1d)	-2385.3307	
(12s,9p,2d)	-2398.1209	
(16s,12p,4d)	-2399.5044	

Discontinuities in the slope are visible for both curves. For the s curve it occurs in going from Be to B. For the p curve it occurs between N and O.

D. Regularity of the Total Energy for Atoms

As the approximate SCF orbitals approach the integral transform representation of the exact atomic orbitals by means of the systematic sequence of (α,β) points given by Equations (2.13) and 2.14), the total energy approaches the HF limit in a very regular fashion. This can be seen in Figures 5a and 5b where the beryllium energies of Schmidt [11] are used. Here the logarithm of the difference between each energy value and the near HF value of the (28s) basis is plotted as a function of 1/N and 1/ln(N!) where N is half the number of s-primitives in the basis.

Because two different linear dependencies seem to dominate at opposite ends of the basis size spectrum a simple analytic expression combining the two such as

$$\begin{aligned} \mathrm{Log}(E_N - E_\infty) = {} & A*N/(1 + \exp[-\alpha(N-2)]) + \\ & B*\ln(N!)/(1 + \exp[\alpha(N-2)]) \end{aligned} \tag{2.15}$$

is capable of fitting the entire curve quite well. The constants A and B are determined linearly while E_∞ and α are determined nonlinearly by minimizing the standard deviation of the fit. The constant 2 appearing in (2.15) may also be varied,

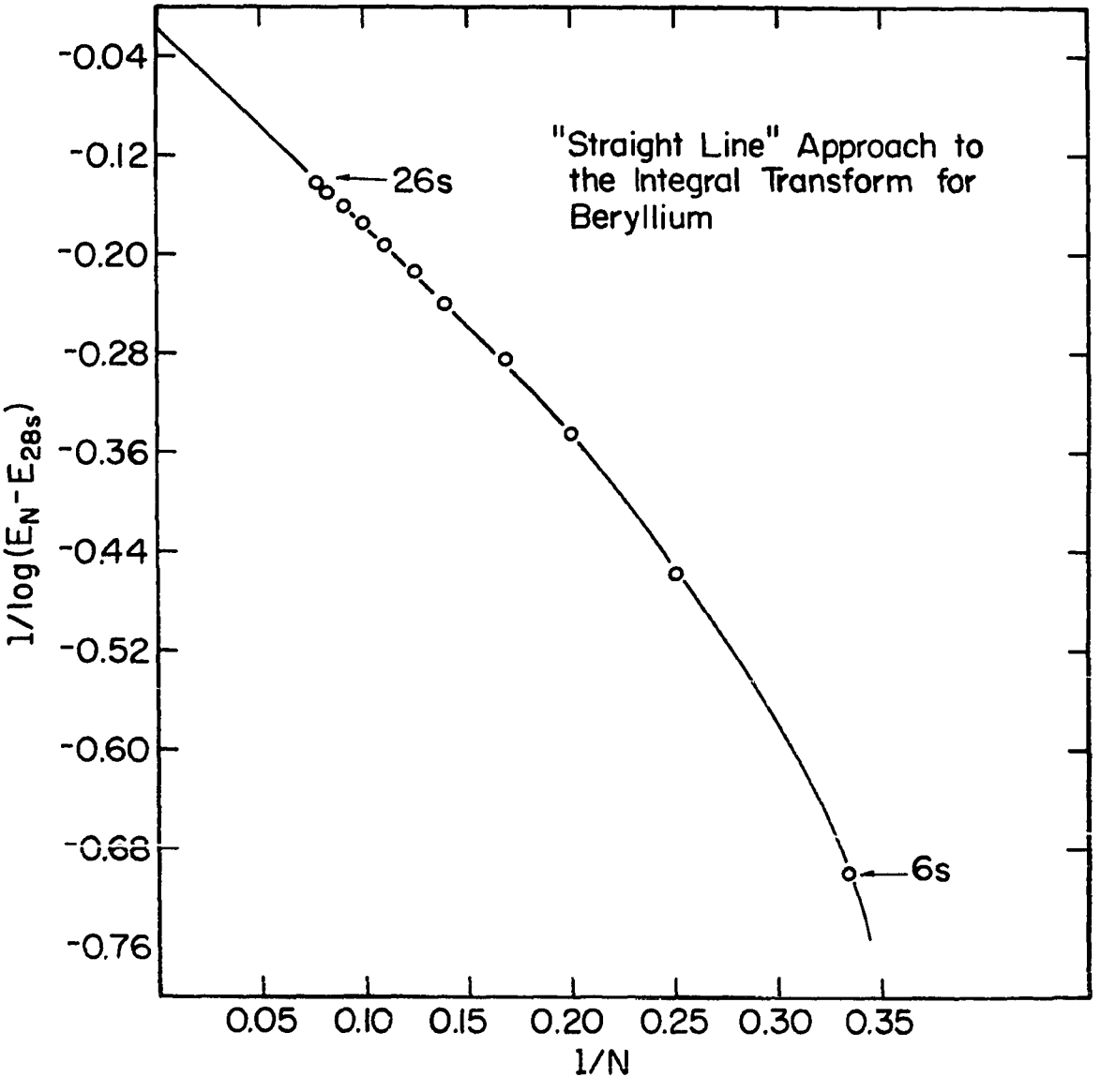


Figure 5a. Dependence of the error in the total energy as an inverse function of the number of basis functions

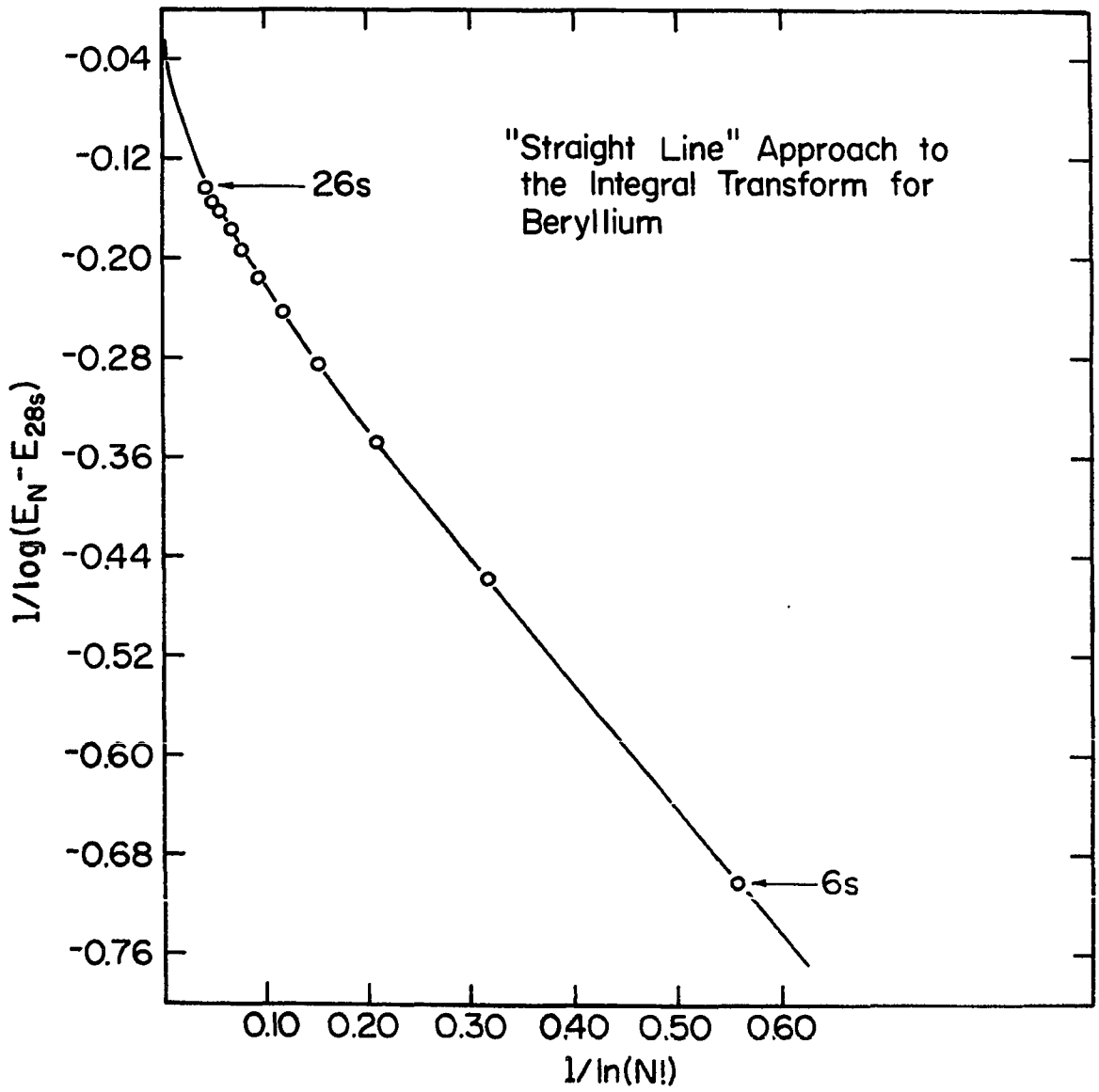


Figure 5b. Dependence of the error in the total energy as an inverse function of the factorial of the number of basis functions

however, E_∞ is quite insensitive to its value. The nine Be points were fit with a maximum deviation of 0.0002 and a typical deviation of 10**(-6) for the larger basis set points with E_∞ = -14.573023, A = -0.551292, B = -0.664424 and alpha = 0.25. Although this technique provides an extrapolation to the HF limit from a limited segment of the converging curve, its usefulness is limited to an improvement no better than one order of magnitude beyond the last computed value employed in the fit. For atoms this degree of improvement is not too difficult to obtain by merely performing the indicated calculations with a larger basis set. The novelty of Equation (2.15) lies in being able to accurately fit energies from such a wide range of basis sets with a two term expansion, and with the help of such a fit, to predict, with confidence lower bounds as well as upper bounds for the exact limiting values. This will be elaborated on in a forthcoming paper by Schmidt and Ruedenberg [12].

Our basis sets are sufficiently large and the energy values sufficiently regular that with the use of Equation (2.15) we can accurately estimate the HF limit for three of the atoms investigated. In applying Equation (2.15) we identify N with N(p), noting that N(s) = 2N(p) + 1 in carbon and oxygen, and N(s) = 2N(p) - 2 in sulfur. The limits are: carbon -37.688617 E_h, oxygen -74.809397 E_h and sulfur -397.50488 E_h. The uncertainty in these values is ±2 in the last digit. These values for the HF limit are in disagreement with the 1968 numerical Hartree-Fock energies of Fischer [13] by 0.00019, 0.00024 and

0.00122 for C, O and S respectively, all lying above the Fischer values. Subsequent numerical calculations [14] are much closer to our estimates of the lower bound.

E. Optimal ETG Molecular Parameters for Uncontracted Calculations

Since the reason for choosing a gaussian primitive basis as opposed to a set of exponential functions is the speed advantage the former gives in molecular multi-center integral evaluation a more pertinent aspect of the ET choice is that optimal ET molecular exponents are derivable from the atomic (α,β) with relative ease. Uncontracted optimizations of the ET (α,β) pairs in the molecules carbon monoxide, methane and acetylene with gaussian basis sets of (6s,3p) up to (14s,7p) demonstrated that the energy differences between the optimal atomic and molecular (α,β) values for large sets were generally less than a millihartree in size. Moreover the optimal (α,β) values for the s-type primitives were very nearly identical for the atom and the molecule after 11s. To a large extent this is to because the majority of primitives for this symmetry are needed to refine the cusp. As these large exponent functions become an increasing percentage of the basis set the (α,β) values which are optimal for the atomic cusp tend to dominate. This domination is aided by the near saturation of the valence region with sufficient functions such that exponent values can deviate considerably from the optimal ones without substantial effect on the total energy.

Optimal (α,β) pairs for p-symmetry in molecules are not observed to converge to the optimal atomic values as rapidly as the s-symmetry, but they do lie within a small region of the atomic values.

Accurate basis sets at the atomic limit are only a minimal requirement for accuracy in molecules. Uncontracted calculations on CO with optimal atomic even-tempered exponents show that there exists an "additional molecular error" beyond that which would be expected from the sum of the atomic errors seen in Table 5. This is illustrated in Table 6. Here the "additional molecular error" is defined as the difference between the error due to basis set truncation within each orbital symmetry for the molecule and the error due to basis set truncation within each orbital symmetry for the two atoms. It is seen that the error generally decreases with increasing basis set size. The magnitude of the molecular error also depends on the particular elements involved and the internuclear separations. The HF limit for CO comes from a hartree extrapolation of the total energies resulting from the largest three bases.

Even with the simplifications inherent in the even-tempered approach, molecular optimization is still very time consuming. Moreover the set of optimal atomic parameters are easily predicted for any size basis while the molecular set is quite unpredictable for small to medium size bases. Because of the similarity between atomic and molecular (α,β) pairs it was

Table 6. Molecular errors for carbon monoxide

(a) Optimal atomic even-tempered parameters

Basis	Energy(a.u.)	Total Molecular Error	Addit. Molecular Error
4s,2p	-110.49971	2217.0	17.8
6s,3p	-112.36763	349.1	5.1
8s,4p	-112.64166	75.0	3.5
10s,5p	-112.69584	20.9	3.8
12s,6p	-112.71028	6.4	1.8
14s,7p	-112.71478	1.9	0.8
16s,8p	-112.71579	1.0	0.5
HF-limit	-112.7167	0.00	0.00

(b) Optimal molecular even-tempered parameters

Basis	Energy(a.u.)	Total Molecular Error	Addit. Molecular Error
4s,2p	-111.02170	1695.0	527.7
6s,3p	-112.39190	324.8	-24.7
8s,4p	-112.64635	70.4	-4.9
10s,5p	-112.69801	18.7	-2.5
12s,6p	-112.71176	4.9	-1.5
14s,7p	-112.71556	1.1	-0.8
16s,8p	-112.71600	0.7	-0.3
HF-limit	-112.7167	0.00	0.00

therefore decided to use the atomic sets in the present investigation.

F. Effective Contracted Orbitals for S and P Symmetries

Having determined the size of the primitive basis set, we must choose a suitable set of contracted orbitals. For this purpose the HF-AO's and the set of unoccupied (virtual) orbitals which result from the LCAO formalism serve quite well.

Using a technique developed by Ruedenberg, Bardo and Cheung [15] for deriving from uncontracted molecular calculations that set of contracted orbitals which optimally reproduces the uncontracted results we investigated the overlap of the space spanned by the atomic SCF AO's (occupied and virtual) with the optimal contracted space. Our investigation involved ET basis sets ranging in size from (6s,3p) to (14s,7p) on the molecules CO, CH_4, C_2H_2 and H_2CO. In all cases studied the space of the 1s, 2s, 2p plus the first several virtual SCFAO's overlapped the space of the most important optimal contracted orbitals to better than 0.995 as shown in Table 7.

Table 7. Transformation matrix between the HF atomic SCFAO's (occupied plus first virtual) and the optimal contracted orbitals of CO.

6s Basis				14s Basis			
Opt.	Contr./AO 1s	2s	3s	Opt.	Contr./AO 1s	2s	3s
1	0.974	0.228	0.008	1	0.476	0.878	0.006
2	-0.229	0.974	0.011	2	0.879	-0.476	0.010
3	0.005	0.012	-0.991	3	0.009	0.043	-0.992

6p Basis			
Opt.	Contr./AO 2p	3p	4p
1	0.979	0.178	0.105
2	-0.196	0.631	0.750
3	0.015	-0.722	0.612

Since the virtual orbitals tend to span the continuum of the atomic HF eigenvalue problem, the energy of these orbitals increases monotonically with their kinetic energy. The lowest-energy virtual orbitals have the lowest kinetic energies. On the other hand, since the most diffuse primitives have the smallest kinetic and potential energies of all primitives, it turns out that the lowest-energy virtual AO's essentially consist of the most diffuse primitives orthogonalized to the occupied SCFAO's. This orthogonalization is unnecessary, however, if the object is merely to span the same space. In fact, Raffenetti [16] was the first to compare a set of HF-AO's plus a diffuse primitive to Dunning's [17] contracted orbitals on the nitrogen molecule and water. He found them to be slightly better than Dunning's. Because of the ease of generating this set and because of their similarity to the HF-AO-virtual space we shall employ this type of basis in the rest of this paper. Table 8 lists the energy losses incurred with this

Table 8. Energy losses with respect to an uncontracted (16s, 8p) basis for a HF-AO-diffuse primitive contraction on CO

Contracted Basis	Energy Loss (millihartrees)
(6s,5p)	0.4
(5s,4p)	2.4
(4s,3p)	5.8
(3s,2p)	32.0
(2s,1p)	149.0

contracted basis for a (16s,8p) primitive basis on CO. The quality of this contraction scheme depends slightly on the internuclear distances involved since the exact separated atoms' coefficients are built in.

G. A Minimal Basis Set Function for Hydrogen

The hydrogen atom basis deserves special attention not only because of its ubiquitous appearance throughout chemistry but, more importantly, because of the substantial energetic effect which results from scaling its minimal basis function. Even though several contraction schemes now in use provide results within a fraction of a millihartree when two or more basis functions are used, it is nevertheless of interest to know which contracted function is most effective when used as a single minimal basis AO.

A common practice is to take its coefficients from the atomic 1s orbital and then determine the optimal scaling factor from the hydrogen molecule. A somewhat better single function is obtained by preserving the primitive exponents from the isolated atom and taking as contraction coefficients those which yield the H_2 molecular orbital resulting from an uncontracted SCF calculation. Also when used in other molecules, this minimal basis function yields a lower energy error per H atom than the scaled 1s AO. This is illustrated in Table 9 which lists the errors for calculations on methane, H_2 and

Table 9. Errors per H Atom in various molecules for hydrogen minimal basis set calculations with respect to an uncontracted H 6s primitive basis[a]

Scale Factor	H_2	C_2H_2	CH_4
1. Coefficients of H minimal basis AO from atomic 1s AO			
1.00	18.7	24.9	20.9
1.19	0.1	4.5	3.6
1.29	4.9	2.6	1.8
1.35	13.0	1.7	2.4
2. Coefficients of H minimal basis AO from uncontracted H_2 SCF MO			
1.00	0.0	4.1	3.1
1.08	5.0	1.3	1.1
1.10	7.0	1.0	1.1
1.13	11.9	0.9	1.2
1.15	17.2	1.2	1.5
3. Coefficients of H minimal basis AO from uncontracted SCF calculations on C_2H_2			
1.00		0.5	1.9
4. Coefficients of H minimal basis AO from uncontracted SCF calculations on CH_4			
1.00		2.2	0.4

[a] All errors are given in millihartrees.

acetylene made with various hydrogen minimal basis set orbitals contracted from six s-primitives, with respect to calculations made with the uncontracted hydrogen 6s basis. In addition to the two contracted orbitals just mentioned the Table also lists some results using minimal basis sets that yield optimal

results in C_2H_2 and CH_4. The carbon basis in these calculations is a (6s,3p) basis. The degree of contraction of the carbon basis influences the errors in Table 9 to less than 0.1 millihartrees.

Since acetylene and methane are usually found at opposite ends of the scaling range, the error in other hydrocarbons are presumably no larger. For use as a single basis function the minimal basis No. 2 with a scaling factor of 1.08 would seem to represent an optimal compromise.

In cases where a hydrogen atom will dissociate from the molecule or where additional accuracy is required, some diffuse primitives must be added to increase the flexibility of the basis. The quality of various basis sets of this type is illustrated in Table 10. The coefficients of the minimal basis

Table 10. Atomic and molecular hydrogen errors with the optimal H_2 contracted orbital and diffuse primitives for a 6s basis[a]

Scale Factor	Contracted Basis	Atomic	Molecular
1.00	one s orbital	18.6	0.4
1.00	two s orbitals	1.9	0.4
1.00	three s orbitals	0.2	0.4
1.00	uncontracted basis	0.2	0.4
1.10	one s orbital	48.5	14.4
1.10	two s orbitals	2.7	1.7
1.10	three s orbitals	0.3	0.5
1.10	uncontracted basis	0.3	0.5

[a]All errors are in millihartrees.

orbital are those from case 2. of Table 9 (i.e., from an uncontracted H_2 calculation with the scale factor unity). The second and third orbital, where present, are the one or two most diffuse single gaussian primitives. The error listed is with respect to the exact value in the atom and with respect to the s-limit of the SCF approximation in the H_2 molecule.

It is apparent that, when at least two s orbitals are used, the choice of a scale factor of unity will give equally satisfactory results, within a millihartree, for the free H atom as well as for the H orbital in a molecule. From the data given in Table 9 for case 2., it can be inferred that this choice will also give millihartree accuracy for hydrogen in other molecules.

H. Polarization Functions

In order to construct symmetry orbitals for use in molecular calculations, admixtures of all primitive functions (or combinations of primitives transforming according to the irreducible representations of the molecule's point group) should be included in the algorithm. While for atoms in the first and second rows this restricts the primitives to be of s or p symmetry, functions of higher angular momentum may mix in for molecules. Such functions allow the MO's to polarize in the direction of the bond and were initially suggested by Nesbet [18]. Polarization functions are known to provide a substantial energy lowering and improvement of expectation

values when compared with similar calculations without such functions. In the CO molecule, approximately three or four sets of even-tempered d primitives are required in order to attain millihartree deviations from the s, p, d basis set limit. Table 11 shows the energy lowering with the inclusion of d

Table 11. Optimal ET d-symmetry parameters and energies for CO

Basis[a]	Carbon		Oxygen		
	Alpha	Beta	Alpha	Beta	Energy
(10s,5p/5s,3p)	--	--	--	--	-112.6958
(10s,5p,1d/5s,3p,1d)	1.00000	1.09660	1.00000	1.03890	-112.7619
(10s,5p,2d/5s,3p,2d)	0.06588	3.89619	0.03088	5.87085	-112.7680
(10s,5p,3d/5s,3p,3d)	0.06856	3.81904	0.04689	4.22070	-112.7704
(10s,5p,4d/5s,3p,4d)	0.06050	3.62110	0.02814	4.11723	-112.7712

[a]Contracted orbitals are the optimal contracted orbitals for 10s,5p.

functions optimized for the CO molecule. The 1d and 2d exponents were optimized with a (6s,3p) and (8s,4p) respectively, instead of with the (10s,5p) basis. Although a satisfactory description of some properties may be obtained without such functions, others, like the internal rotation in hydrogen peroxide, require that they be present in the basis set. They likewise have a strong influence on the potential energy surfaces of the HNO system as will be seen in section II of this work. A recent study by Poirier and Kari [19] indicates that for the computed one-electron properties of first and

second row hydrides there is no economic justification for including f-symmetry polarization functions. For CO they lower the total energy by at most 6 millihartrees, since the actual restricted HF limit for CO lies within a couple millihartrees of -112.7892 E_h, obtained with a very large s, p, d, f Slater-type orbital (STO) basis [20].

I. Regularity of the Total Energy and Dipole Moment for CO

In the case of the free atom a systematic approach to the complete basis provided sufficient regularity in the total energies that extrapolation to the integral transform limit became feasible. Similar behavior is found in the carbon monoxide molecule. However, the increase in basis set size in going from an atom to a molecule precludes the use of as large a set as was used in the atoms. Table 12 lists the results of Hartree extrapolations on the total energies obtained with s, p, and s, p, d basis sets.

Another commonly computed molecular property is the dipole moment. Since this property is rather sensitive to the basis set's ability to properly span a region of space other than near the nucleus it was of interest to see if the use of an energy optimized ET basis would allow an extrapolation of the values obtained with some smaller bases. In Table 13 the values of the dipole moment from polarized and nonpolarized ET bases are reported. In some cases additional diffuse primi-

Table 12. Hartree extrapolations of CO total energies

Basis	Energy	E(limit)
(6s,3p)	-112.3919	--
(8s,4p)	-112.6464	--
(10s,5p)	-112.6980	-112.711
(12s,6p)	-112.7118	-112.717
(14s,7p)	-112.7156	-112.717
(16s,8p)	-112.7160	-112.716
Basis	**Energy**	**E(limit)**
(6s,3p,1d)	-112.4919	--
(8s,4p,2d)	-112.7231	--
(10s,5p,3d)	-112.7704	-112.783
(12s,6p,4d)	-112.7815	-112.785

tives with an s exponent of 0.06 and a p exponent of 0.03 were added to the basis to help in describing the region of space far from the nucleus.

The HF limit value is close to 0.276D and the experimental value obtained by microwave spectroscopy is -0.112D. The fact that the HF value has incorrect sign is not of concern to us for this work. What seems evident from these results is that the value of this property is too highly dependent on diffuse primitives in the basis to allow extrapolation. Even though basis sets A and B or C and D must converge to the same limit they are still far enough apart in their values that a simple extrapolation would seem of questionable merit.

Table 13. Dipole moments from various ET basis sets on CO[a]

Basis A		Basis B	Basis C		Basis D
(6s,3p)	-112.362	-112.426	(6s,3p,1d)	-112.491	--
	0.641	0.581		0.088	--
(8s,4p)	-112.640	-112.649	(8s,4p,2d)	-112.723	--
	0.552	0.567		0.336	--
(10s,5p)	-112.696	-112.701	(10s,5p,1d)	-112.770	--
	0.506	0.543		0.212	0.258
(12s,6p)	-112.7092	--		--	--
	0.479	--		--	--

[a] Every entry contains the total energy (in hartrees) in the first row and the dipole moment (in Debyes) in the second row. Basis set B contains the functions in Basis A plus some additional diffuse primitives. Basis D consists of the functions in Basis C plus some additional diffuse primitives.

III. A HYBRID GAUSSIAN INTEGRAL SCHEME

A. Objective

Numerous studies done at the Hartree-Fock SCF level show the effects of basis set truncation and give an indication of the need for large flexible sets to obtain chemical accuracy (~1 millihartree/atom). Only large sets are capable of reliably yielding results which truly reflect the level of inherent physical and mathematical approximations in the theory employed in the calculation. When inadequate bases are used it is difficult to ascertain whether disagreements with experimental results arise from the basis set or other theoretical approximations.

Since a large number of primitives are needed to describe the cusp behavior near the nuclei, it might be hoped that the omission of such primitives would produce an error that is quantitatively transferable to other geometries or states. If this were true the error would cancel for many cases of chemical interest. That such is not the case, however, is evidenced by the lack of parallelness between energy surfaces of small diffuse primitive sets and those of large sets. Various model potential and pseudo-potential methods are currently being explored for circumventing the lengthy cusp expansion.

In the procedure developed here the AO basis for the molecular calculations is large enough to guarantee at least millihartree/atom accuracy in the total energy. However, only the intra-atomic integrals are actually calculated with this large basis. Moreover, they are merely retrieved from an atomic archive. All inter-atomic integrals are calculated from a shorter substitute basis which approximates the large basis sufficiently well to insure an accuracy on the order of a millihartree. The essence of our investigation is to demonstrate that such an approach is indeed feasible by developing a particular implementation for several specific systems. The scheme is applicable at small as well as large internuclear distances. Although our principal objective is to facilitate ab initio calculations with large basis sets, it is also possible to use the technique with smaller basis sets, albeit with a smaller savings in time.

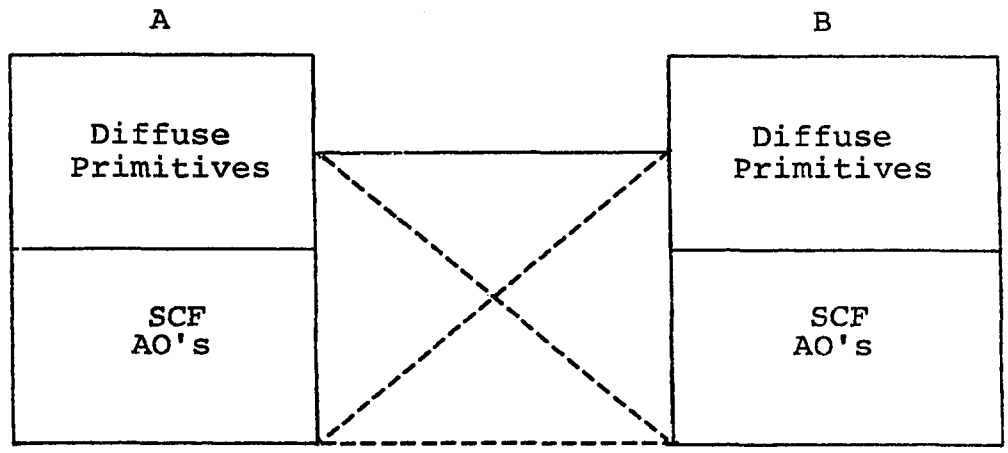
Another obstacle to obtaining chemical accuracy in molecular calculations based on analytic expansions in terms of gaussian primitives is the computation of very large numbers of two-electron repulsion integrals. For molecules possessing no exploitable symmetry the number of such integrals is proportional to (N**4)/8, where N is the number of primitives. This dependency on N currently precludes high accuracy basis sets for asymmetric molecules with more than four first row atoms. Even with much smaller basis sets large

scale ab initio calculations on medium sized molecules containing eight to twelve atoms, are only tractable with the use of various algorithms for pretesting blocks of integrals to partially avoid computing those which are negligibly small in magnitude [21-23]. This procedure can reduce the integral problem to approximately a cubic in N, and it can, likewise, be used in addition to the approximation outlined in the preceding paragraph.

B. A Fitting Procedure for the Small Basis

The conceptual formulation of the integral merging scheme is simple. A set of radial functions expanded in terms of a small number of primitives is fitted to a set of accurately determined HF-AO's, generally given by an expansion in terms of a much larger set of primitives. In a molecular calculation the numerous multi-center integrals would be generated with the AO's from the small set, while all one-center integrals would come from the proper set of AO's with the large primitive set. The approximated integrals should deviate from the "exact" multi-center integrals by some acceptable tolerance. All "exact" one-center integrals from the large basis (computed in terms of primitives all on one center) could be generated once and then stored. The integrals from these atomic calculations would be merged with the approximate multi-center integrals

for use in molecular calculations. Figure 6 shows a diagrammatic breakdown of which integrals would be approximated for a hypothetical diatomic AB using this scheme.



--- approximated

—— exact

Integrals between AO's in each rectangle are exact.

Figure 6. Diagramatic breakdown of integrals for a diatomic AB molecule computed with the merging technique

In order to test this idea we choose as "large" bases a (6s) primitive basis on hydrogen, a (16s,8p) primitive basis on carbon, nitrogen and oxygen and a (22s,12p) primitive basis on sulfur, all of which yield approximations to the exact atomic SCF wavefunctions with energy errors of less than

a millihartree. These atomic SCF calculations employed the even-tempered parameters given in Table 14, from which the

Table 14. Even-tempered parameters used in the "large" basis sets

	Carbon		Nitrogen		Oxygen	
Ns/Np	Alpha	Beta	Alpha	Beta	Alpha	Beta
16	0.04133	2.4255	0.05603	2.4026	0.07501	2.4177
8	0.03096	2.6642	0.04451	2.6711	0.05318	2.7374

	Hydrogen			Sulfur	
Ns	Alpha	Beta	Ns/Np	Alpha	Beta
6	0.03199	3.2577	22	0.05228	2.1914
			12	0.03644	2.4153

expansion of the occupied SCF AO's are easily reproduced. These SCF AO's are then fitted by shorter expansions containing the following numbers of primitives: (4s,3p), (6s,4p), (7s,5p) in C, N, and O and (7s,5p), (8s,6p) in S.

Coefficients and exponents for the fitting functions come from a simultaneous least squares, nonlinear fitting of the HF-AO's defined in the accurate basis set. The set of Ns AO's is first deorthogonalized according to the method of Raffenetti and Ruedenberg [24] to give the characteristic cuspless parts for the 2s and 3s AO's. With a straight minimization of the

sum of the individual deviations, one arising from each of the AO's, a problem of local minima is encountered. To circumvent this problem the quantity which was minimized was taken as:

$$[\gamma(\sigma_1 + \sigma_2) - (\sigma_1\sigma_2)^{1/2}]^2 \qquad (3.1a)$$

where $$\sigma_i^2 = \int dr r^2 (\phi_i^{acc} - \tilde{\phi}_i)^2 \qquad (3.1b)$$

The parameter gamma determines which of the two factors will dominate the fitting procedure. At gamma = 0.5 the standard deviations for the two orbitals are forced to be equal. For gamma >>0.5 the sum of the two deviations is being minimized. In sulfur the deviation of the 1s was ignored. Once optimal exponents were found the orthogonal set of AO's was least squares fit. In practice a value of gamma = 0.510 - 0.502 was found best.

The 2p orbital was fitted with a weighting factor of $1 + (1/r)/\langle 1/r \rangle$. The value of $\langle 1/r \rangle$ was the appropriate one for each particular atom. In sulfur the 2p and 3p were not weighted.

A simple least squares fit was used for the hydrogen function. Some improvement was found by subsequent nonlinear variation of the exponents so as to obtain good energies for C_2H_4 and C_2H_6. Table 15 contains the normalized fitting functions which have so far been tested. For hydrogen two expansion lengths were fitted to the normalized optimal H_2

Table 15. Small basis contracted orbitals fitted to the accurate AO's

Hydrogen s-symmetry

3s-Expansion		4s-Expansion	
Exponents	1s AO	Exponents	1s AO
0.148553	0.466251	0.129816	0.354942
0.655308	0.563014	0.441380	0.524776
4.427656	0.114657	1.655008	0.227469
		10.337910	0.039041

Carbon s-symmetry

4s-Expansion			6s-Expansion		
Exponents	1s AO	2s AO	Exponents	1s AO	2s AO
0.202623	-0.020492	0.817607	0.138121	-0.001569	0.471038
0.811746	0.092856	0.290926	0.444623	0.005631	0.636181
5.754027	0.697760	-0.281660	2.001122	0.154740	-0.033565
33.389878	0.362834	-0.073112	5.827941	0.531046	-0.212924
			20.804388	0.360747	-0.078197
			113.717501	0.102334	-0.024723

7s-Expansion		
Exponents	1s AO	2s AO
0.122369	0.002238	0.365526
0.354639	-0.008228	0.636622
0.977973	0.046462	0.133139
3.672232	0.410214	-0.199845
11.287713	0.462357	-0.119769
41.339238	0.198329	-0.048259
226.646517	0.042435	-0.008021

Table 15. Continued

Carbon p-symmetry

3p-Expansion Exponents	2p AO
0.147838	0.480664
0.605057	0.562737
3.098134	0.180847

4p-Expansion Exponents	2p AO
0.109285	0.306434
0.363130	0.543669
1.305269	0.316617
6.372827	0.071145

5p-Expansion Exponents	2p AO
0.089629	0.202677
0.256613	0.470321
0.767172	0.386495
2.578531	0.153578
12.217748	0.028280

Nitrogen s-symmetry

4s-Expansion Exponents	1s AO	2s AO
0.329011	-0.022418	0.923565
1.630404	0.151620	0.162535
9.597651	0.708962	-0.291192
55.385702	0.297975	-0.056177

6s-Expansion Exponents	1s AO	2s AO
0.211099	-0.000723	0.523025
0.680192	0.008060	0.594303
3.702469	0.291556	-0.104707
11.450813	0.525890	-0.190281
41.979337	0.266764	-0.056515
230.153400	0.060764	-0.015497

7s-Expansion Exponents	1s AO	2s AO
0.137648	-0.000603	0.213538
0.368286	0.002022	0.623866
1.061993	0.009878	0.322098
3.869077	0.306304	-0.167548
11.908892	0.519810	-0.161337
43.634667	0.256764	-0.062001
237.621933	0.058217	-0.013066

Table 15. Continued

Nitrogen p-symmetry

3p-Expansion Exponents	2p AO	4p-Expansion Exponents	2p AO	5p-Expansion Exponents	2p AO
0.208183	0.472236	0.161404	0.316113	0.116362	0.167814
0.894184	0.577238	0.546277	0.537646	0.336080	0.451520
4.614036	0.180851	1.965585	0.316878	1.005332	0.412265
		9.459355	0.071462	3.358313	0.181068
				15.664872	0.034607

Oxygen s-symmetry

4s-Expansion Exponents	1s AO	2s AO	6s-Expansion Exponents	1s AO	2s AO
0.386277	-0.019023	0.803408	0.257402	0.000011	0.453908
1.545005	0.093029	0.312281	0.843619	0.002122	0.649143
10.509882	0.694443	-0.291847	3.416036	0.143546	-0.011381
60.314388	0.364047	-0.081221	10.465813	0.540072	-0.231562
			37.621872	0.364914	-0.086041
			206.035910	0.101639	-0.026007

7s-Expansion Exponents	1s AO	2s AO
0.226946	0.002105	0.349596
0.667978	-0.007057	0.637572
1.868227	0.048697	0.158121
6.800774	0.410281	-0.206522
20.764500	0.461482	-0.130158
75.689975	0.195948	-0.050462
414.465533	0.041555	-0.008549

Table 15. Continued

Oxygen p-symmetry					
3p-Expansion		4p-Expansion		5p-Expansion	
Exponents	2p AO	Exponents	2p AO	Exponents	2p AO
0.268528	0.476345	0.193895	0.305329	0.157876	0.203772
1.183311	0.569557	0.685443	0.532386	0.468683	0.444734
6.129938	0.195015	2.513505	0.337723	1.425085	0.400592
		11.989289	0.080587	4.766168	0.177319
				22.046385	0.033475

Table 15. Continued

Sulfur s-symmetry

7s-Expansion

Exponents	1s AO	2s AO	3s AO
0.114570	0.001923	0.004073	0.249590
0.251073	-0.007188	-0.017052	0.560886
0.516087	0.010630	0.047910	0.389115
2.124309	-0.025737	0.563349	-0.381585
5.304948	0.065939	0.527235	-0.216545
35.887364	0.606993	-0.286144	0.089537
167.934566	0.460702	-0.156889	0.045704

8s-Expansion

Exponents	1s AO	2s AO	3s AO
0.114570	0.054160	0.003055	0.242476
0.251073	-0.016544	-0.011412	0.594288
0.550208	0.021989	0.041373	0.373527
2.074844	-0.037735	0.544466	-0.386120
5.226022	0.073224	0.548925	-0.222491
33.615957	0.529467	-0.256656	0.079678
123.620370	0.447169	-0.158888	0.046966
677.015440	0.123568	-0.033874	0.009872

Sulfur p-symmetry

5p-Expansion

Exponents	2p AO	3p AO
0.088020	0.032336	0.207422
0.212592	-0.084300	0.502224
0.587405	0.156428	0.426991
4.216851	0.704031	-0.174541
19.004673	0.377793	-0.108104

6p-Expansion

Exponents	2p AO	3p AO
0.088020	0.008009	0.206883
0.212593	-0.021284	0.505160
0.598660	0.063398	0.434205
2.799767	0.487347	-0.105775
9.196934	0.501222	-0.149470
35.452004	0.166538	-0.040729

function. Although it was possible to generate fitted functions with less than four s-primitives which gave less than a millihartree error for H_2, these were not optimal for C-H bonds.

For all the fitting basis sets a limited amount of additional nonlinear variation of the exponents was made at the equilibrium geometry and at a point on either side of the equilibrium point to see if a better fit could be obtained. Generally the final exponents were close to those obtained by the procedure described earlier.

C. Minimal Basis Set Results

Extensive tests were conducted on CO and N_2 with the integral merging technique. Both minimal and augmented basis sets were employed. The minimal basis set calculations involved six runs at various internuclear separations with all integrals computed exactly and with the two-center ones approximated by use of the merging scheme. A (16s,8p) ET basis was used for the exact calculations and for the one-center integrals which were stored for use with the merging technique calculations. At each point the deviation in total energy, one- and two-electron components of the total energy, kinetic and potential energies, and the orbital energies were determined. In addition, the average deviation between the two sets of coefficients was computed. The values found for CO are contained in Table 16, those for N_2 are listed in Table 17.

Table 16. Energy and orbital coefficient differences between merged and exact minimal basis set calculations on CO with a (16s,8p) ET basis

R	E	E1	E2	T	V	3s	4s	5s	1pi
1.90	0.0	-2.0	2.0	4.4	-4.4	0.4	0.1	-0.2	0.1
						3.3	4.3	6.7	0.5
	0.7	2.4	-1.7	1.5	-0.9	0.5	-0.6	0.4	0.0
						3.6	4.9	8.2	7.8
	4.5	-20.0	24.6	55.7	-50.7	4.7	1.9	-4.5	2.4
						68.2	12.4	12.4	8.0
2.13	0.0	-1.5	1.5	3.3	-3.3	0.3	-0.1	-0.1	0.1
						2.4	3.1	3.6	0.9
	0.9	2.3	-1.4	0.3	0.6	0.0	-0.3	0.5	0.1
						3.1	4.6	10.5	2.6
	2.0	-11.0	13.0	23.8	-21.8	3.5	-0.3	-3.7	1.5
						51.2	65.8	70.5	7.4
2.30	0.1	-1.2	1.3	2.7	-2.6	0.3	-0.1	0.0	0.1
						2.1	2.7	1.6	0.9
	0.5	0.9	-0.4	0.6	-0.1	0.0	-0.1	0.4	0.1
						3.2	4.8	7.9	4.2
	0.8	-5.9	6.6	4.1	-3.4	3.2	-1.4	-3.5	1.0
						46.8	41.4	35.6	8.1
2.50	0.1	-1.0	1.1	2.1	-2.0	0.2	-0.1	0.1	0.1
						2.1	2.4	2.9	0.4
	-0.1	-1.5	1.4	1.9	-2.0	0.0	0.2	0.1	-0.1
						4.0	4.8	4.7	5.1
	0.1	0.6	0.5	-15.0	15.1	2.9	-2.2	-3.3	0.6
						45.4	22.6	16.8	8.7
2.70	0.0	-1.0	1.0	1.7	-1.7	0.2	0.0	0.1	0.1
						2.2	2.3	1.9	0.5
	-0.5	-2.2	1.8	2.2	-2.6	-0.1	0.2	-0.1	0.1
						5.0	6.1	5.7	3.3
	1.0	5.0	-6.0	-22.1	23.1	2.5	-2.0	-3.1	0.4
						44.8	26.7	27.8	9.0

All energy deviations are in millihartrees. The basis sets are given in the order (7s,5p), (6s,4p) and (4s,3p) top to bottom. Each entry in the four righthand columns contains the orbital energy deviation on top and the standard deviation in the orbital coefficients X(10**4) beneath it for the three sigma and 1pi occupied valence orbitals. The five lefthand columns are total, one-, two-electron, kinetic and potential energies.

Table 17. Energy and orbital coefficient differences between merged and exact minimal basis set calculations on N2 with a (16s,8p) ET basis

R	E	E1	E2	T	V	3s	4s	5s	1pi
1.9	-0.6	1.7	-2.3	-8.3	7.7	0.5	-0.3	0.2	0.1
						1.5	2.6	13.1	0.1
	1.8	-0.5	2.3	4.2	-2.5	0.5	0.1	-0.1	0.2
						6.8	2.5	21.1	0.3
	0.2	29.7	-29.5	-33.4	33.6	-4.7	-3.6	5.4	-1.5
						15.6	29.5	42.3	3.0
2.07	-0.2	2.7	-2.9	-9.5	9.3	-0.3	-0.5	0.2	-0.1
						0.9	1.2	8.2	0.1
	0.9	-5.4	6.2	10.3	-9.5	0.5	0.9	0.4	0.2
						0.6	1.4	10.2	0.4
	0.5	31.6	-31.1	-9.4	9.9	-4.9	-4.0	3.5	-1.1
						26.2	9.0	46.4	2.6
2.3	0.1	2.1	-2.0	-7.2	7.3	-0.2	-0.3	0.1	-0.1
						0.5	2.0	2.5	0.1
	-0.1	7.6	-7.7	9.2	-9.3	0.4	1.2	0.5	0.3
						6.0	1.5	6.5	0.8
	1.4	37.3	-38.7	15.4	-14.0	-6.2	-4.5	1.2	-1.2
						40.8	34.0	64.5	1.4
2.5	0.1	0.4	-0.3	-3.3	3.4	-0.1	-0.1	0.1	0.1
						0.7	2.3	0.7	0.1
	-0.4	-4.6	4.2	3.0	-3.4	0.3	0.7	0.3	0.2
						6.3	1.1	5.7	1.3
	1.5	38.2	-36.7	28.1	-29.6	-7.1	-4.4	0.0	-1.3
						48.6	34.9	74.4	0.0

See material below Table 16.

As was noted previously the fitting procedure describes the near nuclear regions poorly. Thus the accuracy of the merging technique generally will decrease with decreasing internuclear separations. As R—>∞ the error will tend to zero since only the exactly computed one-center integrals will be left. Multiple bonding, such as in CO and N_2, represents a difficult test of the technique since the equilibrium bond lengths are so short, 2.13 and 2.07 bohr respectively. Some single C-O bonds distances, as in methanol, are as large as 2.7 bohr while some N-N bond distances, as in N_2H_2, are nearly 2.8 bohr.

An integral by integral analysis of the 319 two-center integrals for minimal basis set CO calculations shows a total absolute difference between the exact and approximated sets ranging from 0.013 E_h for the (7s,5p) to 0.195 E_h for the (4s, 3p). The total difference was 0.005 for the (7s,5p), and -0.012 for the (4s,3p). Only 38 integrals had an error larger than 0.0001 a.u. for the (7s,5p), the maximum being 0.0008 a.u. For the (4s,3p), 34 integrals exceeded 0.001 a.u. in error, the maximum being -0.006 a.u. The greatest difficulty seemed to lie with hybrid integrals involving the 1s function. This is probably a result of the poor description of the cusp. Of course each integral enters the Fock matrix multiplied by some element of the density matrix. Therefore no straightforward connection exists between the error in a given integral and the final error in the total energy. As is evident from the

results in Tables 14 and 15 the procedure utilizes the deviations which do exist in the integrals in such a manner that a high degree of cancellation of error occurs. In most instances the total and orbital energies are an order of magnitude more accurate than the one and two electron components of the total energy.

Tests of the 3s and 4s fittings of the optimal contracted orbital from H_2 are shown in Table 18 for scale factors of 1.10 and 1.14. Although the differences in total energy are nonzero for the 4s fit, as mentioned earlier this choice of exponents gave good C-H bond results. The error which does exist is quite stable over a wide range of internuclear distances.

Table 18. Minimal basis set total energy differences between merged and exact calculations on H_2 with a 6s basis[a]

R	Scale Factor = 1.10	Scale Factor = 1.14
1.2	0.5	0.5
	1.8	1.9
1.4	0.5	0.5
	0.9	0.7
1.6	0.5	0.5
	-0.2	-0.6
1.8	0.5	0.4
	-0.9	-1.2

[a]Top entry is from the 4s fit, bottom is from the 3s fit. All entries are in millihartrees.

For sulfur a (7s,5p) fit was made of a (22s,12p) ET set of AO's. The deviations in total energy, one-electron component

of the total energy, and orbital energies for this fit are listed in Table 19.

Table 19. Minimal basis set differences between merged and exact calculations on S_2 with a 22s,12p basis[a]

R	E_T	El	3s	4s	5s	9s	1pi	2pi	7pi
3.4	-0.4	-9.4	1.2	-0.9	-0.3	-0.6	0.4	0.0	-0.7
3.6	-0.1	1.0	0.4	-0.6	1.0	-0.5	-0.2	-0.2	-0.9
3.8	0.1	-4.4	-0.2	-0.6	0.9	-0.4	0.1	-0.1	-0.7

[a]All energies are in millihartrees.

In an effort to judge the accuracy of the method on polyatomic molecules further minimal basis set calculations were run on formaldehyde, ethylene and ethane with all the fitted basis sets. Although the error found for these systems with the smallest fitting bases were significantly larger than in the diatomic cases the error were constant in going from ethane to ethylene. This suggests that heats of reactions for similar systems may be predicted fairly well with this small basis. The deviations in total energy and the average absolute deviations in the orbital energies are listed in Table 20.

D. Extended Basis Set Results

When the contracted basis set is larger than the minimal AO set, then the merging technique uses the most diffuse primitives from the fitting basis sets to augment the atomic orbitals. Generally these primitives are not contained in the

Table 20. Minimal basis set differences between merged and exact calculations on formaldehyde, ethylene and ethane[a]

Basis	ΔE	orbital dev.
(4s,3p/3s)	-0.1	2.5
	-8.0	3.5
	-8.9	4.7
(4s,3p/4s)	-5.0	2.1
	-6.7	2.5
	-6.4	3.1
(6s,4p/3s)	0.3	0.5
	2.2	0.8
	2.0	1.2
(6s,4p/4s)	-0.2	0.3
	1.8	0.3
	2.0	0.2
(7s,5p/3s)	0.3	0.7
	-0.2	0.9
	-0.9	1.7
(7s,5p/4s)	0.0	0.2
	0.1	0.3
	0.0	0.3

[a]All energies are in millihartrees.

original large basis set so that additional primitives must be added to the large basis in order to run a check on the accuracy of the merging technique. For example, the (3s,2p) contraction on CO with the large (16s,8p) basis would require one additional s-primitive and one set of p-primitives per atom to test the merged calculations against. In the largest case the (5s,4p) contraction was tested against a full ab initio

calculation done with (19s,11p) primitives. Table 21 shows the energy losses with various contractions using the merging

Table 21. Energy losses for contracted basis set merged calculations on CO[a]

R(bohr)	(3s,2p) Contr. ΔE_T	(4s,3p) Contr. ΔE_T	(5s,4p) Contr. ΔE_T
2.0	0.4	-0.3	0.2
	1.0		
	11.3		
2.13	0.3	-0.1	0.1
	0.9	-0.4	
	8.5		
2.2	0.2		
	0.7		
	6.7		
2.5	0.2		
	0.3		
	3.2		

[a]All energies are in millihartrees.

technique. The relationship between the uncontracted (16s,8p) CO energy and the energy obtained with various HF-AO contraction lengths using the diffuse primitives from the fitting basis sets is given in Table 22. All values given in this table are from full ab initio calculations.

In an effort to expedite extended basis set calculations on sulfur the two most diffuse s- and p-type primitives from the (22s,12p) set were used as the most diffuse primitives in

Table 22. Energy differences between an uncontracted (16s,8p) ET basis and the HF-AO-diffuse primitives contraction scheme on CO with diffuse primitives from the fitting basis sets[a]

Contraction	Fitting Basis	E(total)	3s	4s	5s	1pi
(5s,4p)	(7s,5p)	-1.9	0.2	0.4	-0.1	0.3
(5s,4p)	(6s,4p)	-3.4	-0.4	1.2	0.2	1.1
(4s,3p)	(7s,5p)	-5.0	0.9	2.5	0.2	1.7
(4s,3p)	(6s,4p)	-6.2	-1.1	1.6	-0.5	0.5
(4s,3p)	(4s,3p)	-20.1	0.7	-0.4	-1.3	-0.5
(3s,2p)	(7s,5p)	-27.0	32.5	6.1	-4.2	7.3
(3s,2p)	(6s,4p)	-19.1	19.4	3.1	-5.0	3.5
(3s,2p)	(4s,3p)	-25.5	8.8	0.7	-2.8	0.5
(2s,1p)	none	-149.9	194.2	53.0	36.4	54.3

[a]All entries are in millihartrees. The four columns on the right contain the deviation in the orbital energies.

the fitting set. Thus, for contractions equal to or less than 5s,4p no new additional primitives need be added to the larger set in doing the one-center integrals. If this had not been done the large basis set would have ballooned to (24s,14p) or larger in the case of some extended basis set calculations. The inclusion of d-symmetry primitives in the basis does not present any unusual problems for the technique. When such functions were added to the basis on CO a merged calculation gave little additional error.

As a further test of the accuracy of the sulfur functions, calculations were done on the SO_2 molecule in its ground state. Two contractions, a (3s,2p/2s,1p) and a (6s,4p/4s,2p), were

chosen. The ab initio values of the total energy are -546.7544 and -547.0035 a.u. for the two contractions respectively. Although the (7s,5p) fitting basis for sulfur performed adequately at the minimal basis set level its performance deteriorated when the additional contracted orbitals were introduced. For this reason two extra fits were made of sulfur, an (8s,6p) and a (9s,7p). The results for the total, kinetic, potential, one and two-electron components of the total energy and the average absolute deviation in the orbital energies is given in Table 23. No attempt was made to

Table 23. Minimal and extended basis set calculations on SO_2[a]

Fitting Basis	E_T	E1	E2	T	V	e
(7s,5p/7s,5p)	-4.2	-0.9	-3.3	26.9	-31.1	0.5
	-9.3	-11.4	2.1	39.7	-49.0	0.6
(8s,6p/7s,5p)	-0.4	7.0	-7.4	-5.9	5.5	0.2
	-1.8	-3.1	1.3	13.8	-15.6	0.2
(9s,7p/7s,5p)	0.3	0.7	-0.4	4.5	-4.2	0.1
	-0.5	-4.1	3.5	27.3	-27.3	0.1

[a]All entries are in millihartrees. The top entry for each basis corresponds to miminal basis set calculation, the bottom to the (6s,4p/4s,2p) calcuation.

improve these last two fits with any further optimization of the exponents with regard to the SO_2 energy values. Our desire was to determine in a rough fashion how rapidly the merged basis calculations would improve as the size of the fitting

basis was enlarged. It can be seen that the kinetic and potential energies of the (9s,7p) fit are worse for the longer contraction than the (8s,6p). This situation could probably be improved by further optimization of the (9s,7p) basis.

E. Calculations on HNO with Smaller Primitive Sets

Thus far the basis sets which were fitted contained as many primitives as were necessary to approach within a millihartree of the HF limit for the respective isolated atoms. This followed from our initial intent to derive an economic procedure for doing calculations at this level of accuracy. However, such long primitive expansions are not a prerequisite for the merging technique.

Much smaller basis sets may be handled in a manner almost completely analogous to the large ones. Limitations on the smallness of the sets to be fit arise from the rapidly decreasing flexibility of the even smaller fitting basis and the lack of gaussians in the large set with a sufficiently large exponent to be classified as essentially "near nuclear" in nature. For example, if only three p-type primitives are used for a second row atom the largest exponent in the energy optimized ET set has a value in the range 3-10. This corresponds to an $\langle r \rangle$ of ~0.6-0.2 a.u. These are fairly difficult to approximate with the merging procedure. For practical purposes the lower limit of usefullness for the merging procedure falls around (8s,4p) for the basis set to be approximated.

If diffuse primitives are to be used in the approximate calculation, they will always be the most diffuse primitives in the fitting basis. However, as noted earlier in the discussion on CO, two choices present themselves with regard to how the exponents of these diffuse primitives can be determined. The most effective choice is obtained if all functions in the fitting basis are determined by optimally approximating the atomic SCF orbitals, thus also determining the diffuse primitives. The only drawback of this procedure is that in comparison calculations with the large basis the diffuse fitting primitives must be added to the primitives of the large basis to obtain the correct extended basis. This might unduly increase the number of primitives in an already large basis. This inconvenience can be avoided by choosing as diffuse primitives those of the large basis. But this implies that the most diffuse primitives in the fitting basis are predetermined, thus limiting the flexibility of the fitting basis and reducing its ability to approximate the large basis accurately.

In general we find that a (6s,3p) fitting basis is needed to insure that errors in the total energy will remain on the millihartree level when the geometry is changed or when one progresses from a minimal basis set calculation to an extended basis set calculation. If we restrict ourselves to a minimal number of contracted orbitals then only a (4s,3p) fitting basis is required. The reduction in the number of primitives from (8s,4p) to (6s,3p), while not as spectacular as that for

a larger basis, still provides a 3:1 reduction in the number of integrals over primitives if the N**4 dependence holds true. Obviously this could be a considerable savings if an energy surface was being computed.

We chose the HNO system to test the accuracy of the merging technique for small bases with the exponent restrictions mentioned above. We had previously performed calculations on this molecule in various geometries with an (8s,4p/4s) quality basis, both with and without polarization functions. Three geometries were chosen, one corresponding to a metastable O-H bond, one corresponding to a barrier with a three center bond, and one near the singlet ground state. The barrier height is ~80 millihartrees and involves relatively close internuclear distances. This represents the most difficult situation for the merging procedure. The metastable state is some 50 millihartrees above the groundstate.

More accurate ab initio calculations done with a (10s,5p, 2d/4s,1p) basis at these same geometries shows a disagreement of ~3 millihartrees with the (8s,4p,1d/4s,1p) basis. Thus we would like the merged calculations to be as accurate. Table 24 shows the deviations in total energy, one and two electron components of the total energy, kinetic and potential energies, and the average absolute deviations in the orbital energies for the three geometries.

Table 24. Contracted basis set merged calculations on HNO[a]

Geometry	E_T	E_1	E_2	T	V	e
H \O-N	0.6 -0.2	4.3 3.2	-3.7 -3.4	-4.9 -4.7	5.5 4.5	0.5 0.4
H O——N	0.3 -2.1	-23.8 -2.7	24.0 0.6	-25.3 -6.1	25.5 4.0	3.2 0.6
H O-N	1.2 0.6	1.2 0.9	0.0 -0.3	-4.4 -4.8	5.6 5.4	0.4 0.3

[a]All values are in millihartrees. e is the average absolute deviation in the orbital energies. The two entries for each geometry correspond to the following large basis set calculations, (8s,4p,1d/4s,1p) contracted to (4s,3p,1d/2s,1p) and (8s,4p/4s) contracted to (4s,3p/2s).

In this molecule not all of the atoms were merged. No attempt was made to approximate any of the hydrogen centered functions. Thus as many or as few of the multicentered integrals may be approximated in any given case, a flexibility which adds to the power of the method.

IV. THE HNO SYSTEM

A. Objective

Many triatomics composed of hydrogen and first row elements have been the object of study by theoreticians. From a theoretical standpoint an obvious practical reason for the continuing interest in this group of molecules is the opportunity afforded to perform highly accurate ab initio calculations. They thus serve as a testing ground for new techniques since limitations in the new algorithms must be measured against more elaborate calculations quite possibly approaching the exact result within several percentage points. Justification for the interest in HNO and a review of past experimental and theoretical work on the molecule is presented in the introduction to the work of Dombek [25]. Only the ground to excited singlet separation has been measured spectroscopically, the value being 1.63 eV. Ishiwata et al. [26] infer a triplet separation of 0.8 eV as a consequence of a mechanism they propose. The lowest singlet (A'-symmetry) and triplet (A"-symmetry) states for the isomerization of linear HON to linear HNO were computed by Dombek at 13 optimal SCF geometries around the NO fragment. Because of the dominance of the SCF configuration it was implicitly assumed that the reaction path could be adequately determined at this level of calculation. To obtain the final set of orbitals and energies she utilized a technique known as the Full Optimized Reaction Space method

(FORS), which consists of alternate multiconfigurational self-consistent-field (MCSCF) and configuration interaction (CI) steps in an attempt to converge to what is, in some sense, the optimal set of configurations at each geometry expressed in terms of an energetically optimal set of orbitals. A (6s, 4p/4s) ETG basis set contracted to (3s,2p/2s) with exponents optimized at various points along the reaction path was employed. The MCSCF procedure is one developed by Cheung, Elbert and Ruedenberg [27].

In the present work an (8s,4p,1d/4s,1p) primitive basis contracted to (4s,3p,1d/2s,1p) is used to compute the ground and first excited singlet states and the lowest triplet state SCF energy surfaces and the corresponding MCSCF isomerization pathways. A least-motion approach of HON to another HON, with both in the doubly occupied singlet groundstate, is also investigated to ascertain the barrier to proton transfer. This reaction is a possible mechanism for the conversion of meta-stable HON to stable HNO which would circumvent the 37 kcal/mole barrier to isomerization in the isolated molecule. The ground-state singlet curve was repeated with an even larger (10s,5p, 2d/4s,1p) primitive basis contracted to (5s,4p,2d/2s,1p) in an attempt to determine how close the (8s,4p,1d/4s,1p) basis is to SCF limit convergence.

B. SCF Calculations and Geometry Optimization

In Dombek's work on the isomerization curves an elliptic coordinate system (R_{ON}, η, ξ) was chosen for the geometry optimizations since an ellipse approximates the path taken by the hydrogen as it travels around the NO fragment. Here η = [R(OH) - R(NH)]/R(ON) and ξ = [R(OH) + R(NH)]/R(ON). Thirteen values of η were chosen at which geometry optimizations with respect to the other two variables were performed. These geometries for the ground and first excited triplet states and the next highest singlet excited along with their SCF energies are listed in Table 25. The SCF electronic configurations for these states are given by properly antisymmetrized functions of the following space products with appropriate spin functions to make a singlet or triplet:

$$^1A' = (is)^4(3a')^2(4a')^2(5a')^2(6a')^2(7a')^2(1a'')^2 \quad (4.1a)$$

$$^1\Delta \text{ (linear)} = (is)^4(3\sigma)^2(4\sigma)^2(5\sigma)^2(1\pi x)^2(1\pi y)^2 [(2\pi x)^2-(2\pi y)^2] \quad (4.1b)$$

$$^3A'' = (is)^4(3a')^2(4a')^2(5a')^2(6a')^2(7a')(1a'')^2(2a'') \quad (4.1c)$$

$$^3\Sigma^- \text{(linear)} = (is)^4(3\sigma)^2(4\sigma)^2(5\sigma)^2(1\pi x)^2(1\pi y)^2 [(2\pi x)(2\pi y)-(2\pi y)(2\pi x)] \quad (4.1d)$$

$$^1A'' = (is)^4(3a')^2(4a')^2(5a')^2(6a')^2(7a')(1a'')^2(2a'') \quad (4.1e)$$

$$^1\Delta \text{(linear)} = (is)^4(3\sigma)^2(4\sigma)^2(5\sigma)^2(1\pi x)^2(1\pi y)^2 [(2\pi x)(2\pi y) + (2\pi y)(2\pi x)] \quad (4.1f)$$

Table 25. Optimal geometries and SCF energies for the lowest singlet, lowest triplet and excited singlet states of HNO with the (8s,4p,1d/4s,1p) basis set

Designation/Eta	State	R(ON)/2	Xi	Energy
A	S(A')	1.200	2.47	-129.6240
(-1.0)	T(A")	1.205	2.47	-129.7069
	S(A")	1.200	2.47	-129.6487
	S''	1.200	2.47	-129.6089
B	S(A')	1.179	2.40	-129.6751
(-0.9)	T(A")	1.231	2.14	-129.7413
	S(A")	1.220	2.33	-129.6811
C	S(A')	1.176	2.24	-129.7053
(-0.7)	T(A")	1.259	1.98	-129.7503
	S(A")	1.220	2.17	-129.7022
D	S(A')	1.194	2.07	-129.6868
(-0.5)	T(A")	1.259	1.98	-129.7503
	S(A")	1.238	2.00	-129.6894
E	S(A')	1.210	1.89	-129.6477
(-0.3)	T(A")	1.262	1.81	-129.7161
	S(A")	1.242	1.86	-129.6572
F	S(A')	1.199	1.81	-129.6207
(-0.1)	T(A")	1.249	1.70	-129.6754
	S(A")	1.227	1.81	-129.6267
M	S(A')	1.185	1.84	-129.6243
(0.0)	T(A")	1.240	1.75	-129.6666
	S(A")	1.220	1.89	-129.6303
G	S(A')	1.161	1.95	-129.6426
(0.1)	T(A")	1.211	1.85	-129.6752
	S(A")	1.232	1.92	-129.6495
H	S(A')	1.130	2.21	-129.7005
(0.3)	T(A")	1.202	2.07	-129.7110
	S(A")	1.232	1.99	-129.6914
I	S(A')	1.115	2.30	-129.7463
(0.5)	T(A")	1.188	2.17	-129.7389
	S(A")	1.236	2.09	-129.7202
J	S(A')	1.106	2.47	-129.7575
(0.7)	T(A")	1.152	2.36	-129.7471
	S(A")	1.234	2.23	-129.7247
K	S(A')	1.111	2.60	-129.7108
(0.9)	T(A")	1.114	2.56	-129.7326
	S(A")	1.132	2.55	-129.6868
L	S(A')	1.121	2.65	-129.6419
(1.0)	T(A")	1.123	2.65	-129.7013
	S(A")	1.121	2.65	-129.6580
	S''	1.121	2.65	-129.6163

where (is) represents the totally symmetric innershell orbitals on O and N. One additional singlet state, a $^1\Sigma^+$ combination corresponding to the $(2\pi x)^2 + (2\pi y)^2$ occupancy in the linear configuration was not included in this study, because it was substantially higher in energy than the other states. Two SCF calculations were performed on this state at the linear geometries. These are indicated as S" in Table 25.

The optimal path for the singlet A' state given in Table 25 is plotted in Figure 7. All three states have very similar curves so that differences between individual paths would not be discernible in a plot of this scale except near the midpoint of the paths. Even here the differences are not great. Variations in the NO bond length are indicated and a set of energy contours from a fixed NO distance of 2.20 bohr is superimposed over the path to indicate the degree of latitude the path could have. Outer contours for the two minimas have a value of -129.680 E_h. The increment between contours is 0.010 E_h. Although variations in the structure of HNO for the three states are not large, to take point J as an example, the use of the groundstate singlet geometry for the other two states would have resulted in an energy loss of 0.006 E_h for the triplet and 0.012 E_h for the excited singlet. No geometry optimizations were attempted with the large (10s,5p,2d/4s,1p) basis.

The geometries of Table 25 can represent the "standard" reaction paths corresponding to the steepest descent from the barrier top only to the extent that the latter approximate

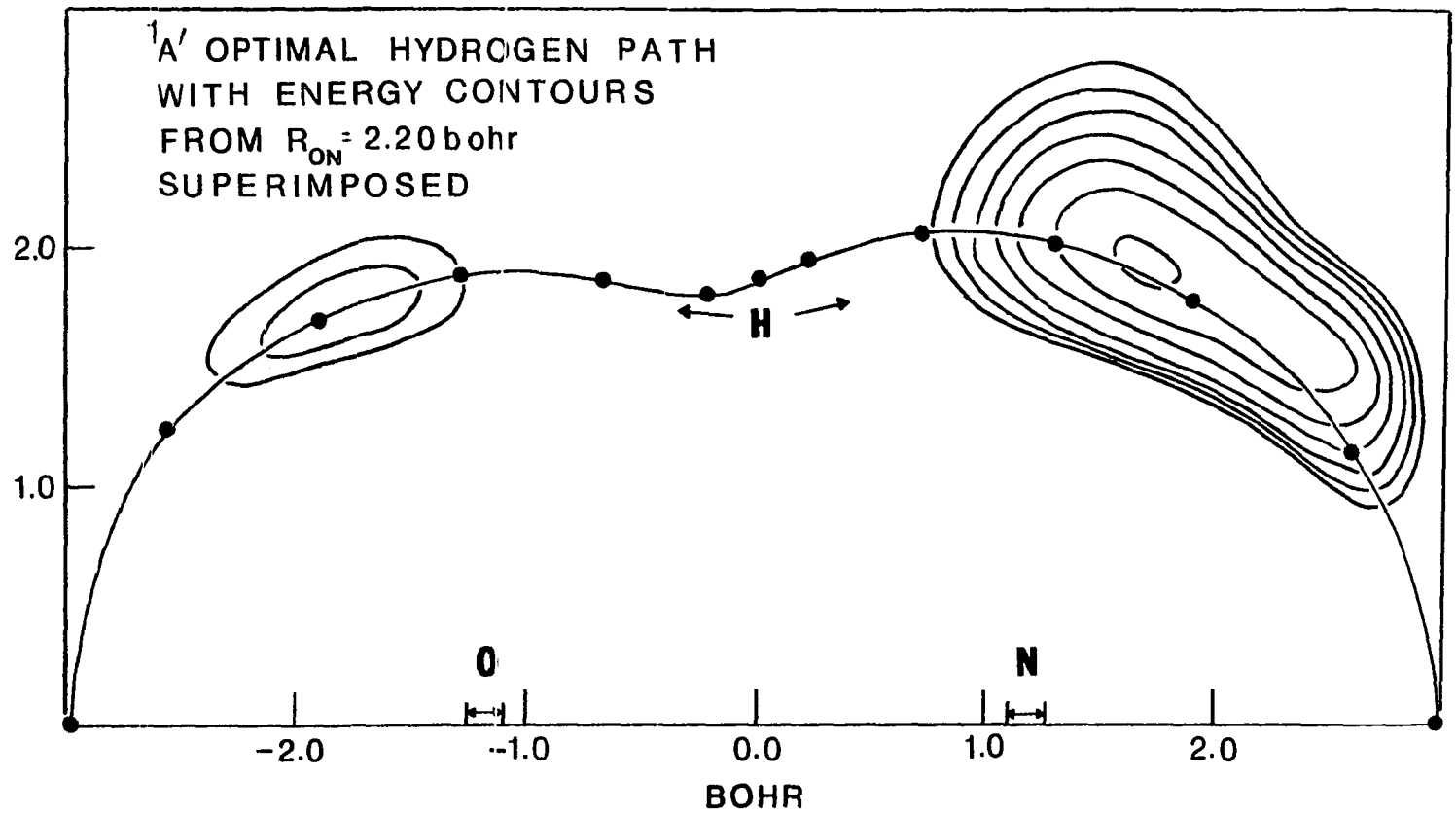
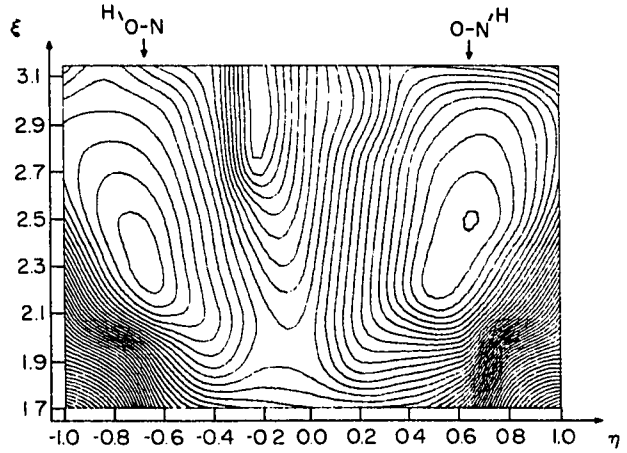


Figure 7. SCF geometry optimized path for the hydrogen atom around the NO fragment for the groundstate of HNO

elliptic curves in the vicinity of the thirteen η values chosen for this series of calculations. In the cases where this approximation holds a determination of the ξ coordinate would amount to an energy optimization along a coordinate orthogonal to the reaction path. How close the geometries of Table 25 are to the standard path can be approximately inferred from Figure 8 where contour and perspective plots of the lowest singlet state for a constant R(ON) value of 2.20 bohr are shown. While the true reaction path is a function of all three geometry variables, a good indication of its position is obtained by looking at a slice of this three dimensional hypersurface which corresponds to a fixed R(ON) value. An elliptical path would be represented by a horizontal line in Figure 8 approximately connecting the two minima. The particular choice of the N-O bond length is a near optimal value from the stable geometry minimum. Contour increments are 0.012 E_h for the contour plot. The lowest contour has an energy of -129.757 E_h for the HNO side. Although a single configuration wavefunction, as was used for these figures, cannot correctly describe the dissociation of groundstate HNO -> H + NO, the paths for this reaction are clearly visible as valleys leading up from the two minimas. Eventually, as the length of the bond to hydrogen increases, the energy surface becomes planar with a value equal to the SCF dissociation energy.

Variation in the total energy as a function of η for the values listed in Table 25 is plotted in Figure 9. Since



HNO-NOH $^1A'$ SCF ISOMERIZATION SURFACE

CONTOUR AND PERSPECTIVE PLOTS IN ELLIPTICAL COORDINATES AT CONSTANT R_{ON} = 2.20 a_o CONTOUR INCREMENT= 0.012 a.u. LOWEST HNO CONTOUR = −129.757 a.u.

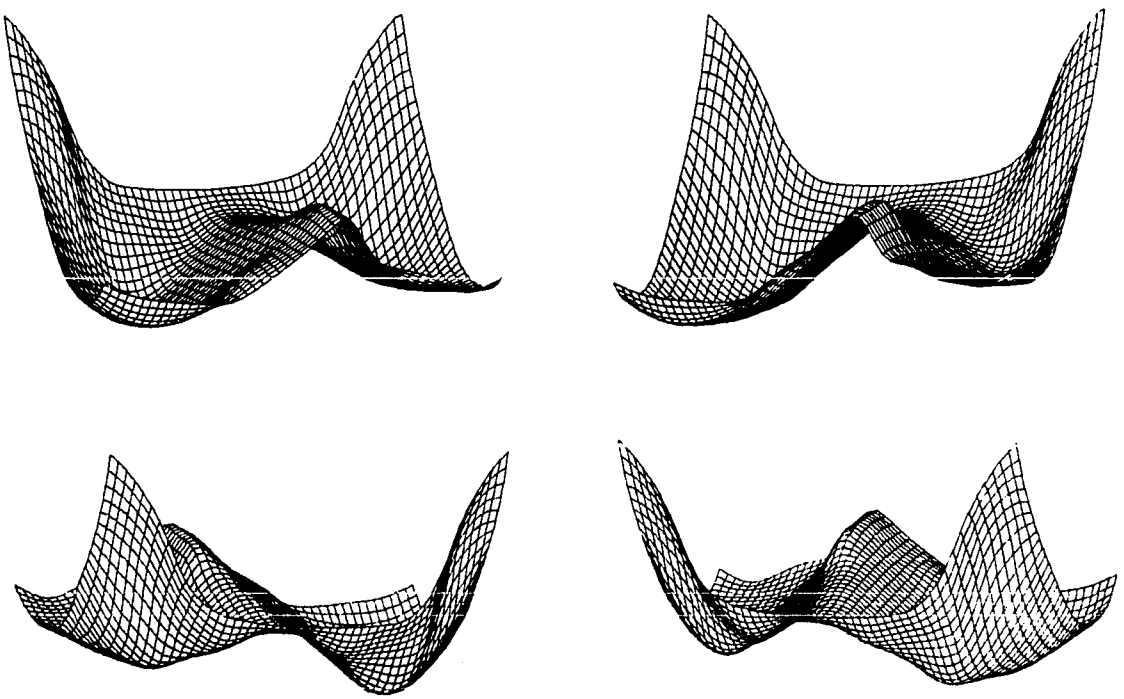


Figure 8. SCF potential surface for the groundstate

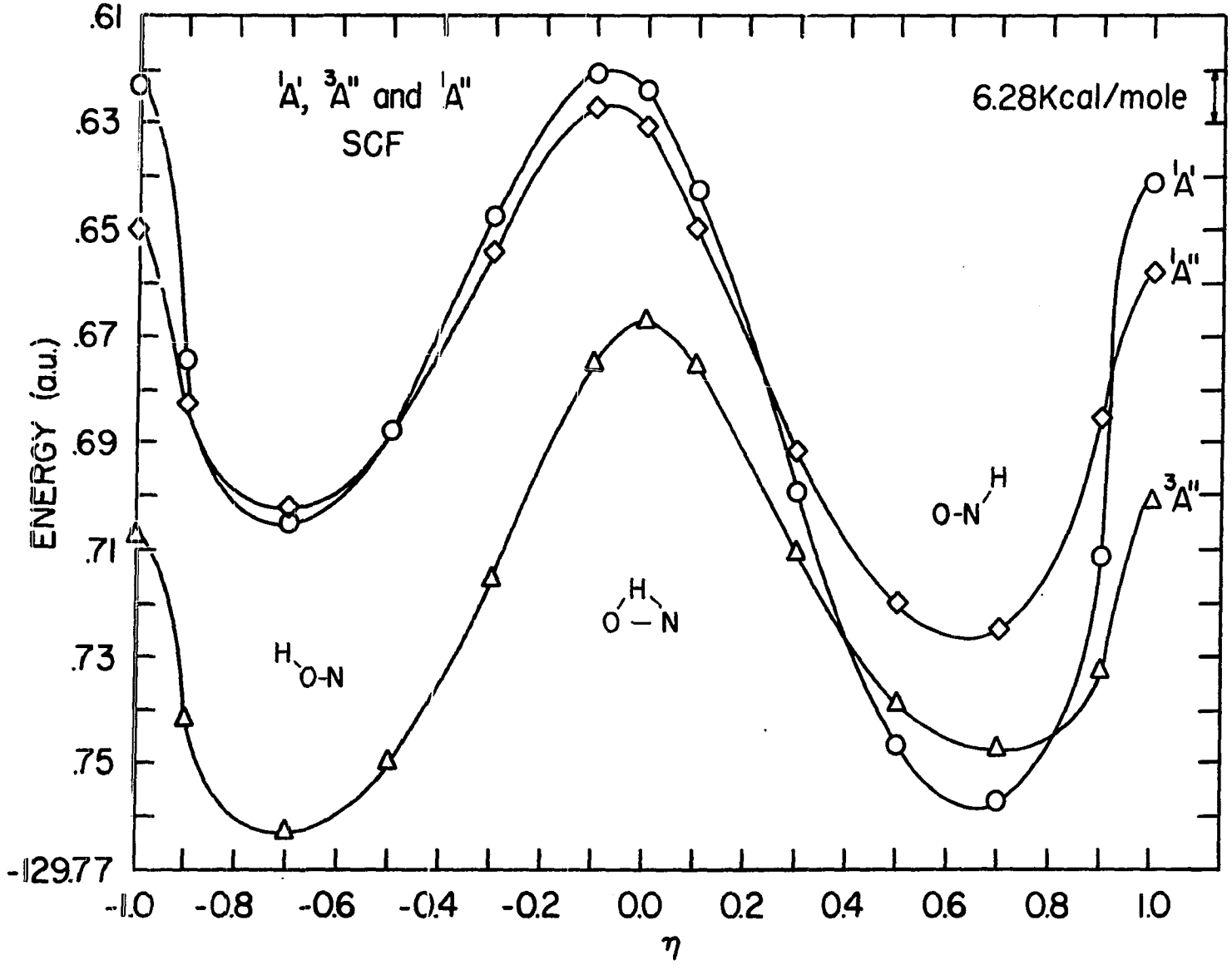


Figure 9. SCF energy curves for the three lowest states of HNO

earlier calculations with a much smaller basis yielded an inverted order for these states at geometry J compared to the values in this table, i.e., the triplet state was lower in energy than the groundstate singlet and the open shell singlet was almost equal in energy to the closed shell singlet, it was desirable to further investigate the effect of the basis set size on these curves. In Figures 10a and 10b the relative similarities of the groundstate and lowest triplet curves are demonstrated by plotting the difference between the rest of the curve and point J for a (6s,3p/3s), (8s,4p/4s), (8s,4p,1d/4s, 1p) and (10s,5p,2d/4s,1p) basis set. Along with the vertical excitation energies this is a more significant indication of the quality of a basis set than the absolute magnitude of the total energy. If extra basis functions merely lowered the curves by a constant amount they would be an expendable luxury for chemists.

Several effects are noticeable. The most dramatic of these is the large destabilization of the linear configurations with respect to the rest of the curve. Part of this is due to the fact that more d-primitives are of the proper symmetry to contribute to the occupied MO's for the nonlinear geometries than for the linear. The rest of the change seems due to the increase in s- and p-primitives. A barrier height reduction of 24 millihartrees (15 kcal/mole) is also obtained in going from the small to the large basis. This could conceivably be reduced even further by enlarging the hydrogen basis set.

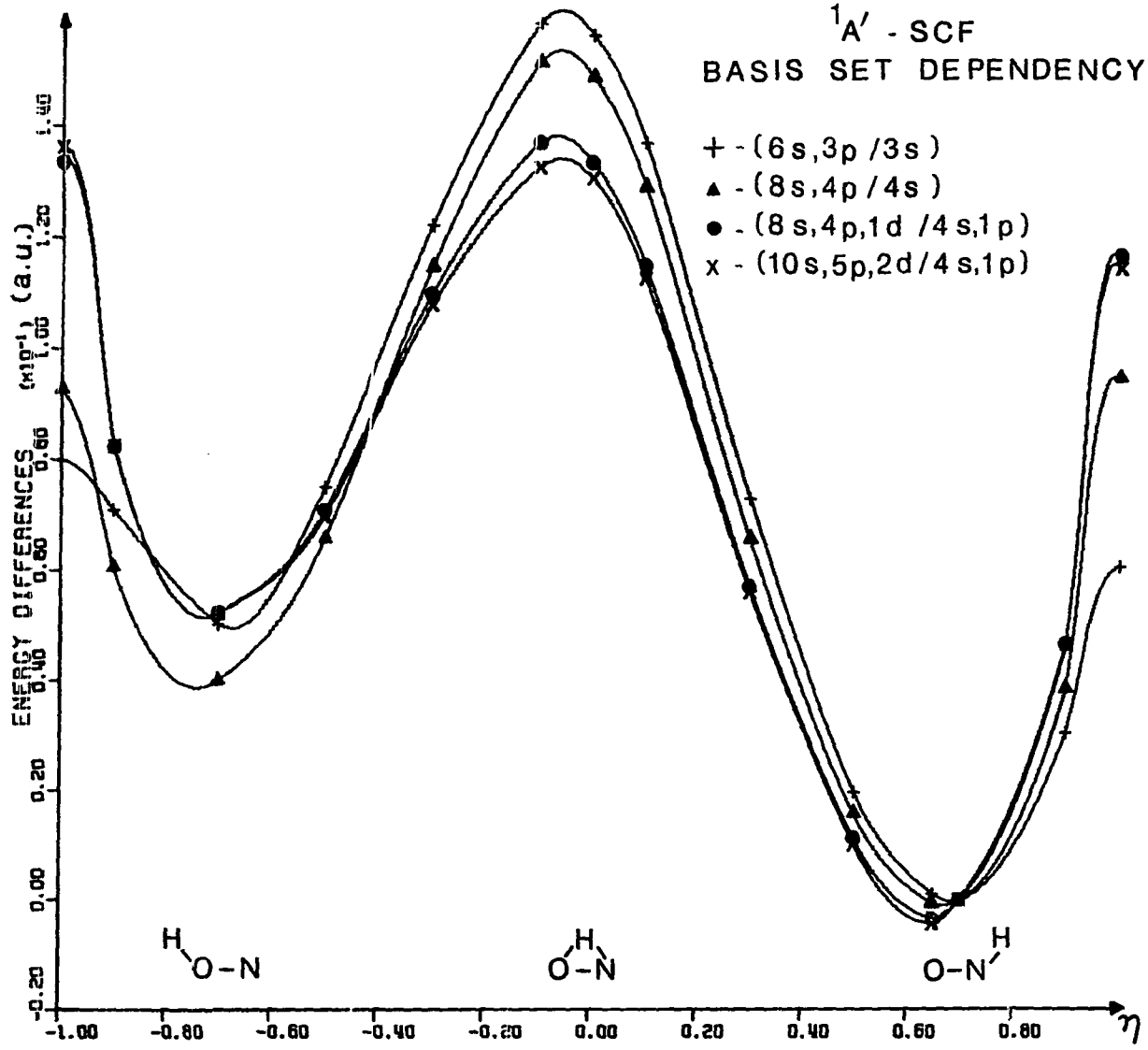


Figure 10a. SCF dependence on the quality of the basis set for the $^1A'$ state

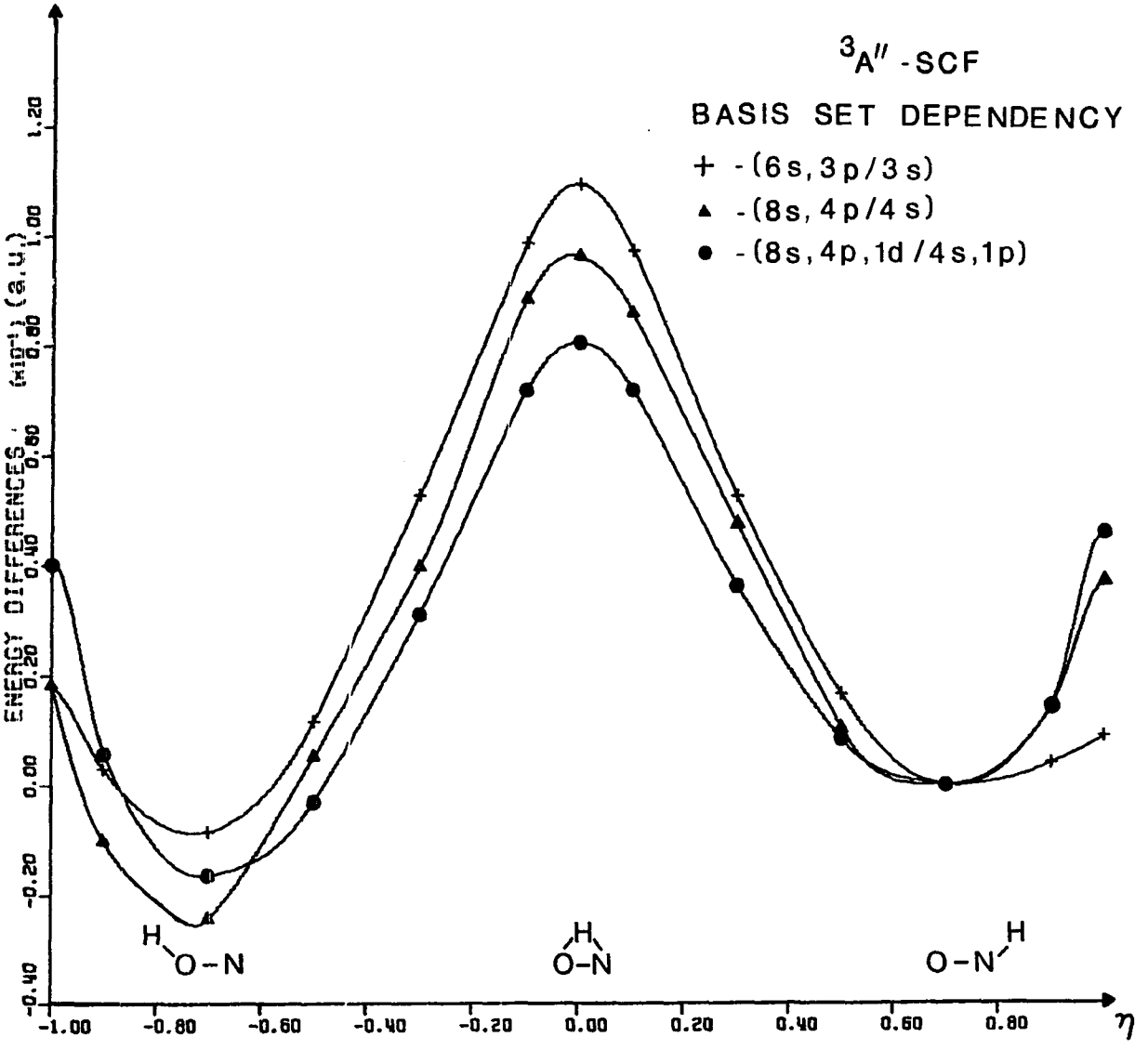


Figure 10b. SCF dependence on the quality of the basis set for the $^3A''$ state

A review of past and present HF-SCF calculations at the optimal geometry and near the experimentally determined geometries for the two singlets is given in Table 26. Using flash photolysis Dalby [28] gives the experimental geometry for the groundstate as R(NH) = 2.007 a.u., R(ON) = 2.288 a.u. and <HNO = 108.5°, and the excited state singlet geometry as R(NH) = 1.958 a.u., R(ON) = 2.345 a.u. and <HNO = 116.2°. The geometries listed for the current work are either the optimal geometry for the indicated basis or the experimental geometry for the groundstate. The latter SCF energies were required for the vertical excitation energy determination. Although the groundstate geometry is predicted quite well, the excited singlet state's NO bond length is larger than the experimental value and the bond angle is predicted to be smaller than that of the groundstate in disagreement with the experimental ordering.

C. The Full Optimized Reaction Space Curves for HNO

Most chemically interesting phenomena can be adequately described, at least qualitatively, by methods capable of recovering a substantial fraction of the valence correlation energy. The orbitals involved are those which are occupied at the HF level along with some low lying virtual orbitals. To meet the above requirement and to facilitate the interpretation of orbital changes which occur during the isomerization reaction, Dombek and Ruedenberg chose to restrict themselves to an

Table 26. Hartree-Fock SCF results near the groundstate geometry for HNO

State	R(NH)	R(NO)	<HNO	Basis	Energy	Ref.
singlet	2.109	2.311	105.1	Minimal-STO	-129.3359	[29]
(A')	1.962	2.496	110.4	(7s,3p/4s,1p)	-129.5880	[30]
	exp.	exp.	109.0	gaussian lobe	(-129.695)	[31]
	exp.	exp.	exp.	gaussian lobe	-129.7344	[32]
	1.871	2.305	109.5	6-31G	-129.7843	[33]
	exp.	exp.	exp.	(6s,3p/4s)	-129.3497	this work
	1.945	2.360	108.9		-129.3522	" "
	exp.	exp.	exp.	(8s,4p,1d/4s,1p)	-129.7555	" "
	1.959	2.224	108.6		-129.7589	" "
	exp.	exp.	exp.	(10s,5p,2d/4s,1p)	-129.8209	" "
	1.959	2.224	108.6		-129.8254	" "
triplet	2.007	2.288	124.0	gaussian lobe	(-129.700)	[31]
(A")	2.007	2.288	120.0	gaussian lobe	(-129.734)	[32]
	2.007	2.288	108.5	(6s,3p/4s)	-129.3491	this work
	1.947	2.618	106.1	(6s,3p/4s)	-129.3730	" "
	2.007	2.288	108.5	(8s,4p,1d/4s,1p)	-129.7466	" "
	1.922	2.201	106.2		-129.7496	" "
singlet	2.007	2.288	120.0	gaussian lobe	(-129.703)	[32]
(A")	2.007	2.288	108.5	(8s,4p,1d/4s,1p)	-129.7171	this work
	1.925	2.460	104.1		-129.7265	" "

Energies within parentheses are estimated from plots contained in the referenced work. These numbers are probably accurate to within a few millihartrees.

orbital space obtained by considering only those orbitals in the valence shell of the isolated atoms. This formal minimal basis consists of the 1s, 2s and sp orbitals for nitrogen and oxygen and the 1s for hydrogen. These eleven functions are divided into nine MO's of A' symmetry, and two of A" symmetry, but the two inner shell orbitals, corresponding to the 1s on nitrogen and oxygen, will be kept doubly occupied throughout all calculations. We are interested in obtaining the best possible wavefunction within this space.

The procedure followed here is the same as that used by Dombek. A complete CI in the space of these nine orbitals is performed. The CI spaces are of dimension 1316, 1722 and 1204 for the groundstate singlet, triplet and excited state singlet respectively. Then based on the expansion coefficients in a natural orbital expansion of this wavefunction the dominant 7-9 configurations are selected for use in an MCSCF calculation to optimize the orbitals in the space of the full AO basis set. Although more configurations could be used, the total energy of the full CI is rather insensitive to any additions. Following this, a second CI is done to see if the important configurations have changed as a result of the MCSCF orbital optimization. If the important configurations have changed from the previous CI, a second MCSCF step on this new list of configurations is done. A final full CI is performed. Since the original HF calculation does not occupy all eleven orbitals a preliminary small MCSCF calculation, with the occupied SCF

orbitals frozen, is used to provide an approximate initial guess for the remaining orbitals in the first CI. The twelve configurations used in this small MCSCF are some obvious ones which provide left-right correlation in the three bonds.

Total energies in the FORS approximation using the (8s, 4p,1d/4s,1p) basis are listed in Table 27. Not all of the

Table 27. Total energies in the FORS approximation for HNO[a]

Point	S(A')	T(A")	S(A")	Point	S(A')	T(A")	S(A")
A	-129.7226	.7674	.7226	G	-129.7851	.7778	(.751)
B	-129.7814	(.811)[a]	.7587	H	(-129.837)	(.804)	(.781)
C	-129.8170	.8363	.7811	I	-129.7574	.8216	.8043
D	(-129.808)	(.824)	(.772)	J	-129.8844	.8364	.8105
E	(-129.783)	(.802)	(.756)	K	-129.8317	.8205	.7899
F	-129.7633	.7791	.7330	L	-129.7583	.7908	.7583
				M	-129.7698	.7751	.7358

[a]Those energies enclosed in parentheses are estimated from the adjacent correlation energies.

thirteen points were actually computed. In an effort to minimize the computational load the correlation energies for several of the geometries were estimated from an average of the values computed at adjacent points. This estimate is probably accurate to a few millihartrees. Energy curves from this table are shown in Figure 11. Compared to the SCF curves

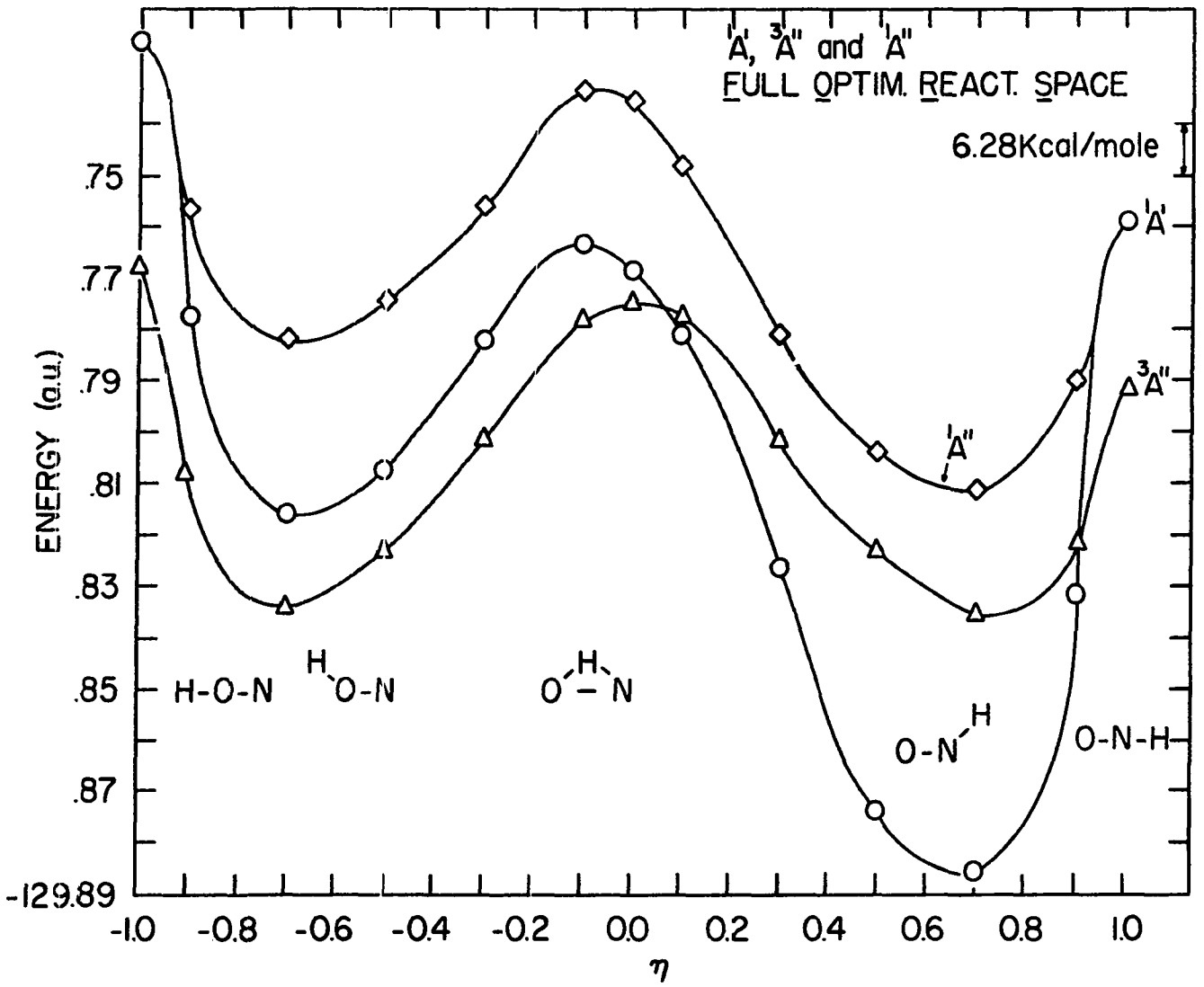


Figure 11. FORS energy curves for the three lowest states of HNO

of Figure 9 the linear geometries are even higher in energy with respect to the groundstate while the barrier to isomerization is lower. Compared to the FORS curves of the (6s,4p/4s) basis the groundstate singlet barrier is reduced from 37.8 Kcal/mole to 33.7 Kcal/mole. The triplet barrier is reduced from 46.9 Kcal/mole to 38.4 Kcal/mole.

In order that a direct comparison between the small and large basis set curves could be made the curves are superimposed at point J in Figures 12a and 12b. The actual difference in total energy for the (6s,4p/4s) and (8s,4p,1d/4s,1p) basis sets is approximately 0.24 E_h. A surprising feature of the FORS plots is the lack of any significant reduction in the barrier since in the SCF approximation the reduction in going from a small basis to a large basis was almost 15 Kcal/mole. In previous studies of the effects of basis set truncation at the CI level, changes in the total energy curves were larger at the CI level than at the SCF level. If similar behavior could be expected for the FORS curve the 34 Kcal/mole A' singlet barrier would be reduced to about 19 Kcal/mole. Figure 13 shows how the correlation energy recovered varies over the reaction path for the two basis sets. On the HON side of the curve the two are in near quantitative agreement, while at the barrier they differ by about 15%.

On the whole the final natural orbital CI in the large and small basis sets showed a great similarity with respect to the ordering of configurations. About the only exception occurred

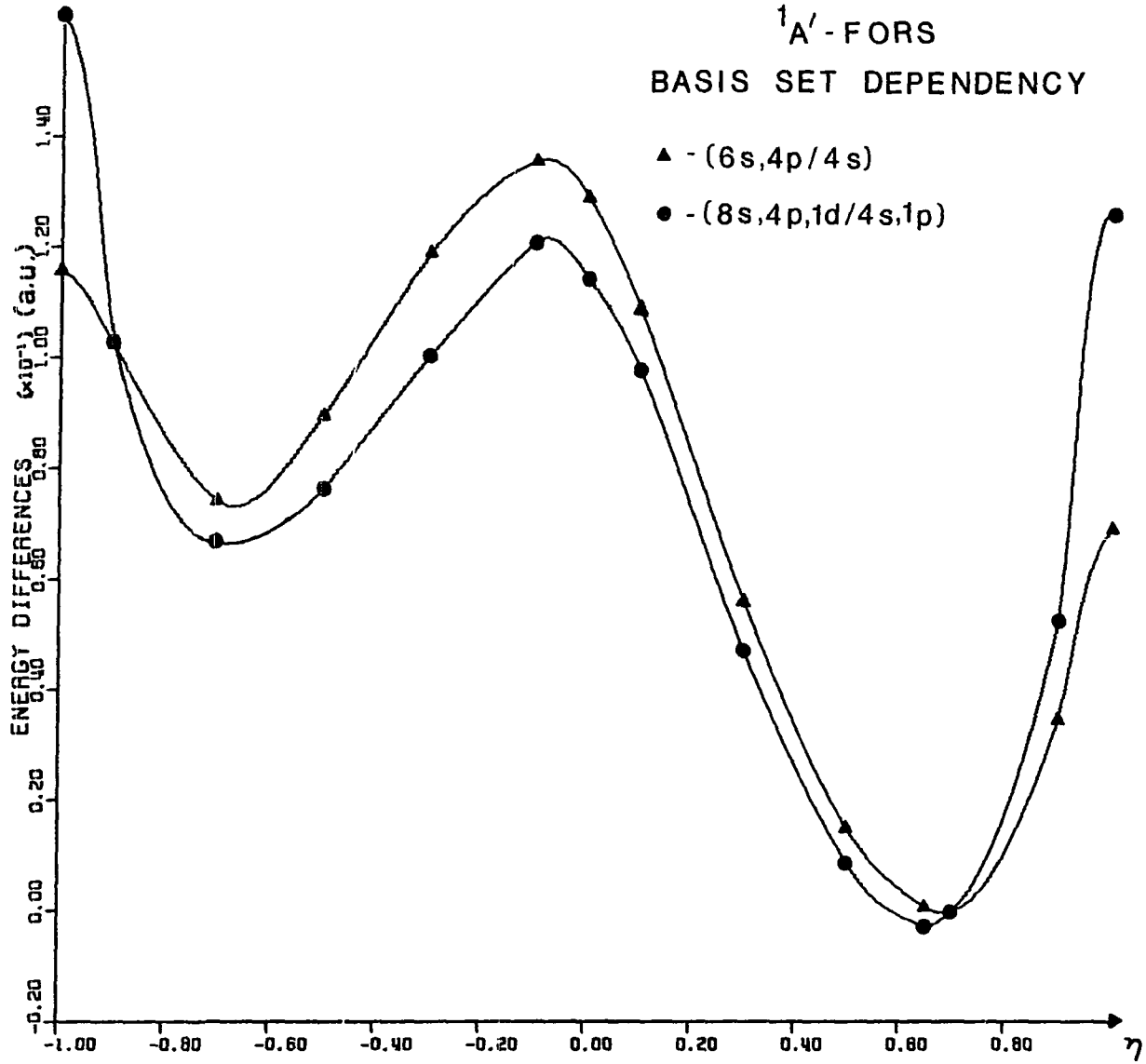


Figure 12a. FORS dependence on the quality of the basis set for the $^1A'$ state

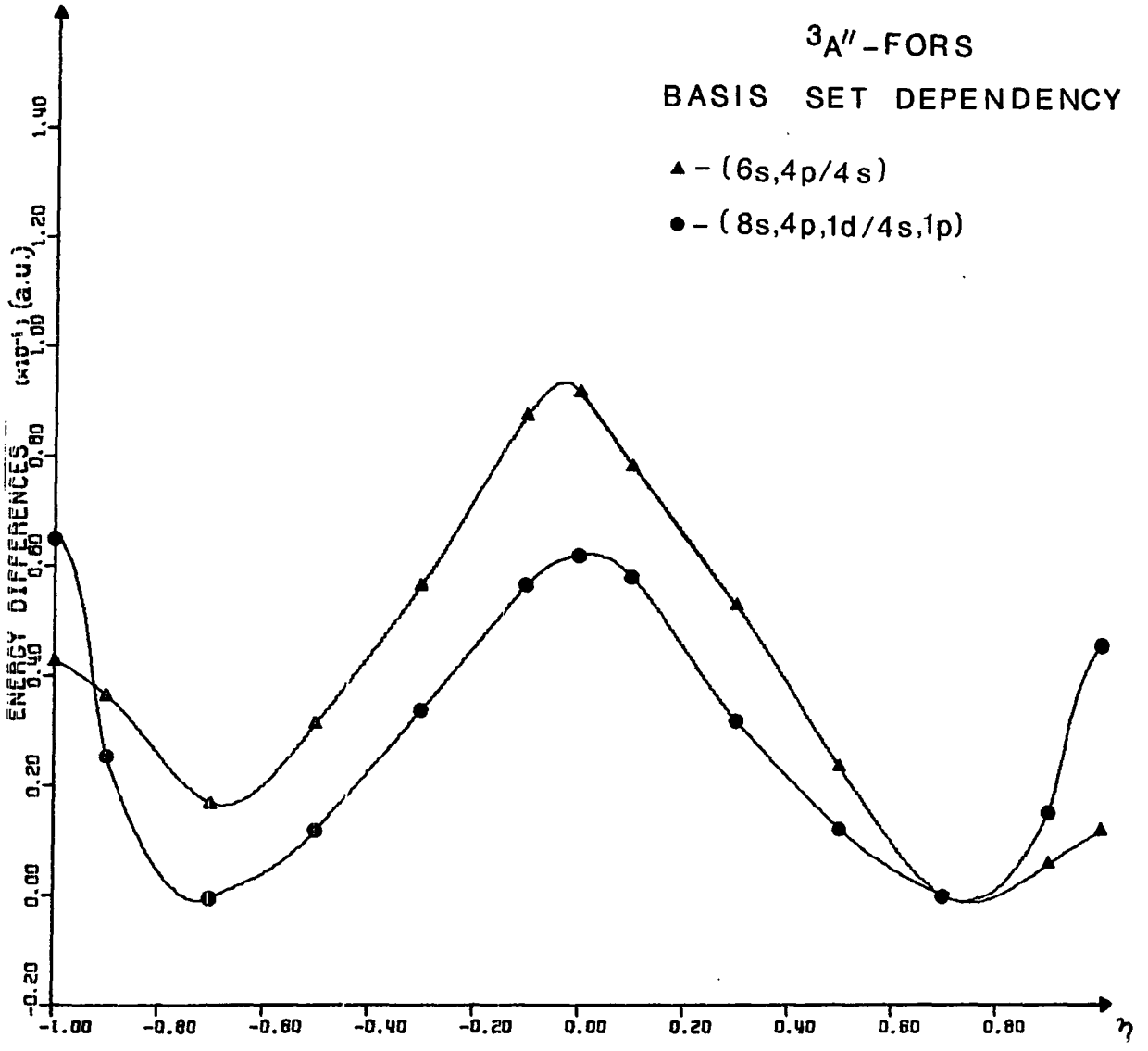


Figure 12b. FORS dependence on the quality of the basis set for the $^3A''$ state

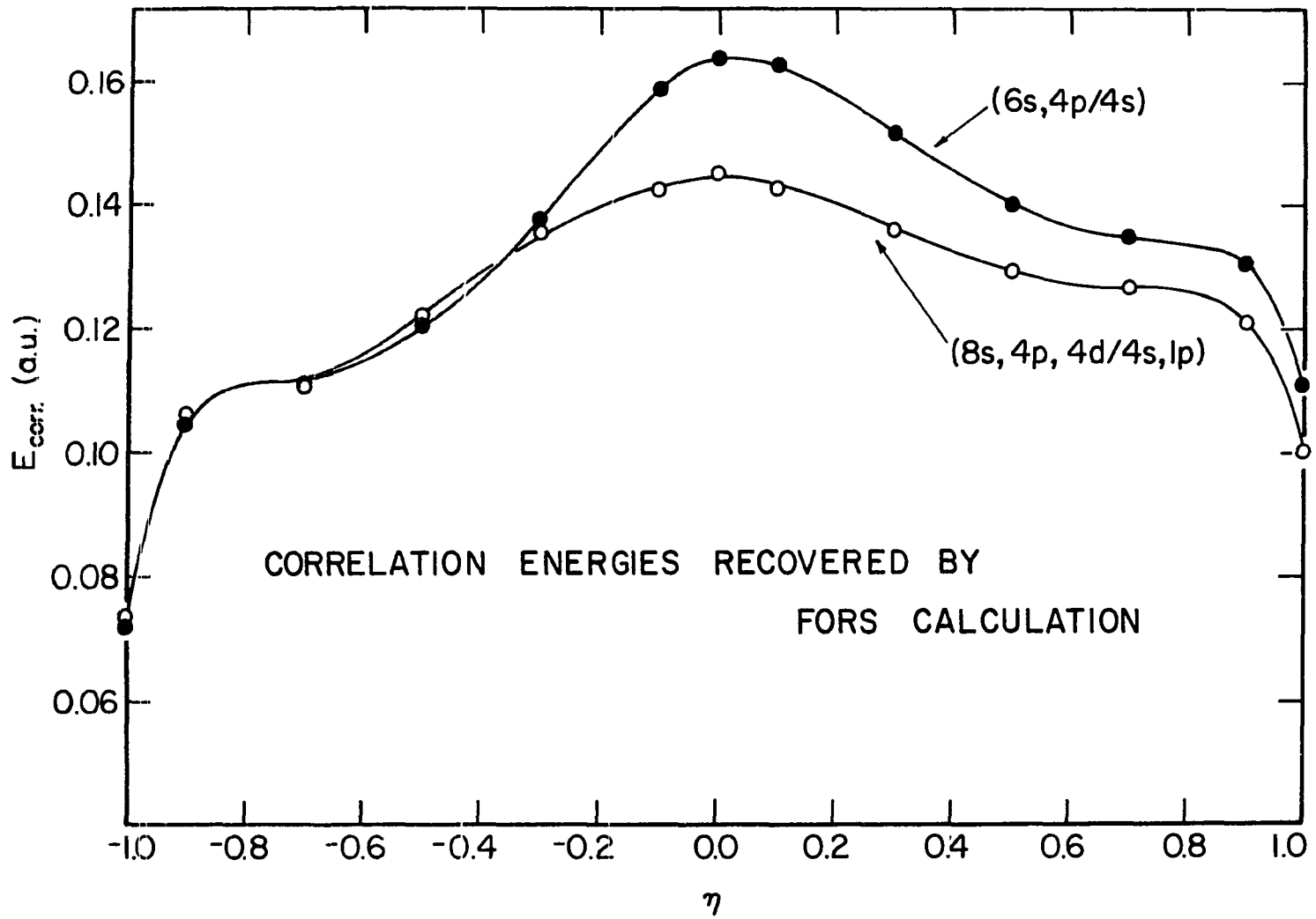


Figure 13. Basis set dependence of the correlation energy recovered in an FORS calculation

near the groundstate geometry for the A' state where the third and fourth most important configurations with expansion coefficients of 0.12 and -0.08 in the small basis do not appear in the top ten configurations for the large basis.

The reason for the unexpected similarity of our FORS results with both basis sets could lie with the use of the conceptual minimal basis set space. In calculations on H_2, LiH and N_2, where an accurate estimate of the entire valence correlation energy was available, this minimal basis set space has recovered from 45%-83% of the valence correlation energy. Since the minimal basis for hydrogen is a single function, compounds containing hydrogen will generally recover a smaller percentage of the correlation energy than compounds without it. An MCSCF-CI calculation by Benioff [34] on NO_2 with the minimal basis set space yielded about 66% of the estimated valence correlation energy. It might be argued that 60%-70% is still too small a fraction to reflect the true effect of altering the size of the basis. On the other hand, perhaps this particular choice of configuration generating orbitals is truly capable of giving quantitative results which are somewhat insensitive to basis set truncation.

A semi-quantitative representation of the groundstate singlet A' FORS energy surface is given in Figure 14. The NO bond distance is fixed at the same value as that chosen for the SCF surface in Figure 8, namely 2.20 bohr. Since an entire grid of points would have been prohibitively expensive, eight

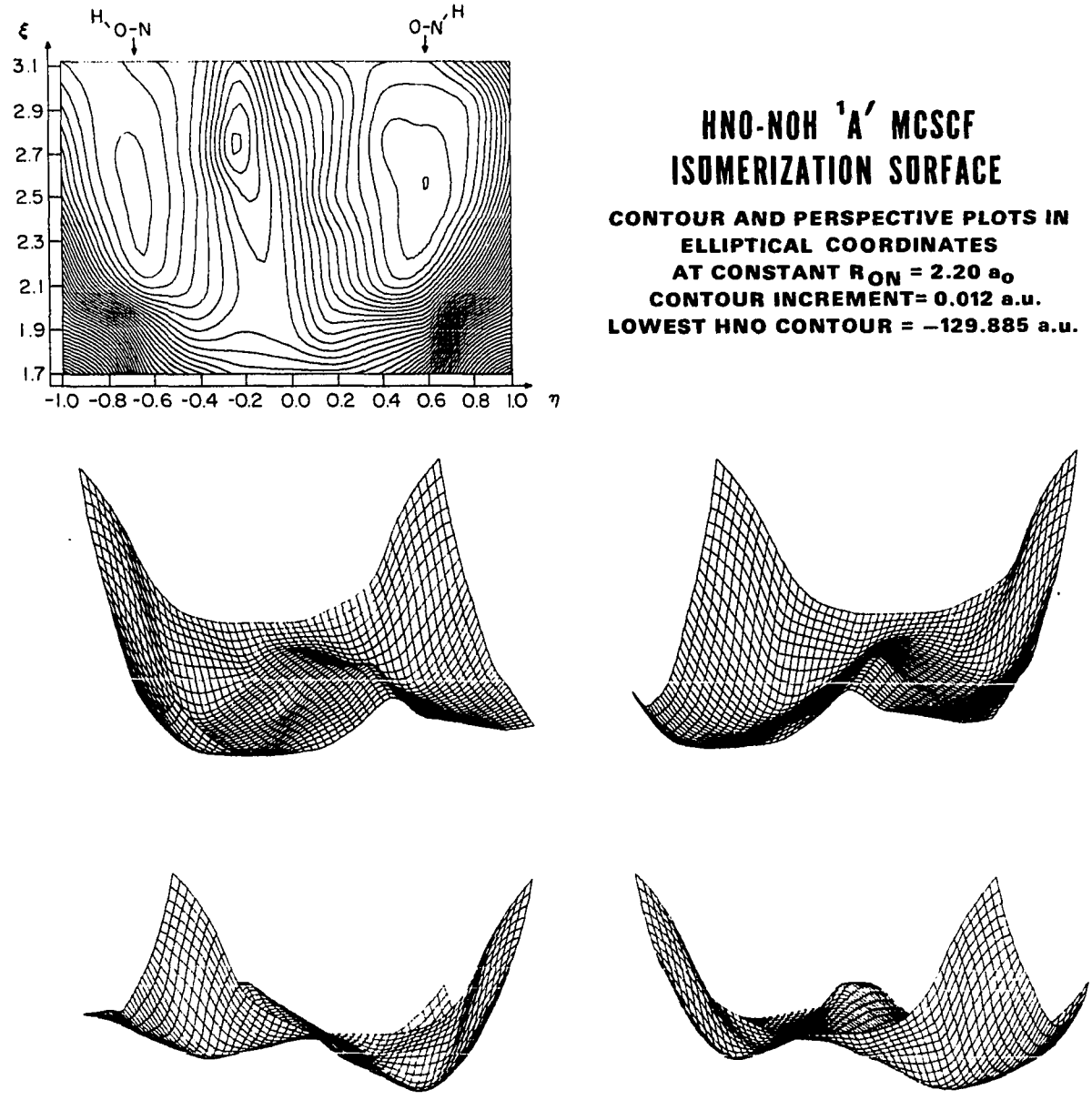


Figure 14. An approximate FORS potential surface for the groundstate

FORS calculations were distributed in a roughly even fashion over the surface. The correlation energy at these eight points was interpolated to obtain an estimate of the correlation energy at the remaining grid points. These estimates were added to the computed SCF energies to fill in the grid. Judging from the magnitude of the change in correlation energy as a function of geometry on this grid, the linear interpolation is probably accurate to ±10 millihartrees. Even though the coefficient of the SCF configuration in the natural orbital expansion of CI wavefunction, which is an indication of the configuration's importance, doesn't drop below 0.9, the differences in the two surfaces are quite substantial. In the transition from an SCF to an MCSCF-CI description of this surface the large barrier separating the HON minimum from the HNO minimum is greatly reduced. More importantly, the barrier to dissociation from HON to H + NO is completely eliminated. The sum of the energy for NO in the identical FORS approximation, given in Table 31 at R = 2.1747 as -129.324 E_h and the energy of the hydrogen atom in this basis, -0.4965 E_h, is -129.821 E_h. This is 0.004 E_h lower than the energy of HON at geometry C listed in Table 27, namely -129.817 E_h.

Inclusion of p-orbitals on the hydrogen in the configuration generating MO's would seem like the next step in alleviating the deficiencies in the formal minimal basis set. This inclusion could substantially lower the HON state with respect to dissociation since such functions would not contribute to

the isolated hydrogen atom. The total correlation energy for a single O-H bond is in the range of 0.23 E_h, from large calculations on the water molecule. The total valence correlation for the hydrogen molecule is 0.041 E_h. A two configuration FORS calculation recovers 45% of this. Adding p orbitals to the configuration generating space increases this to about 75%. If similar behavior could be expected in HON, the molecule would again be predicted to be stable, quite possibly by some tens of kilocalories per mole. The barrier to intramolecular isomerization would be further reduced.

Even with quite large s,p basis sets, such as the (10s,5p), the singlet A' and triplet A" states, as determined by HF-SCF, lie within a few millihartrees of each other near the experimental geometry. The inclusion of polarization functions seems necessary to provide a quantitatively correct spacing of the levels from a HF-SCF calculation. The triplet differs from the closed shell singlet in that one electron from the a' irrep is excited into the a" irrep. However, only one third as many of the polarization functions are of a" symmetry. Thus the ground-state benefits more from the inclusion of polarization functions to the basis.

The experimental vertical excitation energy from the groundstate to the excited singlet is 1.63 eV [26]. Although the phosphorescent emission from the triplet has never been observed, some experimentalists have estimated it to be in the neighborhood of 0.8 eV from certain kinetic considerations. Our

vertical excitation energies are 0.24 eV and 1.05 eV for the triplet and excited singlet at the HF-SCF level and 1.23 eV and 2.16 eV at the FORS level. Both values seem high by about 0.5 eV.

A previous CI calculation of this spectrum by Wu, Buenker and Peyerimhoff [32] would appear fortuitously close, reporting a singlet-singlet separation of 1.60 eV and a singlet-triplet separation of 0.71 eV. This almost certainly results from their particular choice of basis set which yields an SCF separation for these three states comparable to our (8s,4p/4s) basis without polarization functions. The correlation energy recovered by the two basis sets used for the FORS calculations does not change drastically, so that a good estimate of how much correlation energy would be recovered with the (8s,4p/4s) can be made. If this estimate is added to the SCF vertical excitation spectrum for this basis a final separation quite close to the Wu, Peyerimhoff and Buenker value is obtained. The small basis tends to underestimate the energy differences while the larger basis tends to overestimate with respect to the currently accepted experimental value.

D. Bimolecular Reactions

Although extensive efforts have not been made to experimentally observe the HON state, the absence of any experimental evidence for the existence of this geometry, when theoretical results indicate that it should be quite stable with respect to

dissociation to H + NO, suggest the possibility of a mechanism for converting this geometry into the more stable HNO. The intramolecular isomerization has been discussed earlier in this work. Perhaps the next simplest process would involve the conversion of two HON molecules to two HNO facilitated by the formation of some hydrogen bonds.

If one HON molecule is approached by a second HON molecule lying in the same plane as the first and oriented such that the N-O bond of the second is parallel to that of the first with the oxygen atoms across the diagonal from each other the system possesses C_2h symmetry and is in an ideal orientation to simply pass the two protons from one to the other. This involves the breaking of two O-H bonds and the formation of two N-H bonds but no pairs of electrons need be broken apart since the bond pairs simply transform into lone pairs. The reaction is in fact a proton transfer.

Like previous studies of multiple proton transfers [35] we find a multiple minima situation on the energy surface. Figure 15 shows the two minima connected by a barrier of approximately 0.017 E_h (10.7 Kcal/mole). The surface shown in this Figure corresponds to an O-N bond distance of 2.52 bohr and a fixed H(z) of 1.96 bohr. A small (6s,3p/4s) basis was used to generate the SCF surface in this Figure, but the accuracy of the results were checked against the large (8s,4p,1d/4s,1p) basis in the region of the barrier. Energy values on this surface are given in Table 28. The surface was constructed by fitting two

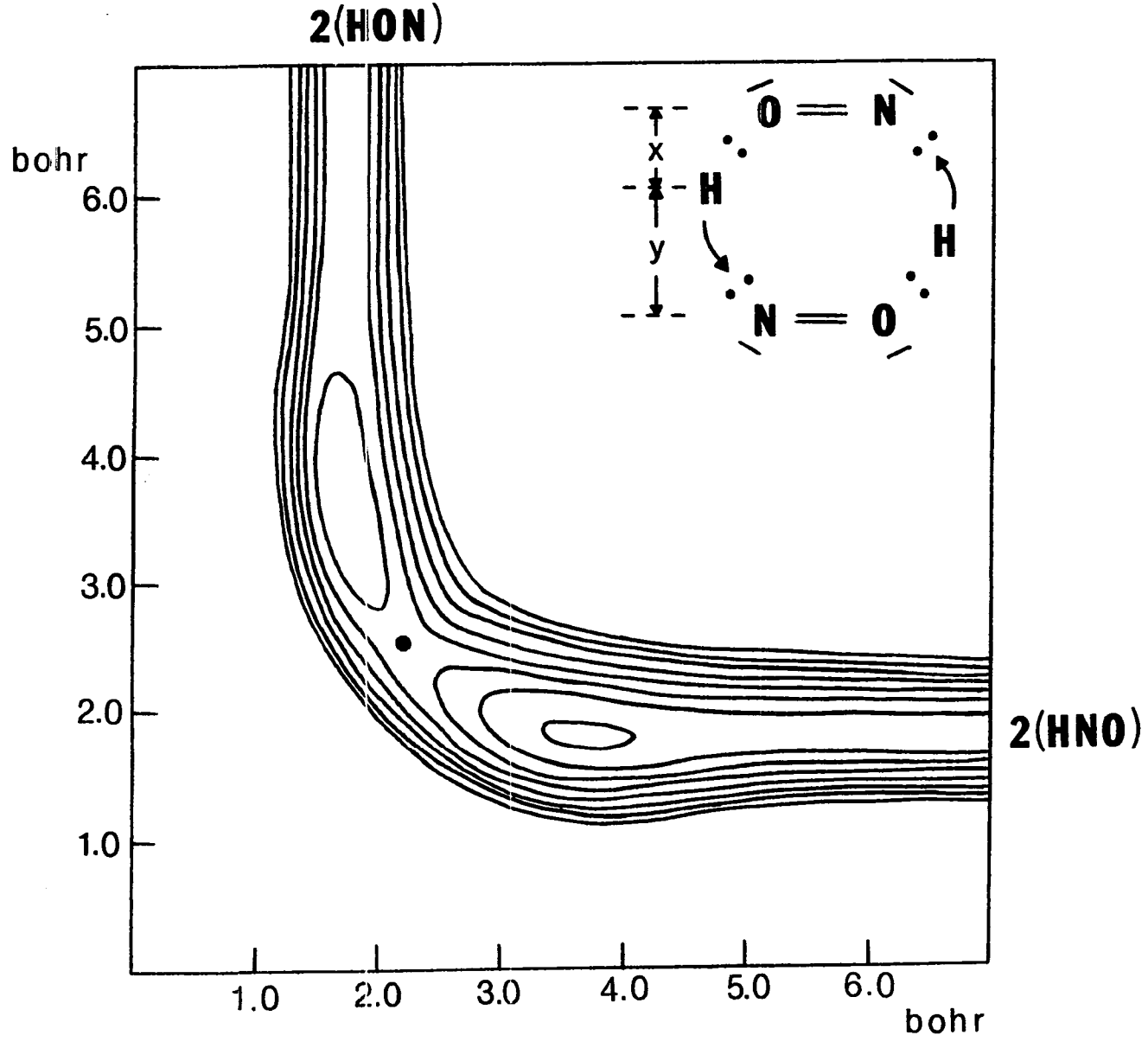


Figure 15. Potential surface for the least motion approach of two HON molecules

Table 28. HON-HON energy differences with respect to the top of the barrier for the surface in Figure 15

R	H(x)	H(y)	Es	El	R	H(x)	H(y)	Es	El
8.0	1.50	6.50	-0.010		5.0	1.50	3.50	0.009	
	1.70	6.30	0.008			1.70	3.30	0.033	0.041
	2.00	6.00	-0.019			2.00	3.00	0.017	
	4.00	4.00	-0.323			2.50	2.50	0.000	0.000
	6.00	2.00	0.032			3.00	2.00	0.044	0.040
	6.20	1.80	0.049			3.20	1.80	0.052	
	6.50	1.50	-0.009			3.50	1.50	0.025	
7.0	1.50	5.50	-0.007		6.0	1.50	4.50	0.000	
	1.70	5.30	0.015			1.70	4.30	0.023	
	2.00	5.00	-0.017			2.00	4.00	-0.002	
	3.50	3.50	-0.219			3.00	3.00	-0.109	
	5.00	1.00	0.038			4.00	2.00	0.044	
	5.20	1.80	0.050			4.20	1.80	0.051	
	5.50	1.50	-0.002			4.50	1.50	0.004	
4.5	1.50	3.00	-0.029		4.0	1.00	3.00	-0.502	
	2.00	2.50	0.002			1.50	2.50	-0.097	
	1.25	1.25	0.006	0.011		2.00	2.00	-0.070	-0.055
	2.50	2.00	0.015			2.50	1.50	-0.095	
	3.00	1.50	-0.017			3.00	1.00	-0.584	

The small (6s,3p/4s) basis gave a barrier of -258.643 Hartrees. The large (8s,4p,1d/4s,1p) basis gave a barrier of -259.367 Hartrees.

overlapping regions of the coordinate plane with a least squares quartic procedure and then piecing the two fits together along the line passing through zero with slope one.

Table 29 gives the optimal geometries of the HON-HON system as the distance between the N-O bonds decreases. All

Table 29. HON-HON optimal geometries and energies

R	R(ON)	H(x)	H(y)	H(z)	Energy	ΔE[a]
Predominantly HON:						
infinite	2.580	infinite	1.69	1.98	-258.6520	32.8
8.0	2.571	6.31	1.69	1.97	-258.6527	32.3
6.0	2.560	4.30	1.70	1.99	-258.6722	20.1
5.0	2.522	3.30	1.70	1.96	-258.6756	17.9
4.7	2.440	2.35	2.35	1.83	-258.6588	28.5
Predominantly HNO:						
4.7	2.440	2.35	2.35	1.83	-258.6588	28.5
5.0	2.381	1.81	3.19	1.88	-258.7102	-3.8
6.0	2.360	1.84	4.16	1.85	-258.7161	-7.5
8.0	2.370	1.83	6.17	1.82	-258.7030	0.8
infinite	2.360	1.84	infinite	1.81	-258.7042	0.0

[a] $\Delta E = E(R) - 2E(HNO)$ in Kcal/mole.

geometrical parameters were optimized for fixed R and under the constraint of C_2h symmetry. H(x) and H(z) are the cartesian coordinates of one hydrogen atom with respect to an origin centered along the N-O bond. The second hydrogen's position is obtained by rotating the position of the first by 180 degrees about the C_2 axis, as seen in Figure 16.

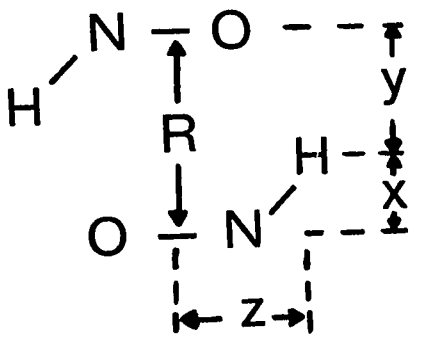


Figure 16. Orientation for HON-HON geometry optimization

The barrier seen in Table 29 is much smaller than that of the intramolecular isomerization and provides a possible conversion mechanism. It should also be noted that the actual energy surface will be somewhat more complex because the singlet A' state in the HON metastable geometry is above the triplet A" and then crosses over to become the groundstate for the HNO geometry. Hence there will be an avoided crossing of the energy surfaces which results from the singlet coupling of the two triplets. There will also be a triplet and a quintuplet coupling for the two triplets, but only the singlet has a possibility of interacting with the singlet calculated here.

However, this approach of the two molecules requires a rather restrictive orientation of the two HON fragments. If the C_2h symmetry is relaxed and we bring the HON's together in such a fashion that the O-H--N-O portion of the system lies in a straight line, another strong hydrogen bond is observed. Table 30 gives the results of a series of SCF calculations with

Table 30. SCF energies for the in-line HON-HON reaction

Distance between HON's	Total Energy
2.00	-258.6032
3.00	-258.6515
4.00	-258.6580
5.00	-258.6547
6.00	-258.6525
8.00	-258.6513

the small (6s,3p/4s) basis for various internuclear distances between the fragments. The geometries of the individual HON units were frozen at the optimal isolated molecule geometries and the two atoms not lying on the line (N and H) were trans to each other. Once the first hydrogen bond is formed, the system has a greater chance of the second hydrogen coming within range to transfer both hydrogens, thus performing the conversion of two HON to two HNO.

E. The Ground State of Nitric Oxide

As indicated in the FORS energy surface of the singlet A' state, dissociation of the HON metastable geometry to H and NO with subsequent formation of HNO appears to be the energetically favored route for uni-molecular isomerization. In order to quantitatively describe this dissociative process it is necessary to compute an FORS wavefunction for the NO molecule using the same (8s,4p,1d) basis set as was used for HNO. Besides its relevance to the HNO system, nitric oxide is of interest in its own right, since it is an important atmospheric emitter. Previous work on this molecule includes a 200 configuration CI calculation using the Iterative Natural Orbital (INO) technique. This was done by Green [36-37] at the experimental equilibrium distance. A 20 configuration optimized valence configuration (OVC) MCSCF calculation followed by a 40 configuration CI along the energy curve was reported by Billingsley [38] using the same large STO basis as Green. Billingsley also reported several expectation values, such as the dipole moment, as a function of the internuclear distance.

To generate the list of configurations needed for the FORS wavefunction we proceed as we did for HNO. Within the formal minimal basis set space of six sigma and four pi orbitals we generate all possible space-spin products consistent with the doublet pi nature of the groundstate. The SCF configuration is

$$^2\pi = (1\sigma)^2 (2\sigma)^2 (3\sigma)^2 (4\sigma)^2 (5\sigma)^2 (1\pi)^4 (2\pi)^1 \qquad (4.2)$$

There are 126 space products and 5 spin functions which yield 252 space-spin products when appropriately combined.

With these 252 configurations we have performed an FORS calculation on NO at the experimental bond distance of 2.1747 bohr with the same quality basis as that used by Green and Billingsley, although the mixture of sigma and pi orbitals was slightly different. Our (14s,7p,2d) ETG primitive basis was contracted to (5s,3p,2d) to give 24 sigma and 10 pi AO basis functions as opposed to 20 sigma and 12 pi in the STO basis of Green. Both sets yielded SCF energies within a millihartree of each other, the ETG energy being -129.289 E_h. An even larger calculation by Green with a near HF limit basis yielded an energy only 5 millihartrees lower. The final FORS energy was -129.406 E_h as contrasted to Billingsley's -129.368 E_h corresponding to an increase of almost 50% in the quantity of correlation energy recovered. Since the FORS energy of the separated atoms lies within less than two millihartrees of the corresponding SCF energy, as shown in the subsequent paragraph, D_e will be the difference between the sum of the atomic SCF energies in this basis, -129.209 E_h, and the molecular FORS energy. This value is 0.197 E_h (5.36 eV) or roughly 80% of the experimental value of 6.61 eV [39]. Billingsley obtained 0.159 E_h (4.33 eV).

However, since we are particularly interested in the dissociation of HON to H and NO we made more detailed calculations with the same basis as was used for the calculations on HON,

namely the (8s,4p,1d) basis. Since one goal of the work by Billingsley was to accurately determine vibrational energy levels up to v=17 we computed a portion of the potential curve with this basis so as to judge the parallelness of the two curves. The sum of the atomic energies in this smaller basis is -129.126 E_h. The largest expansion coefficients from the final natural orbital CI in the FORS calculation and the total energies at the three internuclear distances chosen by Billingsley are given in Table 31 for our (8s,4p,1d) basis. Orbital occupancies for the valence sigma and pi orbitals in the configurations listed assume that the first through third sigma orbitals are doubly occupied.

At a distance of 10 bohr all of the FORS occupied orbitals with the exception of the fifth and sixth sigma orbitals cleanly separate into isolated atomic orbitals on nitrogen and oxygen. These two natural orbitals remain a plus-minus combination of the two p(z) AO's. Since no such separation occurs in Billingsley's wavefunction, his final CI expansion has a great many more nonzero terms than the FORS wavefunction. If the two remaining nonlocalized natural orbitals are replaced by localized p(z) orbitals the FORS-CI expansion would contain only three terms in the dissociation limit. The groundstate potential energy curve is shown in Figure 17.

The difference in energy between R = 2.1747 bohr and R = 3.0747 is 0.127 E_h in the current work and only 0.118 E_h in the

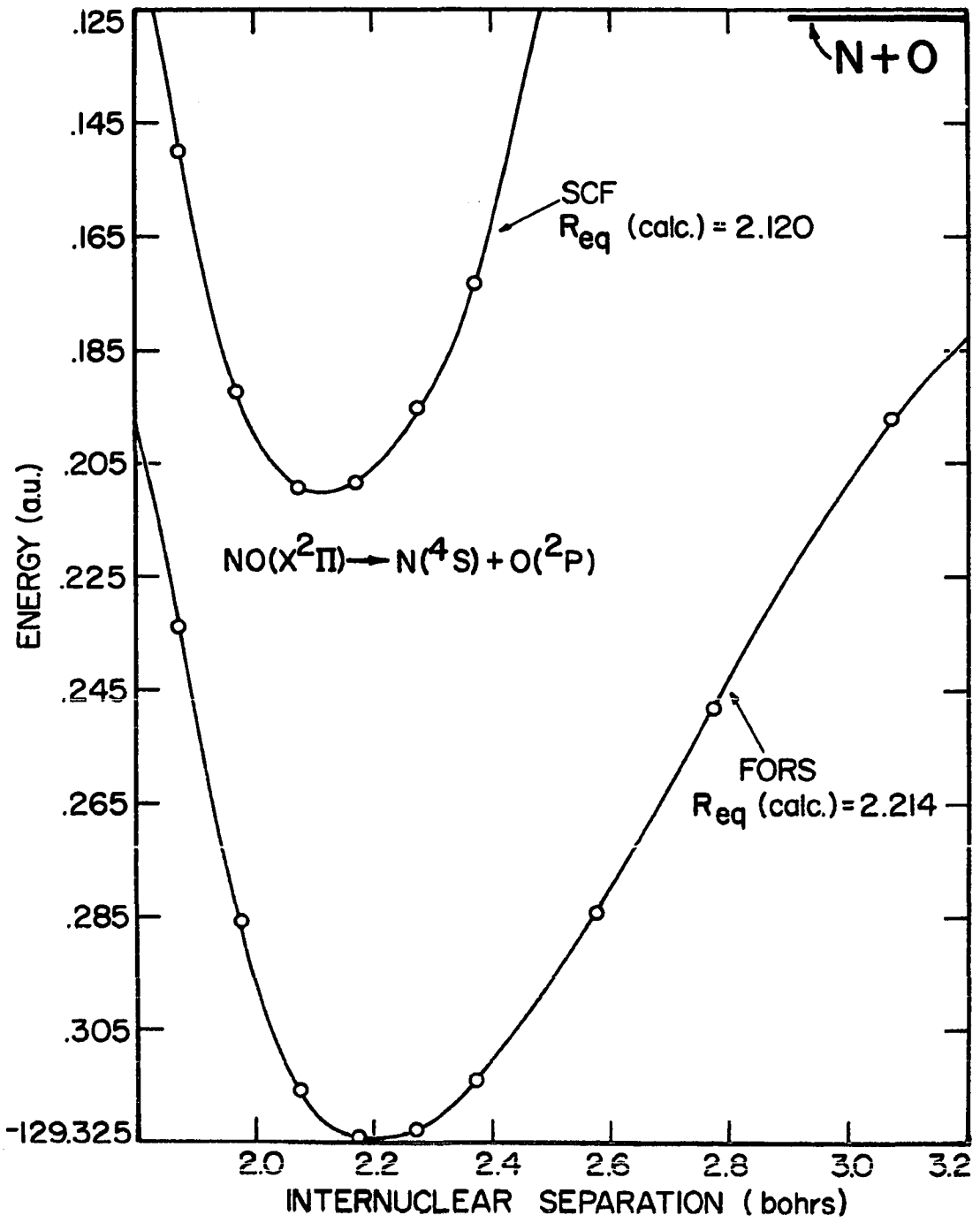


Figure 17. Groundstate potential curves for NO

OVC calculation. This discrepancy could affect the computed vibrational levels.

Table 31. CI expansion coefficients of the leading natural orbitals and total energies for the ground state of NO

Orbital Occupancy						CI Expansion Coefficients		
4s	5s	6s	1π	2π	spin[a]	R=2.1747	R=3.0747	R=10.00
2	2	0	4	1	1	0.966	0.840	0.000
2	0	2	4	1	1	-0.059	-0.130	0.000
2	2	0	2	3	1	-0.143	-0.272	0.000
2	2	0	3	2	1	-0.040	0.258	0.500
					2	-0.011	0.096	0.289
2	0	2	3	2	1	0.002	-0.055	-0.500
					2	0.000	-0.011	-0.289
2	2	0	2	3	1	-0.092	-0.178	0.000
					2	-0.093	-0.099	0.000
2	2	0	1	4	1	0.071	0.105	0.000
2	1	1	3	2	1	0.044	0.058	0.000
					2	0.082	0.117	0.000
					3	0.010	0.056	0.408
					4	-0.010	-0.063	-0.408
0	2	0	4	3	1	-0.059	-0.046	0.000
SCF Energy						-129.208	-128.979	-128.868
FORS Energy						-129.324	-129.197	-129.128

[a]The 5 electron spin functions are defined (using $\theta_- = (\alpha\beta-\beta\alpha)/\sqrt{2}$, $\theta_+ = (\alpha\beta+\beta\alpha)/\sqrt{2}$) as

$\theta_1 = \theta_-\theta_-\alpha$.

$\theta_2 = \theta_-\{\theta_+\alpha-\sqrt{2}\alpha\alpha\beta\}/\sqrt{3}$.

$\theta_3 = \{\theta_+\theta_-\alpha-\sqrt{2}\alpha\alpha\theta_-\beta\}/\sqrt{3}$.

$\theta_4 = \{\alpha\alpha\beta\beta+\beta\beta\alpha\alpha - 1/2\ \theta_+\theta_+\}\alpha/\sqrt{3}$.

The dissociation energy for the (8s,4p,1d) basis is essentially the same as that obtained with the larger (14s,7p,2d) basis. The computed equilibrium distance is 2.214 bohr, slightly too large. Most of this error would be eliminated by using the larger basis since the calculated equilibrium distance at the SCF level is shorter by 0.02 bohr for this basis.

V. LITERATURE CITED

1. C. C. J. Roothaan, Rev. Mod. Phys. 23, 69 (1951).

2. C. Schwartz, in Methods in Computational Physics, edited by B. Alder (Academic, New York, 1963), Vol. 2, p. 241.

3. K. Ruedenberg, R. C. Raffenetti, and R. D. Bardo, "Energy Structure and Reactivity", Proceedings of the 1972 Boulder Conference on Theoretical Chemistry (Wiley, New York, 1973), p. 164.

4. S. Huzinaga, Approximate Atomic Functions (University of Alberta Press, 1971).

5. R. Kikuchi, J. Chem. Phys. 22, 148 (1954).

6. D. M. Bishop and R. L. Somorjai, J. Math. Phys. 11, 1150 (1970).

7. W. J. Taylor, J. Math. Phys. 19, 52 (1978).

8. R. C. Raffenetti, J. Chem. Phys. 59, (1972).

9. R. D. Bardo and K. Ruedenberg, J. Chem. Phys. 59, 5956 (1973).

10. R. C. Raffenetti, Int. J. Quantum Chem. 9, 289 (1975).

11. M. W. Schmidt, Dept. of Chemistry, I. S. U., (unpublished).

12. M. W. Schmidt and K. Ruedenberg, (to be published).

13. C. F. Fischer, in Some Hartree-Fock Results for Atoms Helium to Radon (University of British Columbia Press, Vancouver B. C., 1968).

14. C. F. Fischer, The Hartree-Fock Method for Atoms: A Numerical Approach (Interscience, New York, 1977).

15. K. Ruedenberg, R. D. Bardo, and L. M. Cheung, (to be published).

16. R. C. Raffenetti, J. Chem. Phys. 58, 4452 (1973).

17. T. H. Dunning, J. Chem. Phys. 53, 2823 (1970).

18. R. K. Nesbet, J. Chem. Phys. 36, 1518 (1962).

19. R. Poirier and R. Kari, Can. J. Chem. 56, 543 (1978).

20. S. Green, J. Chem. Phys. 52, 3100 (1970).

21. E. Clementi, Proc. Nat. Acad. Sci. USA 69, 2942 (1972).

22. H. Popkie and J. Kaufman, Int. J. Quantum Chem. 10, 569 (1976),

23. H. Kashiwasa and T. Takada, Int. J. Quantum Chem. 22, (1977).

24. R. C. Raffenetti and K. Ruedenberg, J. Chem. Phys. 59, 5950 (1973).

25. M. Dombek, Ph.D. Thesis, Iowa State University (1977) (unpublished).

26. T. Ishiwata, H. Akimoto, and I. Tanaka, Chem. Phys. Letters 27, 260 (1974).

27. L. M. Cheung, S. T. Elbert, and K. Ruedenberg, Int. J. Quantum Chem. (submitted) (1978).

28. F. W. Dalby, Can. J. Phys. 36, 1336 (1958).

29. J. E. Williams and J. N. Murrell, J. Amer. Chem. Soc. 93, 7149 (1971).

30. A. W. Salotto and L. Burnelle, J. Chem. Phys. 52, 2936 (1970).

31. G. A. Gallup, Inorg. Chem. 14, 563 (1975).

32. A. A. Wu, S. D. Peyerimhoff, and R. J. Buenker, Chem. Phys. Letters 35, 316 (1975).

33. P. C. Hariharan and J. A. Pople, Theor. Chim. Acta (Berlin) 28, 213 (1973).

34. P. A. Benioff, J. Chem. Phys. 68, 3405 (1978).

35. S. Scheiner and C. W. Kern, Theoretical Study of Proton Transfers Between Base Pairs of DNA (to be published).

36. S. Green, Chem. Phys. Letters 13, 552 (1972).

37. S. Green, Chem. Phys. Letters 23, 115 (1973).

38. F. P. Billingsley II, J. Chem. Phys. 62, 864 (1975).

39. A. G. Gaydon, Energies and Spectra for Diatomic Molecules, 3rd Ed., (Chapman and Hall, London, 1968).

VI. ACKNOWLEDGEMENTS

The author wishes to thank the many people who have contributed, either directly or indirectly, to this work. My special appreciation goes to my family and friends in Dubuque, Iowa for bearing those early years of scientific exuberance and to a gifted, but sometimes cantankerous, high school chemistry teacher, Sr. Marjorie Moran. Appreciation is also extended to my wife, Marianne, who many times softened the disappointments and disillusionments of the past few years with her presence.

Essential computer programs and software assistance were provided by Drs. Richard Raffenetti, Lap Cheung, Ken Sundberg and Steve Elbert. Dr. Elbert's proficiency at improving the performance of existing programs and his concern for making the user's task as simple as possible was both greatly appreciated and is an indication of his excellence. The perspective plots were drawn with a graphics package provided by Drs. Nelson Beebe and Harold McIntosh.

The author is grateful to Mike Schmidt for numerous helpful discussions and an uncanny ability to spot input errors. Dr. Klaus Ruedenberg is thanked for his instruction over the years and his guidance in preparing this work.

Finally, Ms. Troi Tranter is to be thanked for the professional job she did in typing this dissertation and in anticipating potential sources of trouble.