On the optimization of Gaussian basis sets

George A. Petersson and Shijun Zhong

Hall-Atwater Laboratories of Chemistry, Wesleyan University, Middletown, Connecticut 06459-0180

John A. Montgomery, Jr. and Michael J. Frisch

Gaussian, Inc., North Haven, Connecticut 06473

(Received 21 June 2002; accepted 3 September 2002)

A new procedure for the optimization of the exponents, \( a_j \), of Gaussian basis functions, \( Y^m_l(\theta, \phi) r^l e^{-a_j r^2} \), is proposed and evaluated. The direct optimization of the exponents is hindered by the very strong coupling between these nonlinear variational parameters. However, expansion of the logarithms of the exponents in the orthonormal Legendre polynomials, \( P_k \), of the index, \( j \): 

\[
\ln a_j = \sum_{k=0}^{\infty} A_k P_k((2j-2)/(N_{\text{prim}}-1) - 1),
\]

yields a new set of well-conditioned parameters, \( A_k \), and a complete sequence of well-conditioned exponent optimizations proceeding from the even-tempered basis set (\( k_{\text{max}} = 1 \)) to a fully optimized basis set (\( k_{\text{max}} = N_{\text{prim}}-1 \)). The error relative to the exact numerical self-consistent field limit for a six-term expansion is consistently no more than 25% larger than the error for the completely optimized basis set. Thus, there is no need to optimize more than six well-conditioned variational parameters, even for the largest sets of Gaussian primitives © 2003 American Institute of Physics. [DOI: 10.1063/1.1516801]

I. INTRODUCTION

Virtually all \textit{ab initio} electronic structure calculations employ expansions in basis sets of atomic orbitals. Current efforts to expand the scope of such calculations to heavier elements and include relativistic corrections require new basis sets of atomic orbitals. It is therefore timely to reconsider the problem of basis set optimization.

The expansion of atomic self-consistent field (SCF) orbitals in increasing numbers of Gaussian functions:

\[
\psi(r, \theta, \phi) = Y^m_l(\theta, \phi) \sum_{j=1}^{N_{\text{prim}}} C_j r^l e^{-a_j r^2}
\]

provides a systematic sequence of approximations, but optimization of the exponents of these Gaussian functions becomes progressively more difficult as the number, \( N_{\text{prim}} \), increases. Jensen has observed \(^2\) that: “In some cases,... optimization resulted in variational collapse, i.e. two basis exponents having the same \( l \) value converge toward the same \([\alpha]\) value, causing the overlap matrix to become singular.” The problem here is that the small energy improvement attainable by varying one of the exponents is exceeded by the effect of round-off error as the two exponents converge. Thus the exponent optimization problem not only grows in size, but also becomes ill-conditioned.

An early attempt to circumvent this problem was the introduction of even-tempered basis sets by Bard and Ruedenberg. \(^3\) The exponents are constrained to form a geometric sequence, \( \alpha_j = \alpha_0 \beta^{j-1} \) \((j=1, 2, \ldots, N_{\text{prim}}) \), with only two variational parameters, \( \alpha_0 \), and \( \beta \). This not only dramatically reduces the dimension of the search space for the optimization, but also completely eliminates the ill-conditioning. The overlap integral between two consecutive members of an even-tempered basis set is

\[
S_{j,j+1} = \left[ \frac{2 \sqrt{\beta}}{1+\beta} \right]^{(2j+3)/2}.
\]

These even-tempered basis sets span the one-electron Hilbert space uniformly (i.e., \( \cos \theta_j \) is constant) and provide a good first-order approximation, but the deviation from the SCF limit can be ten times the error for the same number of fully optimized Gaussian primitives (vide infra). Huzinaga and co-workers addressed this deficiency by introducing two additional parameters, \( \gamma \) and \( \delta \), with their well-tempered basis sets. \(^4-6\)

\[
\alpha_j = \alpha_0 \beta^{j-1} \left[ 1 + \gamma (j/N_{\text{prim}})^{\delta} \right], \quad j = 1, 2, \ldots, N_{\text{prim}}.
\]

The deviation of the well-tempered basis sets from the SCF limit is typically between two and three times the error for the same number of fully optimized Gaussian primitives. This is a substantial advantage over the even-tempered basis sets, but there is still room for improvement.

The very broad range of Gaussian exponents included in a basis set makes it convenient to deal with the \( \ln \alpha \) rather than the exponents themselves. The even-tempered basis set is then a linear function of the index, \( j \):

\[
\ln \alpha_j = \ln \alpha_0 + j \ln \beta, \quad j = 1, 2, \ldots, N_{\text{prim}}.
\]

A comparison of even-tempered and well-tempered basis sets with the corresponding fully optimized basis set for krypton (Fig. 1) suggests that the well-tempered function in Eq. (3) is not the optimum four parameter generalization of the two parameter even-tempered function in Eq. (4). The figure clearly indicates a definite curvature of the optimum function (i.e., an increased spacing between exponents) at large and small \( \alpha \), whereas Eq. (3) only introduces this modification for the large \( \alpha \).
II. EXPANSION OF $\ln \alpha_j$ IN LEGENDRE POLYNOMIALS

The even-tempered function in Eq. (4) can be generalized to an expansion of $\ln \alpha_j$ as a polynomial in the index, $j$:  

$$\ln \alpha_j = \sum_{k=0}^{k_{\text{max}}} B_k j^k.$$  

(5)

Setting $k_{\text{max}}$ equal to 1 gives an even-tempered basis set. Setting $k_{\text{max}}$ equal to 3 gives a four-parameter function (Fig. 2) that better approximates fully optimized exponents than the well-tempered function of Eq. (3) does. The deviation of these third-degree polynomial basis sets from the SCF limit is generally no more than 60% larger than the error for the same number of fully optimized Gaussian primitives (vide infra).

Klobukowski examined polynomials of degree four and found a substantial improvement over well-tempered basis sets for neon and argon. He then generated basis sets for H–Ne and Al–Ar, but he did not pursue the general polynomial expansion in Eq. (5), with other values for $k_{\text{max}}$.

Increasing $k_{\text{max}}$ to $N_{\text{prim}} - 1$ gives a set of fully optimized exponents, $\{\alpha_j\}$. However, if we examine these higher values of $k_{\text{max}}$, we discover that we have traded one ill-conditioned problem for another even more ill-conditioned optimization. The overlap integral between two successive powers of $j$ is

$$f_0^{N_{\text{prim}} j^n j^{n-1} dj} \left[ f_0^{N_{\text{prim}} j^n j^{n-1} dj} \right]^{1/2} = \sqrt{1 - \frac{1}{4n^2}}.$$  

(6)

For example, the overlap of $j^{20}$ with $j^{19}$ is 0.999 687 5, making $B_{19}$ almost redundant with $B_{19}$. This leads to a divergent sequence of large oscillating (i.e., $B_{k} \sim -B_{k-1}$) coefficients in Eq. (5).

Mapping the indices, $j = 1, 2, ..., n$, onto the interval $-1$ to 1, permits a well-conditioned expansion in the orthonormal Legendre polynomials, $P_k$:

$$\ln \alpha_j = \sum_{k=0}^{k_{\text{max}}} A_k P_k \left( \frac{2j - 2}{N_{\text{prim}} - 1} \right)$$  

with the $k_{\text{max}} + 1$ variational parameters, $\{A_k\}$. The two-term polynomial expansion ($k_{\text{max}} = 1$) again corresponds to an even-tempered basis set. We now have a complete sequence of well-conditioned exponent optimizations proceeding from the even-tempered basis set ($k_{\text{max}} = 1$) to a fully optimized basis set ($k_{\text{max}} = N_{\text{prim}} - 1$). Linear dependencies are impossible and the problem of “variational collapse” has been eliminated—exponents $\alpha_j$ and $\alpha_{j+1}$ do not cross (vide infra).

The general patterns of convergence for these expansions of $\ln \alpha$ in Legendre polynomials are very consistent from one case to another. The krypton atom provides a good representative case to illustrate these patterns. The even-tempered basis set is of course the fully optimized basis set if we use only two primitives. However, if we include enough Gaussian primitives to be useful for electronic structure calculations, the dependence of $\ln \alpha_j$ on the index, $j$, quickly develops the curvature seen in Fig. 2. The general shape of the function becomes almost independent of the size of the basis set [Fig. 3(a)]. In fact, if we rescale both axes, the curves can almost be superimposed [Fig. 3(b)]. The very
large exponents required for the cusp and the region around the nucleus are not needed for either $p - \lim_{r \to 0} \varphi_p(r) = r + \cdots$ or $d - \lim_{r \to 0} \varphi_d(r) = r^2 + \cdots$ functions. However, the exponents of higher angular momentum functions exhibit the same basic shape of $\ln \alpha_j$ versus $j$, but the curvature becomes progressively more pronounced—especially for the smaller exponents (Fig. 2). We can therefore focus on the $s$ functions for our illustrative examples.

**A. Convergence of $E(N_{\text{prim}})$ with $k_{\text{max}}$**

The SCF energy obtained with an even-tempered basis is equal to the energy obtained with fully optimized exponents if we use only two Gaussian primitives. However, for the $(15s11p6d)$ krypton basis set, the deviation of the even-tempered basis set from the SCF limit is almost four times the error for the same number of fully optimized Gaussian primitives. The deviation of the $(24s17p12d)$ even-tempered basis set from the SCF limit grows to more than ten times the error for the same number of fully optimized Gaussian primitives (Fig. 4).

A four-term polynomial is exact for four primitives and deviates from the SCF limit by only 30% more than the fully optimized $(15s11p6d)$ krypton basis set, but this error grows to more than 50% for the $(24s17p12d)$ krypton basis set (Fig. 4). The deviation of a six-term polynomial from the SCF limit remains less than 25% from the $(15s11p6d)$ krypton basis set through the $(18s13p8d)$, $(21s15p10d)$, and $(24s17p12d)$ krypton basis sets. The six-term polynomial (i.e., through $P_5$) is therefore our choice for the standard expansion of $\ln \alpha$ in Legendre polynomials. Further optimization of the Gaussian exponents will substantially increase the effort, while giving an energy improvement that is of no practical value (Fig. 4). Since we plan to construct sequences of basis sets to be used for extrapolation to the SCF limit, increasing the SCF error by a consistent factor of $\sim 20\%$ for each of these basis sets will have an insignificant effect on the extrapolated atomic SCF energy. The effect on molecular energies is insignificant, even if we do not extrapolate (vide infra).

**B. Variation of $A_k$ with $k_{\text{max}}$**

The coefficients, $\{A_k\}$, of the first eight Legendre polynomials, $\{P_k\}$, tend to occur in pairs of comparable magnitude (Fig. 5), but opposite sign if we list the exponents in decreasing order. For example, the variationally optimized coefficients, $A_0$ and $A_1$, of $P_0$ and $P_1$ are 7.56 and $-5.68$. 

---

**Fig. 3.** The qualitative behavior of the optimized Gaussian $s$ exponents is unchanged between the $(15s11p6d)$ and $(33s23p18d)$ Kr basis sets. (a) The shape of the $\ln \alpha_j$ vs $j$ curve is quite similar for these two very different basis sets. (b) If we rescale both axes, the variation of the $s$ exponents for the $(15s11p6d)$ and $(33s23p18d)$ Kr basis sets are almost identical.

**Fig. 4.** The four fully optimized Kr basis sets, $(15s11p6d)$, $(18s13p8d)$, $(21s15p10d)$, and $(24s17p12d)$, give progressively smaller SCF errors indicated by the four dashed horizontal lines (—). Partial optimization of each of these four basis sets approaches the limit for that number of primitives as $k_{\text{max}}$ in Eq. (7) increases to the maximum value, $N_{\text{prim}} - 1$. The first point ($k_{\text{max}} = 1$) for each size of basis set represents an even-tempered basis set.
and the coefficients, $A_2$ and $A_3$, of $P_2$ and $P_3$ are 0.41 and −0.31 for the $s$ exponents of a (15s11p6d) krypton basis set. They show very little variation with the number of terms in the expansion, $k_{\text{max}}$. This has the practical advantage of allowing us to first optimize $A_0$ and $A_1$ with $k_{\text{max}} = 1$; then optimize $A_2$ and $A_3$ with $k_{\text{max}} = 3$, making only minor adjustments to $A_0$ and $A_1$; next optimize $A_4$ and $A_5$ with $k_{\text{max}} = 5$, making only minor adjustments to $A_0$, $A_1$, $A_2$, and $A_3$;

e etc. Expansions with more than 8 terms lose this systematic pattern and the smaller coefficients (i.e., $A_8$, $A_9$, etc.) become strongly dependent on the length of the expansion, making the approach to full optimization of larger basis sets much more difficult. Even though we have removed the ill-conditioning, the higher-order coefficients are still strongly coupled.

It is the dominance of $A_0$ and $A_1$ that prevents "variational collapse." If the magnitude of $A_3$ were comparable to $A_1$, the behavior shown in Fig. 6 would be possible. Hence, it is important to optimize $A_0$ and $A_1$ before introducing $A_2$ and $A_3$. The expansion of $\ln a$ in Legendre polynomials avoids "variational collapse" by coupling such behavior to a complete mismatch to the behavior of optimum exponents (Fig. 6). A polynomial of degree 3 or higher in the index, $j$, can accommodate "variational collapse" (i.e., $\alpha_j = \alpha_{j+1}$), but to do so it must differ drastically from the optimum set of exponents, $\{\alpha_j\}$. If we have a reasonable initial guess, variational optimization of the Legendre coefficients, $A_k$ in Eq. (7), will not increase the deviation of the $\alpha_j$ from their optimum values.

C. Variation of $A_k$ with $N_{\text{prim}}$

The coefficients, $\{A_k\}$, of the first six Legendre polynomials, $\{P_k\}$, show a more substantial variation with the number, $N_{\text{prim}}$, of primitive Gaussian functions in the basis set (Fig. 7). The most efficient way to optimize a sequence of basis sets appears to be to first optimize the coefficients, $\{A_k\}$, for the smallest basis set as described in the preceding.
section. Then, using these coefficients as an initial guess, reoptimize $A_0$ and $A_1$, followed by $A_2$ and $A_3$, and finally $A_4$ and $A_5$ for the next larger basis set.

The optimized exponents themselves are not very useful as an initial guess for constructing a larger basis set. The larger exponents tend to interleave so as to be midway between exponents of the basis set with one fewer primitives.\textsuperscript{1,2} Although the smaller exponents often decrease only moderately with increasing numbers of primitives, occasionally a new smaller exponent appears, making the smaller exponents unreliable as an initial guess.\textsuperscript{1,2}

D. Variation of $A_k$ with atomic number, $Z$

The coefficients, $\{A_k\}$, of the first six Legendre polynomials, $\{P_k\}$, show systematic variations with the atomic number, $Z$, for the ground states of the neutral atoms (Fig. 8). Beyond hydrogen and helium, all atoms show the pairing noted in Fig. 5. The smooth overall general patterns are complicated by the filling of shells and the increasing size of the basis set that is required to achieve a given accuracy (~1 mhartree/electron for Fig. 8). As one would expect, the constant term, $A_0$, increases with increasing nuclear charge, reflecting the contraction of the core electrons. The linear term, $A_1$, becomes increasingly negative, but this is partially an effect of increasing the size of the basis set (Fig. 8). Note that $A_1$ is virtually the same for carbon and nitrogen using nine Gaussian $s$ primitives, and then becomes slightly more negative, but is again virtually the same for fluorine and neon with ten Gaussian $s$ functions. This pattern is repeated for the “quadratic” and “cubic” terms, $A_2 P_2$ and $A_3 P_3$, but the “quartic” term, $A_4 P_4$, tends to decrease as a shell is filled and increase with principal quantum number (Fig. 8). The coefficients of the higher-order Legendre polynomials do not show any obvious trend other than to be small. These patterns beyond $A_0$ and $A_1$ are of no practical importance, since the above-described optimization algorithm requires only an initial guess for $A_0$ and $A_1$.

E. Relativistic effects

The effects of relativistic contraction on the basis set optimization can be illustrated by comparison of Douglas–Kroll–Hess relativistic SCF functions\textsuperscript{9,10} with their nonrelativistic Hartree–Fock counterparts for the Xe atom (Fig. 9). The relativistic basis set requires additional $s$, $p$, and $d$ functions to achieve the same accuracy (~1 mhartree/electron). As one would expect these added functions have the large exponents required to modify the behavior near the nucleus (Fig. 9).

III. RESULTS

The quantitative performance for the expansion of exponents in Legendre polynomials can be evaluated by several criteria. First, we compare the four-parameter Legendre expansion with the corresponding well-tempered basis set. Next, we compare the six-parameter Legendre expansion to full optimization of the same basis set. Then, we examine the convergence to the exact SCF limit. Finally, we consider a case for which conventional optimization failed. The optimizations reported all employ a modified Newton–Raphson method in which gradients are evaluated numerically from finite differences. The Hessian, second derivative matrix, is initialized with diagonal second derivatives and is subse-
A. Well-tempered versus four-parameter Legendre expansions

Huzinaga, Klobukowski, and Tatewaki reported\(^5\) the energy obtained from a krypton (16s13p8d) well-tempered basis set for which the \(p\) and \(d\) exponents were constrained to be a subset of the \(s\) exponents. We have optimized this well-tempered basis set without this constraint and find the energy of the well-tempered (16s13p8d) krypton basis set is again about 8 mhartree above that for the fully optimized exponents (Table I). Optimization of the (16s13p8d) Kr basis set with a four-parameter expansion of the \(ln\ a\) in Legendre polynomials reduces this error to less than 4 mhartree, with no increase in computational effort. This pattern is repeated with the (14s11p6d) and (21s16p12d) krypton basis sets (Table I). Optimization of all three krypton basis sets using four-term expansions in Legendre polynomials yields energies with errors 40%, 63%, and 57% larger than those for the corresponding fully optimized basis sets (Table I). Thus, even truncating the Legendre expansion with only four terms introduces an error that is consistently smaller than the inherent error from the finite size of the basis set.

The constraint of the \(p\) and \(d\) exponents to be a subset of the \(s\) exponents is significantly more restrictive than the use of a Legendre expansion of degree 3 (Table I), and thus seems ill-advised. We have therefore allowed complete independence for the basis functions of differing angular momentum.

TABLE I. Comparison of SCF energies (hartree a.u.) from even- and well-tempered basis sets with those from optimized Legendre expansions and fully optimized Gaussian basis sets for a krypton atom.

<table>
<thead>
<tr>
<th>Optimization</th>
<th>Constraint</th>
<th>Basis set</th>
<th>14s11p6d</th>
<th>16s13p8d</th>
<th>21s16p12d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Well-tempered</td>
<td>(a_s = a_p = a_d)</td>
<td></td>
<td>(-2752.032 4^a)</td>
<td>(-2752.015 436)</td>
<td>(-2752.052 176)</td>
</tr>
<tr>
<td>Fully optimized</td>
<td></td>
<td></td>
<td>(-2752.040 259^a)</td>
<td>(-2752.039 988)</td>
<td>(-2752.054 332)</td>
</tr>
<tr>
<td>Even-tempered</td>
<td>None</td>
<td></td>
<td>(-2751.871 434)</td>
<td>(-2752.015 436)</td>
<td>(-2752.052 176)</td>
</tr>
<tr>
<td>Well-tempered</td>
<td></td>
<td>(\Sigma A_k P_k, k \leq 3)</td>
<td>(-2751.978 106)</td>
<td>(-2752.039 988)</td>
<td>(-2752.054 332)</td>
</tr>
<tr>
<td>Fully optimized</td>
<td></td>
<td>(\Sigma A_k P_k, k \leq 5)</td>
<td>(-2751.997 801)</td>
<td>(-2752.045 453)</td>
<td>(-2752.054 583)</td>
</tr>
<tr>
<td>Error ratio(^6)</td>
<td></td>
<td></td>
<td>1.398</td>
<td>1.634</td>
<td>1.570</td>
</tr>
<tr>
<td>Exact SCF limit</td>
<td></td>
<td></td>
<td>(-2752.054 977)</td>
<td>(-2752.054 977)</td>
<td>(-2752.054 977)</td>
</tr>
</tbody>
</table>

\(^a\)Reference 5.
\(^6\)[\(E(k \leq 5) - E(SCF\ limit)\)/\(E(k < N_{\text{prim}}) - E(SCF\ limit)\)].

B. Convergence to the \(N_{\text{prim}}\) basis set limit

The cc-pV\(N\)z, “correlation consistent polarized valence \(n\)tuple-\(\zeta\)” basis sets developed by Dunning and co-workers\(^14\)–\(^19\) provide a convenient reference set of optimized Gaussian basis sets. The equivalent expansions in \(N_{\text{prim}}\) parameter Legendre polynomials are in agreement with the energies from the cc-pVQZ, quadruple-\(\zeta\) correlation consistent basis sets\(^14\)–\(^19\) to within \(\pm 1\) mhartree, with the exception of krypton (Table II).

A detailed comparison reveals that the \(d\) exponents of the cc-pVQZ krypton basis set were not completely optimized. In fact, the \(d\) exponents from a six-parameter Legendre expansion give an energy 2 mhartree below the cc-pVQZ energy. The 12 mhartree error in the krypton cc-pVQZ energy is of no consequence for practical applications of these basis sets. A more serious incomplete optimization was uncovered for the krypton cc-pVDZ basis set. The cc-pVDZ energy (\(-2751.974 872\))\(^19\) is 39 mhartree (25 kcal/mol) above the energy (\(-2752.014 064\)) from the \(k_{\text{max}} = N_{\text{prim}} - 1\) fully optimized Legendre expansion. Once again, it is the \(d\) exponents of the cc-pVDZ krypton basis set that are responsible for the entire discrepancy. These two examples demonstrate the superior reliability of exponent optimizations employing expansions in Legendre polynomials.

If we have six or fewer Gaussian functions, the six-term

TABLE II. Comparison of SCF energies (hartree a.u.) obtained from optimized six-parameter Legendre expansions with those from fully optimized Gaussian basis sets.

<table>
<thead>
<tr>
<th>Basis set</th>
<th>He</th>
<th>Ne</th>
<th>Ar</th>
<th>Kr</th>
<th>Xe</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\Sigma A_k P_k, k \leq 5)</td>
<td>(-2.861 514 01)</td>
<td>(-128.543 75)</td>
<td>(-526.816 674)</td>
<td>(-2752.054 687)</td>
<td>(-7232.138 184)</td>
</tr>
<tr>
<td>(\Sigma A_k P_k, k \leq N_p)</td>
<td>(-2.861 514 23)</td>
<td>(-128.543 470)</td>
<td>(-526.816 781)</td>
<td>(-2752.054 726)</td>
<td>(-7232.138 217)</td>
</tr>
<tr>
<td>cc-pVQZ</td>
<td>(-2.861 514 23)</td>
<td>(-128.543 470)</td>
<td>(-526.816 780)</td>
<td>(-2752.054 714)</td>
<td>NA</td>
</tr>
<tr>
<td>SCF limit</td>
<td>(-2.861 680 01)</td>
<td>(-128.547 098)</td>
<td>(-526.817 513)</td>
<td>(-2752.054 977)</td>
<td>(-7232.138 362)</td>
</tr>
<tr>
<td>Error ratio(^6)</td>
<td>1.001 2</td>
<td>1.026</td>
<td>1.145</td>
<td>1.146</td>
<td>1.23</td>
</tr>
</tbody>
</table>

\(^6\)[\(E(k \leq 5) - E(SCF\ limit)/E(k < N_{\text{prim}}) - E(SCF\ limit)\)].
C. Computational effort

The number of SCF energy evaluations required to achieve convergence for a particular basis set is illustrated in Fig. 10, where we compare a conventional optimization of “scale factors,” $\delta_j$, for each individual exponent, $\alpha_i = \delta_j \times \alpha_{ij}^0$, with an optimization based on Legendre polynomials of increasing order, $k_{\text{max}}$ in Eq. (7). Both use even-tempered basis sets as the initial guess. The advantage of the Legendre expansion grows with the size of the basis set, increasing to about an order of magnitude for the (21s15p10d) krypton basis set, and substantially more than an order of magnitude for the (24s17p12d) krypton basis set (Fig. 10).

Although the full optimization of basis sets containing up to 27s, 19p, and 14d Gaussian functions (Table II) is clearly both possible and reliable using expansions of $\ln \alpha_j$, in $N_{\text{prim}}$ Legendre polynomials, we do not recommend this procedure. Such applications require dozens of iterations, each of which requires $2N_{\text{prim}}$ numerical derivatives, $\partial(\mathcal{H})/\partial \alpha_k$ and $\partial^2(\mathcal{H})/\partial \alpha_k^2$. In contrast, the six-parameter optimization

Legendre expansion will be exact. The helium (7s) basis set thus shows only a 0.12% increase in the error if we employ a six-parameter Legendre expansion (Table II). This error gradually increases to 23% of the inherent basis set error as we increase the number of primitive Gaussian functions in the basis set sequence: He (7s), Ne (12s6p), Ar (16s11p), Kr (21s16p12d), and Xe (27s19p14d). Even with the largest xenon (27s19p14d) basis set, the error from truncating the Legendre expansion with six terms is less than one-fourth of the error from limiting the primitive Gaussian basis to “only” (27s19p14d).

![Figure 10](image)

**FIG. 10.** The four fully optimized Kr basis sets, (15s11p6d), (18s13p8d), (21s15p10d), and (24s17p12d), give progressively smaller SCF errors indicated by the four dashed horizontal lines (---). Optimization of each of these four basis sets using six-term Legendre polynomial expansions (---) approaches the six-term limit indicated by a dotted horizontal line (---). Increasing the order of the polynomial expansion (---) then converges to the limit for each basis set (---). Conventional optimization of the individual Gaussian exponents (---) is much more slowly convergent.

**TABLE III.** Convergence of the SCF energies (hartree a.u.) obtained from optimized six-parameter Legendre expansions$^a$ to the SCF limit.

<table>
<thead>
<tr>
<th></th>
<th>He</th>
<th>Ne</th>
<th>Ar</th>
<th>Kr</th>
<th>Xe</th>
</tr>
</thead>
<tbody>
<tr>
<td>2Z</td>
<td>-2.859 895 4</td>
<td>-128.540 925</td>
<td>-526.796 326</td>
<td>-2752.015 249</td>
<td>-7232.078 845</td>
</tr>
<tr>
<td>3Z</td>
<td>-2.861 514 0</td>
<td>-128.546 374</td>
<td>-526.814 538</td>
<td>-2752.050 546</td>
<td>-7232.131 045</td>
</tr>
<tr>
<td>4Z</td>
<td>-2.861 660 6</td>
<td>-128.546 985</td>
<td>-526.816 983</td>
<td>-2752.054 388</td>
<td>-7232.137 251</td>
</tr>
<tr>
<td>5Z</td>
<td>-2.861 677 3</td>
<td>-128.547 078</td>
<td>-526.817 413</td>
<td>-2752.054 885</td>
<td>-7232.138 184</td>
</tr>
<tr>
<td>6Z</td>
<td>-2.861 679 6</td>
<td>-128.547 094</td>
<td>-526.817 490</td>
<td>-2752.054 962</td>
<td>-7232.138 325</td>
</tr>
<tr>
<td>8Z</td>
<td>-2.861 680 0</td>
<td>-128.547 098</td>
<td>-526.817 513</td>
<td>-2752.054 977</td>
<td>-7232.138 362</td>
</tr>
<tr>
<td>Num$^b$</td>
<td>-2.861 680 0</td>
<td>-128.547 098</td>
<td>-526.817 513</td>
<td>-2752.054 977</td>
<td>-7232.138 362</td>
</tr>
</tbody>
</table>

$^a$Z-He(5s), NE(10s6p), Ar(12s8p), Kr(15s11p6d), and Xe(18s13p8d). 3Z: He(7s), NE(12s8p), Ar(14s10p), Kr(18s13p8d), and Xe(21s15p10d). 4Z: He(9s), NE(14s10p), Ar(16s12p), Kr(21s15p10d), and Xe(24s17p12d). 8Z: He(17s), NE(22s18p), Ar(24s20p), Kr(33s23p18d), and Xe(36s25p20d).

$^b$Numerical Hartree–Fock from Ref. 20.
TABLE IV. Convergence of SCF energies (hartree a.u.) to the \( s,p \) limit of the \( N_2 \) molecule (\( R_{NN} = 2.068 \) bohr a.u.).

<table>
<thead>
<tr>
<th>Limit</th>
<th>Basis set</th>
<th>( \alpha_j = \sum A_j P_j \ ; k \leq 5 )</th>
<th>Full opt</th>
<th>Error ratio(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( s, p ) limit</td>
<td>15s8p</td>
<td>-108.910 759</td>
<td>-108.910 828(^b)</td>
<td>1.091</td>
</tr>
<tr>
<td></td>
<td>16s9p</td>
<td>-108.910 816</td>
<td>-108.910 874</td>
<td></td>
</tr>
<tr>
<td></td>
<td>17s10p</td>
<td>-108.911 479</td>
<td>-108.911 507</td>
<td></td>
</tr>
<tr>
<td></td>
<td>18s11p</td>
<td>-108.911 587</td>
<td>-108.911 588</td>
<td></td>
</tr>
<tr>
<td></td>
<td>19s12p</td>
<td>-108.911 587</td>
<td>-108.911 588</td>
<td></td>
</tr>
<tr>
<td></td>
<td>20s13p</td>
<td>-108.911 587</td>
<td>-108.911 588</td>
<td></td>
</tr>
<tr>
<td></td>
<td>23s15p</td>
<td>-108.911 587</td>
<td>-108.911 588</td>
<td></td>
</tr>
<tr>
<td></td>
<td>26s18p</td>
<td>-108.911 587</td>
<td>-108.911 588</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)\( E(\sum A_j P_j \ ; k \leq 5) - E(\mathrm{SCF} \ s,p \ \text{limit}) \)/\( E(\text{full opt}) - E(\text{SCF} \ s,p \ \text{limit}) \).
\(^b\)Reference 2.

generally requires about ten iterations involving just six such derivatives, which can reduce the computation time by more than an order of magnitude, while increasing the error in the SCF energy by less than 25% (Figs. 4 and 10). The computational effort for each energy evaluation increases as \( (N_{\text{prim}})^4 \), so Fig. 10 understates the increase in computer time required for the larger basis sets.

D. Convergence to the CBS SCF limit

We have employed Legendre polynomials of degree 5 to optimize a sequence of basis sets for the noble gas atoms (Table III). The number of primitive Gaussians for the “double \( \zeta \)" sets: He (5s), Ne (10s6p), Ar (12s8p), Kr (15s11p6d), and Xe (18s13p8d), were selected to give an accuracy of between 0.6 and 1.2 hartree/atom for each atom. The increment (He+2s, Ne+2s2p, Ar+2s2p, Kr+3s2p2d, and Xe+3s2p2d) used to form the sequence: 2Z, 3Z, …, 6Z, was also selected to give a consistent error per electron for each of these atoms, but the range grows to between 0.2 and 1.2 \( \mu \)hartree/atom for the 6Z basis sets. The SCF limits obtained with very large, 8Z, basis sets: He (17s), Ne (22s18p), Ar (24s20p), Kr (33s23p18d), and Xe (36s25p20d), are in complete agreement with numerical Hartree–Fock calculations\(^20\) (Table III). This verifies the reliability of using Legendre polynomials of degree 5 to optimize very large basis sets and lends credibility to the Douglas–Kroll–Hess relativistic SCF limits we plan to determine for the first four complete rows of the periodic table (i.e., H through Xe).

E. The \( N_2 \) molecule

Studies of the \( s, p \) limit for the \( N_2 \) molecule provide a clear example of the advantages the expansion of \( \ln \alpha \) in Legendre polynomials has relative to direct exponent optimization. Jensen reported\(^2\) the energy of \( N_2 \) with a fully optimized (15s8p) Gaussian basis set, but he was forced to constrain the exponents for the (16s9p) basis to avoid “variational collapse.” He concluded that this partially optimized (16s9p) energy (-108.910 851 hartree) was “a somewhat crude upper limit,” but nevertheless “our best estimate” of the \( s, p \) limit.

The expansion of the \( \ln \alpha \) for the (15s8p) basis set in six Legendre polynomials gives an energy (-108.910 759 hartree) that is only 9% farther from the \( s, p \) limit than the energy (-108.910 828 hartree) from the fully optimized (15s8p) basis set (Table IV). Increasing the \( s, p \) basis to (16s9p) and beyond poses no problem for the Legendre expansions leading to a value (-108.911 588 hartree) for the \( s, p \) limit that is significantly lower than Jensen’s value (Table IV). However, we suspect that coupling between \( s \) and \( p \) functions on adjacent centers giving approximate \( \sigma \) and \( \pi \) components of \( d, f, g \), etc., functions makes limits referenced to atomic angular momentum poorly defined for molecules, and thus do not attach much significance to our result. Nevertheless, this example clearly demonstrates the advantages of the expansion of \( \ln \alpha \) in Legendre polynomials.

Atomic functions differing in angular momentum are almost independent, and can therefore be optimized separately. The molecular coupling of different atomic angular momentum functions requires simultaneous optimization of all exponents and makes the optimization of molecular basis sets much more slowly convergent than the optimization of the same basis set for the constituent atoms. Expansions in Legendre polynomials make the optimization substantially less difficult, but we still do not recommend such optimizations for one particular molecule.

TABLE V. Convergence of SCF energies (hartree a.u.) to the CBS SCF limit of the \( N_2 \) molecule (\( R_{NN} = 2.068 \) bohr a.u.).

<table>
<thead>
<tr>
<th>Basis set</th>
<th>( \alpha_j = \sum A_j P_j \ ; k \leq 5 )</th>
<th>Full optimization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calculated</td>
<td>Extrapolated(^a)</td>
<td>Calculated</td>
</tr>
<tr>
<td>9s5p1d</td>
<td>-108.969 780</td>
<td>-108.969 780</td>
</tr>
<tr>
<td>11s7p2d1f</td>
<td>-108.990 688</td>
<td>-108.993 469</td>
</tr>
<tr>
<td>13s9p3d2f1g</td>
<td>-108.993 313</td>
<td>-108.993 781</td>
</tr>
<tr>
<td>17s13p5d4f3g2h1i</td>
<td>-108.993 792</td>
<td>-108.993 819</td>
</tr>
<tr>
<td>21s17p7d6f5g4h2i</td>
<td>-108.993 817</td>
<td>-108.993 817</td>
</tr>
<tr>
<td>Numerical</td>
<td>-108.993 826(^b)</td>
<td>-108.993 826(^b)</td>
</tr>
</tbody>
</table>

\(^a\)Equation (8).
\(^b\)Reference 21.
Practical calculations that are easily extended to polyatomic molecules employ basis sets that have been optimized for the constituent atoms and include progressively higher angular momentum functions (Table V). Such expansions allow extrapolation based on the neon convergence (Table III):

$$E(N_2)_{\text{SCF}} = E(N_2)_{nZ} + \frac{[E(N_2)_{\text{SCF}} - E(N_2)_{nZ}]}{[E(N_2)_{nZ} - E(N_2)_{(n-1)Z}]}$$

(8)

to a value within a few microhartree of the numerical SCF limit (Table V). We shall examine these extrapolations in detail in a subsequent publication.

Full optimization of the \((9s5p1d)\) basis set gives an energy for the nitrogen atom that is 0.50 kcal/mol (0.801 mhartree) lower than the energy from optimization with a six-term Legendre expansion. The full optimization of this atomic basis set lowers the \(N_2\) molecular energy by 1.22 kcal/mol. This increases the calculated \((9s5p1d)\) SCF dissociation energy of the \(N_2\) molecule from 112.71 to 112.92 kcal/mol, which is a trivial correction relative to the SCF limit dissociation energy, 120.45 kcal/mol. The effects of full optimization are even smaller with the larger basis sets. In fact, full optimization for the nitrogen atom raises the energy of the \((17s13p5d4f3g2h1i)\) basis set for the \(N_2\) molecule by 4 \(\mu\)hartree (Table V). Complete optimization of the atomic basis set by using \(N_{\text{prim}}\) terms in the expansion of \(\ln \alpha\) in Legendre polynomials offers no advantage whatever over using just six terms, \textit{when the atomic basis set is applied to a molecule} (Table V).

IV. CONCLUSIONS

The error relative to the exact numerical SCF limit for a six-term expansion of \(\ln \alpha\) in Legendre polynomials is consistently less than 25% larger than the error for the completely optimized basis set, even for \(N_{\text{prim}} > 20\). If two or more basis sets (i.e., two or more values of \(N_{\text{prim}}\)) are used to extrapolate to the SCF limit, full optimization of the Gaussian exponents appears to offer no advantage at all over optimization of a six-term expansion of \(\ln \alpha\) in Legendre polynomials. Thus, there is no need to optimize more than six well-conditioned variational parameters, even for the largest sets of Gaussian primitives.

ACKNOWLEDGMENTS

The authors are grateful to the Department of Energy and Gaussian, Inc., for their support of this research.