Regular article

Relativistic and nonrelativistic finite nucleus optimized double zeta basis sets for the $4p$, $5p$ and $6p$ elements

Kenneth G. Dyall

Thermosciences Institute, NASA Ames Research Center, Mail Stop 230-3, Moffett Field, CA 94035-1000, USA

Received: 30 March 1998 / Accepted: 12 June 1998 / Published online: 19 October 1998

Abstract. Relativistic and nonrelativistic valence double zeta basis sets have been optimized at the self-consistent field (SCF) level with a Gaussian nuclear charge distribution for the $4p$, $5p$ and $6p$ elements. The basis sets were optimized on the weighted average energy of the nonrelativistic p^n configuration. Each basis set has been supplemented with a correlating d function optimized in multireference singles and doubles CI calculations out of the SCF reference configuration set, a diffuse s and p function optimized at the SCF level for the negative ion for all except the rare gases, and a diffuse correlating d function optimized in multireference singles and doubles CI calculations out of the negative ion SCF reference configuration set. These basis sets are therefore the equivalent of the correlation-consistent double zeta basis sets of Dunning and coworkers. In addition to the valence correlation and diffuse functions, each basis set has also been supplemented with two f functions, one for correlation and one for polarization of the sub-valence d shell, and prescriptions given for obtaining s and p functions for sub-valence shell correlation.

Key words: Gaussian basis sets $-$ Relativistic basis sets $$ p Block – Double zeta – Correlation consistent – Finite nucleus

1. Introduction

The burgeoning interest in the heavy elements among quantum chemists brings with it two related demands: quantum chemistry programs which can handle relativistic effects, and Gaussian basis sets for use in these programs. Although effective core potential (ECP) methods have dominated the scene until now, the feasibility of all-electron methods has improved with the rapid increase in computing power. It is therefore timely to tackle the second of the two demands. Already there are some even-tempered basis sets available [1, 2] and some obtained by fitting techniques [3, 4]. The former are rather large for the purposes of molecular calculations on anything but the smallest systems, and fitting procedures do not necessarily produce the best quality basis sets. The trade-off between quality and size is best met by energy-optimized basis sets.

Recently, a method for relativistic energy-optimized basis sets was presented [5], and a computer program developed for the purpose of basis set optimization. The program has been extended to permit the optimization of correlating functions in CI calculations and to optimize nonrelativistic basis sets. This paper reports the first set of relativistic energy-optimized basis sets from this program, for the $4p$, $5p$ and $6p$ elements. The basis sets are intended to match the correlation consistent basis sets of Dunning [6].

The use of relativistic all-electron methods brings with it the scope for comparison with nonrelativistic methods to obtain relativistic corrections to properties. Such comparisons should be done in equivalent basis sets. For molecular properties, particularly in small basis sets, a calculation of the basis set superposition error (BSSE) can be essential to obtain good values of the properties. When comparing relativistic and nonrelativistic properties, then, it is important to ensure that the BSSE is the same (or very similar) for the two methods, otherwise the computed relativistic correction will contain a basis set error. Simply using a nonrelativistic basis set in a relativistic calculation does not ensure this. Therefore nonrelativistic basis sets which match the relativistic basis sets in quality have also been generated.

2. Method

The extension of nonrelativistic approaches to the optimization of basis sets for relativistic calculations is not entirely straightforward. First, one has to decide which quantum number to optimize on $-\ell$, j or κ . For the widest use and the most efficient integral evaluation with current programs, an ℓ -optimized basis set is probably to be

This article is supplemented by an internet archive which can be obtained electronically from the Springer-Verlag server located at http://link.springer.de/journals/tca

preferred, although the indications of Ref. [5] are that a j-optimized basis set would have superior properties, at least for the small size of a double zeta basis. In this work, ℓ -optimization has been chosen.

A second issue is the choice of the state on which to optimize the basis set. In nonrelativistic basis optimizations, it is common to take the ground state of the atom, in LS coupling. In the relativistic regime, it is not always possible or desirable to select a suitable single state. Choice of one spin-orbit component could unduly bias the calculation towards that state, where a balanced treatment of spin-orbit effects could be important for the description of molecular effects. Moreover, the formation of molecular bonds can often be viewed as preceded by a hybridization, which corresponds to an atomic excitation. For these reasons, the energy function for the optimization was taken to be the degeneracy-weighted average of the energies of the states arising from the principal nonrelativistic configuration, s^2p^n . The basis functions for the occupied spinors were determined in self-consistent field (SCF) calculations on this set of states.

A third issue is that of the nuclear model. Here, we have chosen the atomic mass number of the principal isotope of each species to determine the exponent of the Gaussian nuclear charge distribution in accordance with a standard prescription [7]. This means that the basis sets are dependent on the neutron number as well as the proton number, though for these basis sets, the truncation error is probably larger than any error that might be incurred by using a different isotope. The finite nuclear model has been used for both the nonrelativistic and the relativistic basis sets.

The size of the basis sets was determined from calculations on the rare gas atoms. In selecting the basis sets, a balance has to be struck between the number of exponents describing the radial maxima and the energy truncation error incurred in increasing the size of the basis set. To determine the size of the sets in each orbital space, basis sets were optimized within each space with different numbers of functions. The starting functions for each basis set were taken from the sets of Fægri ($[8]$; K. Fægri Jr., personal communication). In each case the largest truncation error for the largest double zeta set occurred in the p space rather than the s space. The p set was therefore taken to be the largest double zeta set. The sets for the other orbital spaces were chosen so that the energy for the addition of one function was approximately equal for each orbital space, except that in the 6p set, the s set was chosen to be the largest double zeta set. To obtain comparable nonrelativistic and relativistic basis sets, in addition to the energy balance and radial maximum description, the coefficients of the radial functions were examined, to obtain the most nearly similar distribution of coefficients over the outer radial maxima. In this way, the sizes of the basis sets arrived at were 15s11p6d for the 4p elements for both relativistic and nonrelativistic basis sets, 19s15p9d and 18s14p9d for the relativistic and nonrelativistic $5p$ element basis sets, respectively, and 24s20p13d8f and 22s17p13d8f for the relativistic and nonrelativistic 6p element basis sets, respectively. The energy truncation error in the p set for the $6p$ elements is particularly large due to the relativistic effects, but any increase in the basis set size produces a triple zeta distribution. The diffuse s and p functions were determined by optimizing these exponents for the negative ion at the SCF level, again using the weighted average energy.

The valence correlating functions were determined in multireference singles and doubles CI (MR-SDCI) calculations out of the reference set used for the SCF calculations, for consistency. Excitations out of the s and p orbitals were considered. Only the d exponent was optimized, in a calculation which included one s, p and d function in the correlating space. The s and p functions were taken to be the outermost primitives of the SCF set, in accordance with the findings of Dunning [6] for the first row. Similarly, the diffuse d function for each basis set was determined in an MR-SDCI calculation on the reference set for the negative ion, with only the diffuse d function being optimized, in a calculation which included the two most diffuse s and p functions (i.e. the outermost SCF functions and the diffuse functions) and the d correlating function as well as the diffuse d function in the correlating space. Diffuse functions were not determined for the rare gas atoms.

Because of the proximity of the d shell to the valence space, particularly for the early members of the p block, a correlating f function was determined for this shell, in a similar series of MR-SDCI calculations. The correlating space included one s , p , d and f function, all of which were optimized. It was found that the s function was close to the geometric mean of the second and third most diffuse functions in the SCF set, i.e. it fell "in the gap" between the *ns* and $(n - 1)s$ functions for the 6*p* elements. Going up the periodic table, this s function approached the third most diffuse function, i.e. the outer SCF primitive in the $(n - 1)s$ maximum. The p function was close to the outer SCF primitive in the $(n - 1)p$ maximum, and the d function was usually between the outermost two d SCF primitives. In addition, for the 6p elements, the f function was close to the outermost f SCF primitive. With this information, a prescription for determining the correlating functions for the $(n - 1)$ shell was formed, which adds a 2s2p2d1f primitive outer core correlating set to the existing sets. For the s and p sets, the outermost SCF primitive from the $(n - 1)$ shell is taken for the first function, and the geometric mean of this and the innermost SCF primitive from the n shell is taken for the second. For the d set, the outermost two d SCF primitives are taken. The MR-SDCI calculations were performed for the $p¹$ and $p⁵$ elements, and a linear interpolation taken for the f functions of the intervening elements, and an extrapolation for the rare gases, for the 4p and $5p$ elements. The outermost f SCF primitive was taken for the 6p elements.

In addition to the f functions determined for correlation, a set of f functions was determined for polarization of the d shell. The exponents were optimized to maximize the magnitude of the second-order perturbation energy for the electric dipole polarizability of the d shell, calculated with a set of configurations involving single $d \rightarrow f$ excitations with respect to the reference set.

3. Results

The total energies and the truncation errors for the basis set SCF calculations with respect to the numerical limit are given in Tables $1-3$ for both the neutral and the negative ion. The truncation error increases slowly with Z across a row, and decreases as Z increases down a column. The negative ion truncation errors are slightly smaller than the neutral values, no doubt due to the increased flexibility in the basis from the extra exponent. The SCF electron affinities deduced from this table do not in fact represent the physical values, because of the state averaging. This explains why, for example, Tl apparently has a negative electron affinity whereas the physical value would be expected to be large due to the relativistic stabilization of the $6p_{1/2}^2$ state, and Pb has an appreciable electron affinity where the physical value, for the same reason, would be expected to be small. The relativistic truncation errors are somewhat larger than the nonrelativistic truncation errors, but are very comparable in magnitude. This is in contrast to the use of nonrelativistic basis sets with relativistic contraction coefficients, where the truncation error can be orders of magnitude larger $-$ see Ref. [5] for an example. The decrease in the truncation error with increasing n is due to the way in which the basis sets have been chosen on the basis of the best double zeta p set. The exponents are generally in increasing sequence across the row, with a few exceptions in the core. These are due to the effect of the nuclear size for the different isotopes $-$ for example, because the mass number of Rn is significantly larger than that of At, there is a drop in the innermost s exponents going from At to Rn due to the shape of the nuclear potential for the two different nuclides. The

Table 1. Total energies in E_h of basis set and numerical SCF calculations and differences between the two for the 4p neutral atoms and negative ions

Table 2. Total energies in E_h of basis set and numerical SCF calculations and differences between the two for the 5p neutral atoms and negative ions

relativistic and nonrelativistic exponents for the 4p elements do not differ significantly in the valence region, and even in the core are not too different. The differences increase with Z , as expected. The $6p$ nonrelativistic basis sets are the same size as those of Fægri [8] and apart from the core region have almost identical exponents. A sample of the exponents is given in Tables $4\overline{-7}$.

To illustrate the differences that would be obtained if the same primitive sets were used in both relativistic and nonrelativistic calculations, some results were obtained for the PbO molecule using the DREAMS [9] package¹, with both the relativistic and nonrelativistic sets used to obtain the contraction coefficients. The Pb basis set was contracted to a valence double zeta set, with the outer-

Table 3. Total energies in E_h of basis set and numerical SCF calculations and differences between the two for the 6p neutral atoms and negative ions

| | | Relativistic | | Nonrelativistic | |
|----------------|-----------------------|-----------------------------|-----------------|------------------------------------|-----------------|
| | | Neutral atom | Negative ion | Neutral atom | Negative ion |
| T1 | | | | | |
| | basis | -20274.830354 | -20274.822523 | -18961.126187 | -18961.123334 |
| | numeric | -20274.850644 | -20274.842721 | -18961.135824 | -18961.132912 |
| | difference | 0.020291 | 0.020198 | 0.009637 | 0.009578 |
| P _b | | | | | |
| | basis | -20913.693854 | -20913.705883 | -19523.251646 | -19523.268606 |
| | numeric | -20913.714332 | -20913.725813 | -19523.261070 | -19523.277785 |
| Bi | difference | 0.020478 | 0.019930 | 0.009424 | 0.009179 |
| | basis | -21565685154 | -21565.721569 | -20094.752029 | -20094.792215 |
| | numeric | -21565.706080 | -21565.741649 | -20094.761394 | -20094.801188 |
| | difference | 0.020925 | 0.020080 | 0.009365 | 0.008973 |
| P _O | | | | | |
| | basis | -22230.991543 | -22231.055407 | -20675.664546 | -20675.730557 |
| | numeric | -22231.013179 | -22231.075893 | -20675.673968 | -20675.739446 |
| | difference | 0.021636 | 0.020486 | 0.009422 | 0.008889 |
| At | | | | | |
| | basis | -22909.785069 | -22909.879008 | -21266.023661 | -21266.117759 |
| | numeric | -22909.807616 | -22909.900086 | -21266.033217 | -21266.126674 |
| | difference | 0.022547 | 0.021078 | 0.009556 | 0.008915 |
| Rn | | | | | |
| | basis | -23602.080766 | | -21865.843344 -21865.853041 | |
| | numeric difference | -23602.104246 0.023480 | | 0.009697 | |
| | | | | | |

Table 4. Exponents of relativistic s functions for the $6p$ elements. Powers of ten are given in parentheses

most s, p and d primitives from the $n = 5$ shell uncontracted. The O basis set used was the aug-cc-pVDZ basis set [6]. The bond length and harmonic frequency were obtained from a quartic fit to five points spaced at $0.05a_0$. The results are reported in Table 8. The bond length is a relatively insensitive property, so only small changes are seen. The changes in the harmonic frequencies are relatively a little larger. The relativistic corrections in comparable basis sets, i.e. DHF in a relativistic basis and NRHF in a nonrelativistic basis, are 0.0132 Å and 82.2 cm^{-1} . The relativistic correction to the frequency is overestimated in both basis sets. One would expect to see

¹DREAMS is a Dirac-based Relativistic Electronic Atomic and Molecular Structure program suite, consisting of an adapted version of MOLECULE, a vectorized integral program developed by J. Almlöf and P. R. Taylor, and a Dirac-Hartree-Fock and MP2 program developed by K.G. Dyall. The DHF program methods are described in Ref. [9] and the MP2 program methods are described in Ref [10].

370

Table 5. Exponents of relativistic p functions for the 6 p elements. Powers of ten are given in parentheses

| TI | P b | Bi | P _O | At | R _n |
|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| $1.372826462(+07)$ | $1.443158465(+07)$ | $1.526785076(+07)$ | $1.609473073(+07)$ | $1.685809768(+07)$ | $1.728241218(+07)$ |
| $2.357212827(+06)$ | $2.507433138(+06)$ | $2.686369414(+06)$ | $2.865167510(+06)$ | $3.035311932(+06)$ | $3.156417284(+06)$ |
| $5.235667825(+05)$ | $5.604442459(+05)$ | $6.045605996(+05)$ | $6.489139356(+05)$ | $6.917456577(+05)$ | $7.252827232(+05)$ |
| $1.362907634(+05)$ | $1.463727873(+05)$ | $1.584833540(+05)$ | $1.707001929(+05)$ | $1.825960304(+05)$ | $1.924125034(+05)$ |
| $4.014808936(+04)$ | $4.315790893(+04)$ | $4.678586981(+04)$ | $5.045048669(+04)$ | $5.403468366(+04)$ | $5.708988285(+04)$ |
| $1.318202105(+04)$ | $1.415244017(+04)$ | $1.532507694(+04)$ | $1.650864236(+04)$ | $1.766782597(+04)$ | $1.867650226(+04)$ |
| $4.787986219(+03)$ | $5.125874354(+03)$ | $5.534552353(+03)$ | $5.945714830(+03)$ | $6.347746116(+03)$ | $6.701904351(+03)$ |
| $1.902844147(+03)$ | $2.030219303(+03)$ | $2.184243089(+03)$ | $2.338416493(+03)$ | $2.488564347(+03)$ | $2.621691632(+03)$ |
| $8.150228768(+02)$ | $8.668209059(+02)$ | $9.294239260(+02)$ | $9.917197702(+02)$ | $1.052074430(+03)$ | $1.105744535(+03)$ |
| $3.696760159(+02)$ | $3.921484222(+02)$ | $4.193200676(+02)$ | $4.462075891(+02)$ | $4.721175139(+02)$ | $4.951794080(+02)$ |
| $1.750437726(+02)$ | $1.853563532(+02)$ | $1.978512072(+02)$ | $2.101471782(+02)$ | $2.219276295(+02)$ | $2.324043636(+02)$ |
| $8.445626704(+01)$ | $8.951928251(+01)$ | $9.569678643(+01)$ | $1.017153882(+02)$ | $1.074211959(+02)$ | $1.124594283(+02)$ |
| $4.204322853(+01)$ | $4.448576449(+01)$ | $4.743563035(+01)$ | $5.028313302(+01)$ | $5.296902201(+01)$ | $5.535572478(+01)$ |
| $2.117207489(+01)$ | $2.242284437(+01)$ | $2.392963090(+01)$ | $2.536794737(+01)$ | $2.671527340(+01)$ | $2.791363738(+01)$ |
| $1.010759872(+01)$ | $1.067517645(+01)$ | $1.135160409(+01)$ | $1.200145350(+01)$ | $1.261551158(+01)$ | $1.317283524(+01)$ |
| $4.903149283(+00)$ | $5.191773072(+00)$ | $5.534612267(+00)$ | $5.867798545(+00)$ | $6.186862214(+00)$ | $6.482614164(+00)$ |
| $1.963288773(+00)$ | $2.090830671(+00)$ | $2.241087187(+00)$ | $2.389950356(+00)$ | $2.534717105(+00)$ | $2.672633531(+00)$ |
| $8.054969909(-01)$ | $8.717179903(-01)$ | $9.531749978(-01)$ | $1.036025175(+00)$ | $1.118393406(+00)$ | $1.198635995(+00)$ |
| $2.004962861(-01)$ | $2.167452903(-01)$ | $2.482931026(-01)$ | $2.823137493(-01)$ | $3.172639217(-01)$ | $3.525855605(-01)$ |
| $5.592788373(-02)$ | $6.582541890(-02)$ | $7.791582245(-02)$ | $9.041135818(-02)$ | $1.031024176(-01)$ | $1.158931466(-01)$ |
| $1.143128111(-02)$ | $1.672807205(-02)$ | $2.131720946(-02)$ | $2.560729551(-02)$ | $2.974781623(-02)$ | |

Table 6. Exponents of relativistic d functions for the $6p$ elements. Powers of ten are given in parentheses

| T1 | P _b | Bi | P_{O} | At | Rn |
|--|--|--|--|--|--|
| $7.802159122(+03)$ $2.039697913(+03)$ $7.304909746(+02)$ $3.051673560(+02)$ $1.396497508(+02)$ $6.729264317(+01)$ $3.352059870(+01)$ $1.656838473(+01)$ $8.081529920(+00)$ $3.880654889(+00)$ | $8.490690047(+03)$ $2.208903548(+03)$ $7.884925742(+02)$ $3.288543231(+02)$ $1.503896186(+02)$ $7.251323542(+01)$ $3.618056107(+01)$ $1.797620156(+01)$ $8.820086427(+00)$ $4.272036395(+00)$ | $9.155673091(+03)$ $2.371438049(+03)$ $8.439102586(+02)$ $3.513859738(+02)$ $1.605666247(+02)$ $7.744846220(+01)$ $3.868871366(+01)$ $1.930121654(+01)$ $9.513742763(+00)$ $4.638285283(+00)$ | $9.815160522(+03)$ $2.531863341(+03)$ $8.983495092(+02)$ $3.734331993(+02)$ $1.704931193(+02)$ $8.225238013(+01)$ $4.112548699(+01)$ $2.058609423(+01)$ $1.018537157(+01)$ $4.992467374(+00)$ | $1.047864557(+04)$ $2.692566365(+03)$ $9.526444534(+02)$ $3.953422901(+02)$ $1.803291610(+02)$ $8.700342743(+01)$ $4.353151720(+01)$ $2.185244499(+01)$ $1.084657371(+01)$ $5.341068552(+00)$ | $1.115204533(+04)$ $2.855023772(+03)$ $1.007306358(+03)$ $4.173241051(+02)$ $1.901717602(+02)$ $9.174901724(+01)$ $4.593121487(+01)$ $2.311325778(+01)$ $1.150427088(+01)$ $5.687863147(+00)$ |
| $1.695634968(+00)$ $6.952405987(-01)$ $2.535764968(-01)$ $1.502677828(-01)$ $4.103564828(-02)$ | $1.889061654(+00)$ $7.920011355(-01)$ $2.968329249(-01)$ $1.754599320(-01)$ $5.577262421(-02)$ | $2.073552597(+00)$ $8.857377074(-01)$ $3.386177681(-01)$ $2.022320671(-01)$ $6.968345404(-02)$ | $2.254052929(+00)$ $9.783342452(-01)$ $3.797490308(-01)$ $2.294832468(-01)$ $9.040700398(-02)$ | $2.433076502(+00)$ $1.070774692(+00)$ $4.206681863(-01)$ $2.569928183(-01)$ $1.093891790(-01)$ | $2.612118578(+00)$ $1.163651826(+00)$ $4.616515015(-01)$ $2.866753176(-01)$ |

Table 7. Exponents of relativistic f functions for the 6p elements. Powers of ten are given in parentheses

larger changes in properties that were more sensitive to the core description, like the electric field gradient at the nucleus, or NMR shielding parameters.

The full tables of basis sets including nonrelativistic SCF, spin-free relativistic SCF and Dirac-Fock SCF coefficients are available in ASCII format from the Internet archive for Theoretical Chemistry Accounts (see below), and also from the basis set archive at Pacific Northwest National Laboratory, http://www.emsl.pnl. gov:2080/forms/basisform.html.

Table 8. Bond lengths in \AA and harmonic frequencies in cm⁻¹ for PbO at the SCF level using the two different basis sets for Pb. Frequencies are evaluated for the most abundant isotopes

| | Basis | DHF | NRHF | $\Delta_{\rm rel}$ |
|------------|--------------|------------|-------------|--------------------|
| r_e | Rel | 1.8988 | 1.9117 | 0.0129 |
| | NR | 1.8979 | 1.9120 | 0.0141 |
| ω_e | Rel | 791.1 | 878.6 | 87.5 |
| | NR | 788.7 | 873.3 | 84.6 |

4. Conclusions

Relativistic and nonrelativistic finite nucleus energyoptimized double-zeta basis sets have been developed for the $4p$, $5p$ and $6p$ elements. These have been supplemented with a correlating d function and diffuse s , p and d functions for a description of the negative ion, generated in the same way as the augmented sets of Dunning and coworkers. An f function for correlation of the $(n - 1)d$ shell was also added, along with a prescription for generating a 2s2p2d1f set for outer core correlation. An f function for polarization of the d shell was also determined. It is hoped that these will find use in relativistic molecular calculations and in comparative studies of relativistic and nonrelativistic calculations.

5. Internet archive

This paper contains an Internet archive in ASCII format. The archive contains the nonrelativistic exponent sets with the coefficients of the Hartree-Fock orbitals and the relativistic exponent sets with the coefficients of the Dirac-Fock and spin-free Dirac-Fock spinors, as well as the diffuse, correlating and polarizing functions. The archive is accessible at no charge at http://link.springer.de/journals/ tca.

Acknowledgements. The author was supported by Pacific Northwest National Laboratory contract BPNL 291140-A-A3 to Eloret. This work was supported through the U.S. Department of Energy by the Mathematical, Information, and Computational Science Division phase II grand challenges of the Office of Computational and Technology Research, and performed under contract DE-AC06-76RLO 1830 with Battelle Memorial Institute. All calculations were performed on the IBM RS6000/595 workstations of the Computational Chemistry Branch at NASA Ames Research Center.

References

- 1. Malli GL, Da Silva ABF, Ishikawa Y (1993) Chem Phys Lett 201:37; (1994) J Chem Phys 101:6829
- 2. Jorge FE, Bobbio TB, da Silva ABF (1996) Chem Phys Lett 263:775
- 3. Lee YS, Baeck KK, Mc Lean AD (1989) J Comp Chem 10:112
- 4. Minami T, Matsuoka O (1995) Theor Chim Acta 90:27
- 5. Dyall KG, Fægri K Jr (1996) Theor Chim Acta 94:39
- 6. Dunning TH Jr (1989) J Chem Phys 90:1007
- 7. Visscher L, Dyall KG (1997) At Data Nucl Data Tables 67:207
- 8. Fægri K Jr (1987) Theor Chim Acta 72:297
- 9. Dyall KG (1994) In: Malli GL (ed) Relativisitic and correlation effects in molecules and solids. Plenum, New York, p 17
- 10. Dyall KG (1994) Chem Phys Lett 224:186