

The Dirac Equation in Quantum Chemistry: Strategies to Overcome the Current Computational Problems

LUCAS VISSCHER

Department of Theoretical Chemistry, Faculty of Sciences, Vrije Universiteit Amsterdam,
De Boelelaan 1083, 1081 HV Amsterdam, The Netherlands

Received 2 August 2001; Accepted 23 September 2001

DOI 10.1002/jcc.10036

Abstract: A perspective on the use of the relativistic Dirac equation in quantum chemistry is given. It is demonstrated that many of the computational problems that plague the current implementations of the different electronic structure methods can be overcome by utilizing the locality of the small component wave function and density. Possible applications of such new and more efficient formulations are discussed.

© 2002 Wiley Periodicals, Inc. J Comput Chem 23: 759–766, 2002

Key words: Dirac equation; quantum chemistry; relativistic effects

Introduction

It is by now well accepted that relativistic effects play an important role in our understanding of molecular properties of heavy elements.¹ The most economical approach to include these effects is the relativistic effective core potential method.^{2–4} This method is based on the chemical insight that relativistic effects on valence orbitals and properties can be well modeled using atomic reference calculations and data. Core properties, however, require an all-electron treatment, and here the dominant methods are the Douglas–Kroll–Heß⁵ and ZORA⁶ methods. All these methods are based on a one- or two-component reformulation of the Dirac equation and make a number of approximations to obtain a computationally feasible algorithm.

The direct use of the Dirac–Coulomb–Breit Hamiltonian in molecular calculations has been subject of rapid development in the last decade of the previous century.^{7–9} At the moment of writing, many standard methods of *ab initio* and density functional computational chemistry have been transferred to the four-component relativistic level, and the chief obstacle for more wide-spread use of these approaches is the rather formidable computational expense required to perform the calculations. An added problem is the lack of systematic and well-tested basis sets and the somewhat primitive user interfaces for most codes. This makes application of the four-component methods still something for specialists.

In this perspective I will sketch ways to overcome these drawbacks and indicate possible application fields that go beyond mere benchmarking of other relativistic approaches.

Theory

From the underlying theory of quantum electrodynamics (QED) an approximately relativistic many-body theory is derived¹⁰ in which the one-body Hamiltonian is given by the Dirac equation^{11,12} and the two-electron interaction is modeled by either the Coulomb, the Gaunt,¹³ or the Breit¹⁴ operator. Higher order QED effects play a role for absolute electron binding energies¹⁵ and in very accurate calculations of lighter elements,¹⁶ but the size of other errors is usually larger, so that these small effects can be neglected in molecular applications. The same is true for the difference between the different two-electron operators. Although the Breit operator is the most accurate relativistic two-electron operator,

$$g^{\text{Breit}} = g^{\text{Coulomb}} + \Delta g^{\text{Gaunt}} + \Delta g^{\text{Breit}} \\ = \frac{1}{r_{12}} - \frac{(\alpha_1 \cdot \alpha_2)}{r_{12}} - \frac{1}{2} (\alpha_1 \cdot \nabla_1)(\alpha_1 \cdot \nabla_2)r_{12} \quad (1)$$

it is also the most awkward one to use in practical calculations because integrals need to be computed by expensive integral derivative techniques. The dominant magnetic part (Δg^{Gaunt}) can be evaluated by standard nonrelativistic integral routines, and is relatively easy to add to the Coulomb operator.

Both two- and four-component approaches rely on the so-called no-pair approximation. Although the prediction of the existence of antiparticles like the positron was a key ingredient of the success of Dirac's relativistic theory, their influence on chemistry is negligible. The formation of electron-positron pairs requires energies

Correspondence to: L. Visscher; e-mail: visscher@chem.vu.nl

that are far beyond the range of chemical reactions. Also, the formation of *virtual* electron-positron pairs as used in the perturbation theory of QED to describe interactions of electrons with electromagnetic fields yields only marginal changes in valence electron energies. These higher order terms of QED can, therefore, be neglected for most chemical purposes. However, in doing so it is important to have an accurate reference potential for which the electron and positron solutions are defined. The key difference between a two-component and a four-component approach is the fact that in the latter the electron and positron solutions of the Dirac equation are not decoupled on beforehand. The advantage of this procedure is that a realistic screened nuclear potential is used to define the no-pair approximation that will minimize higher order QED effects. The disadvantage is that matrix elements involving the positronic solutions are to be formed explicitly and iteratively. One way of looking at the different realization of the no-pair approaches is to focus on the relation between the large and small components of the 4-spinors. Transformation to an effective two-component formalism is only possible once this relation is established. Below I will discuss three ways to reach this stage.

The first possibility is to use relativistic Density Functional Theory.^{17–22} A common pragmatic approach is then to ignore current–current interactions and employ a nonrelativistic exchange–correlation potential V^{XC} to supplement the nuclear attraction V^N and Coulomb repulsion V^C potentials on the diagonal of the Dirac operator. This gives the relation

$$\psi_i^S = (V_N + V_C + V_{XC} - \epsilon_i - 2c^2)^{-1}(c\vec{\sigma} \cdot \vec{p})\psi_i^L \quad (2)$$

between the large and small components of the wave function. In the ZORA approach,^{6,23} one realizes that the orbital energies ϵ_i are small compared to the rest mass energy, so that they can be neglected. An additional simplification is possible by replacing the molecular V^C and V^{XC} potentials by the sum of the atomic potentials^{24,25} or by a model potential.²⁶ With these assumptions the inverse operator of Eq (2) can easily be formed, and an efficient two-component formalism for molecular calculations can be derived. Relativistic DFT without these approximations requires an iterative procedure to obtain V_C , V_{XC} , and ϵ_i and reformulation in a two-component formalism is more involved.^{27–29} In that case, four-component implementations^{19,20} may become competitive in terms of computational expense.

The second approach to decouple the electron and positron solutions is to use a Hartree–Fock procedure. In the restricted Hartree–Fock scheme one obtains the relation

$$\psi_i^S = (V_N + V_C - K^{SS} - \epsilon_i - 2c^2)^{-1}(c\vec{\sigma} \cdot \vec{p} - K^{SL})\psi_i^L \quad (3)$$

between the large and small components of the wave function. Compared to Eq. (2), explicit nonlocal exchange operators K

$$\begin{aligned} K^{SL}\psi_j^L(r_1) &= \sum_i^{n_{\text{occupied}}} \left(\int \frac{\psi_i^{L\dagger}(r_2)\psi_j^L(r_2)}{r_{12}} d\mathbf{r}_2 \right) \psi_i^S(r_1) \\ K^{SS}\psi_j^S(r_1) &= \sum_i^{n_{\text{occupied}}} \left(\int \frac{\psi_i^{S\dagger}(r_2)\psi_j^S(r_2)}{r_{12}} d\mathbf{r}_2 \right) \psi_i^S(r_1) \end{aligned} \quad (4)$$

appear at two places, while the electron correlation effects contained in V_{XC} are neglected. Electron correlation can be reintroduced at the no-pair level, but this will not influence relation (3) because the spinor basis is fixed.

Electron correlation effects on the decoupling of electron and positron solutions are taken into account in the relativistic Multi-configuration Self-Consistent Field (MCSCF) formalism. This method has since long been applied in numerical atomic calculations where a proper choice of boundary conditions at the nucleus³⁰ is sufficient to distinguish between electron and positron solutions. Recently, an implementation³¹ of MCSCF in the algebraic approximation³² was reported that is also suited for molecular calculations. The decoupling between electron and positron solutions is achieved by choosing suitable orbitals (e.g., the Hartree–Fock orbitals of the second decoupling scheme) to start a direct second-order optimization of the total energy. This is based on the mini-max principle³³ where the energy is minimized with respect to electron–electron-type rotations and maximized with respect to electron–positron-type rotations. The procedure gives no explicit relation between the large and the small component of individual spinors, but one can at any point of the MCSCF procedure invoke the no-pair approximation and neglect higher order coupling effects by ignoring further electron–positron rotations. The most common application will be to use a relatively small configuration space to determine suitable orbitals for a subsequent multireference correlation approach in the no-pair approximation.

In all three realizations of the four-component no-pair approximation the most straightforward scheme is to expand the large and small components of the spinors in a basis set of the type $\{(\chi_\mu^L, 0)^T, (0, \chi_\mu^S)^T\}$. The small component basis set is then to be related to the large component via the free particle relation $\chi_\mu^S = (\vec{\sigma} \cdot \vec{p})\chi_\mu^L$ to ensure proper resolution of identity in the evaluation of relations (2) or (3). The resulting matrix formulation of the HF problem is

$$\begin{pmatrix} \mathbf{V}_N^{LL} + \mathbf{J}^{LL} - \mathbf{K}^{LL} - \epsilon_i \mathbf{S}^{LL} & c \Pi^{LS} - \mathbf{K}^{LS} \\ c \Pi^{SL} - \mathbf{K}^{SL} & \mathbf{V}_N^{SS} + \mathbf{J}^{SS} - \mathbf{K}^{SS} - (\epsilon_i + 2c^2) \mathbf{S}^{SS} \end{pmatrix} \times \begin{pmatrix} \mathbf{C}_i^L \\ \mathbf{C}_i^S \end{pmatrix} = 0 \quad (5)$$

and solutions are found via a self-consistent-field procedure. The DFT matrix formulation is almost identical except that the K operators are to be replaced by a diagonal \mathbf{V}_{XC} matrix formed in both the large and the small component basis. The MCSCF procedure requires among others evaluation of an orbital gradient that can be also formulated in terms of auxiliary Fock matrices.^{31,32}

After the orbitals are obtained it is natural to work in the molecular spinor representation and write the many-body Hamiltonian in second quantized form as

$$\hat{H} = \sum_{P,Q} Z_Q^P \hat{E}_P^Q + \frac{1}{4} \sum_{P,Q,R,S} V_{RS}^{PQ} \hat{E}_{PQ}^{RS} \quad (6)$$

where Z_Q^P and V_{RS}^{PQ} represent the one-electron and antisymmetrized two-electron matrix elements. The \hat{E}_P^Q and \hat{E}_{PQ}^{RS} operators are replacement operators as defined by Paldus.³⁴ The advantage of this reformulation is that nonrelativistic theory, and various real-

izations of relativistic theory within the no-pair approximation only differ in the definition of the quantities Z and V .

These differences show up when considering the permutation symmetry and algebra (real or complex) of the Z and V matrices. This depends on the inclusion of spin-orbit interactions in the orbital optimization step. Without spin-orbit coupling pure spin orbitals are obtained, and the matrices Z and V are real. Inclusion of spin-orbit coupling gives true 4-spinors and complex matrices. Here, the inclusion of the Δg^{Gaunt} and Δg^{Breit} operators in V does also play a role because the symmetry of these operators differs from that of the dominant operator g^{Coulomb} . Because the Δg^{Gaunt} operator is time-antisymmetric with respect to inversion of time for one of the interacting particles, whereas the Coulomb operator is time-symmetric, incorporation of magnetic two-electron interactions leads to an additional reduction of permutation symmetry for V . Provided that these reductions of permutation symmetry and change of algebra are accounted for, electron correlation can be included at the no-pair level using adaptations of standard nonrelativistic methods like Configuration Interaction (CI),^{35,36} Many-Body Perturbation Theory (MBPT),³⁷ or Coupled Cluster (CC)^{38,39} theory.

Computational Aspects

In this section I will outline the computational bottlenecks in the current implementation of four-component methods in the program package DIRAC⁴⁰ and indicate how these may be overcome. I will thereby focus on approximations that are suited for relatively large molecules. Use of additional point group symmetry in the case of diatomic molecules is also possible, and will be implemented in the near future.

Hartree–Fock

The main reason that four-component calculations have acquired a reputation of being prohibitively expensive for production calculations lies of course in the fact that an additional small component basis set is required to expand the Hartree–Fock equations. This basis could be of just the same size as the large component basis, which in turn, is comparable in size to an all-electron nonrelativistic basis set for the same atom, but this implies use of j -adapted spinors^{41,42} rather than the conventional Cartesian or spherical Gaussian-type basis sets for which integral evaluation routines are readily available. In practice, this means that the basis set in which the molecular one- and two-electron integrals are evaluated is roughly three to four times larger than the corresponding nonrelativistic basis set. Still, this is in itself not a major problem because transformation and diagonalization of large Fock matrices is only prohibitive in the limit of large molecules where linear scaling techniques are used to reduce the computationally more intensive n^4 (with n the number of basis functions) work associated with the integral evaluation. Our implementation⁴³ of the nonrelativistic Lévy–Leblond formalism⁴⁴ that—apart from skipping the calculation of two-electron integrals over small component basis functions—does all the work required for a normal four-component relativistic calculation gives comparable timings as the nonrelativistic DALTON program. This indicates that the additional work associated with the manipulation of large and complex matrices is relatively minor.

It is thus the formation of the small \mathbf{J}^{SS} , \mathbf{K}^{SL} , and \mathbf{K}^{SS} contributions to the Fock matrix in equation that makes Dirac–Hartree–Fock calculations so much more demanding than one- or two-component calculations. The problem is that although the small component *density* is very localized around the nuclei the small component *basis functions* are not. In regions where the molecular potential is approximately constant the small component wave function becomes directly proportional to the first derivative of the large component function. This means that large and small component *basis functions* need to fall off exponentially with the same exponent. This is realized by the kinetic balance condition, which automatically generates diffuse functions in the small component basis set if they are present in the large component basis. The fact that the diffuse small component functions are essential to get a proper kinetic energy of valence electrons does not mean that they are also important for the proper description of electron repulsion. Because the coefficients of diffuse small component functions in the occupied molecular orbitals are of the order α , they will hardly contribute to the small component electron density. This observation makes physically justified approximation possible. A crude first approximation is to neglect the entire $\mathbf{G}^{SS,SS}$ class of integrals [see definition in eq. (9) below] and replace it by a simple point charge model.⁴⁵ This method also neglects the influence of the more substantial small-component density around the nuclei, but this does not appear to cause large errors in the resulting energies and wave functions. Following this line of thought it should be appropriate to also approximate the screening of the nuclei due to the small-component density in the \mathbf{J}^{LL} part of the matrix by this model. Together with the first step this amounts to simply reducing the nuclear charge by the amount of charge present in the small-component basis functions associated with this nucleus. Evaluation of the remaining contributions to the \mathbf{J}^{SS} , \mathbf{K}^{SL} , and \mathbf{K}^{SS} will then dominate the computational cost of the Hartree–Fock procedure. Below I will consider these other contributions in detail. The relevant parts of the Fock matrix become

$$\mathbf{F}^{SS} = \mathbf{V}_N^{SS} - 2c^2 \mathbf{S}^{SS} + \mathbf{I}^{SS,LL} \cdot \mathbf{D}^{LL} \quad (7)$$

and

$$\mathbf{F}^{SL} = c \Pi^{SL} + \mathbf{I}^{SL,LS} \cdot \mathbf{D}^{SL} \quad (8)$$

with the two-electron integral matrix \mathbf{I} being block diagonal in L and S

$$\begin{aligned} \mathbf{I}_{\mu\nu,\kappa\lambda}^{VW,XY} &= G_{\mu\nu,\kappa\lambda}^{VW,XY} \delta_{VW} \delta_{XY} - G_{\mu\lambda,\kappa\nu}^{VY,XW} \delta_{VY} \delta_{XW} \\ G_{\mu\nu,\kappa\lambda}^{VW,XY} &= \iint \frac{\chi_\mu^{Vi}(r_1) \chi_\nu^{Wj}(r_1) \chi_\kappa^{Xi}(r_2) \chi_\lambda^{Yj}(r_2)}{r_{12}} d\mathbf{r}_1 d\mathbf{r}_2 \end{aligned} \quad (9)$$

Crucial is the screening of the nuclear charge by the large component density given by the last term of equation (7). The contributions $\mathbf{I}^{SS,LL} \cdot \mathbf{D}^{LL} = \mathbf{G}^{SS,LL} \cdot \mathbf{D}^{LL}$, and \mathbf{V}_N^{SS} are of similar magnitude and opposite in sign. Both are (for not too compact basis functions) small compared to $2c^2 \mathbf{S}^{SS}$. Because the Fock matrix is formed in a nonorthogonal and rather dense basis small

errors are, however, magnified in the transformation to the orthogonal basis. Approximation schemes should, therefore, take care to keep the balance between the different terms. As an example, we consider a scheme⁴⁶ in which a one-center approximation was introduced to neglect integrals $G_{\mu\nu,\kappa\lambda}^{SS,LL}$ if μ and ν belong to different centers. Already in a calculation on butadiene with a modest cc-pVDZ contracted basis⁴⁷ this resulted in positronic-type eigenvectors with maximum eigenvalues of $+10^7$ au. These were falsely assigned to the electronic set by the energy selection criterium used in the program, and lead to severe convergence problems. An *ad hoc* remedy was to also neglect the corresponding multicenter $V_{N,\mu\nu}^{SS}$ integrals. This moved the erroneous positron levels to around -10^5 au, and stabilized the self-consistent field process. This simple technique makes the computational scaling of $\mathbf{G}^{SS,LL}$ integral calculation an n^3 process which (for large n) should become insignificant relative to the n^4 process of the $\mathbf{G}^{LL,LL}$ integral calculation. For the largest systems that was calculated—a polyethylene chain with 20 carbon atoms—we indeed observed that the evaluation of the $\mathbf{G}^{LL,LL}$ contribution to the Fock matrix took less CPU time (38 min per SCF iteration) than the $\mathbf{G}^{LL,LL}$ class (44 min per SCF iteration). Within this approximation it is also trivial to include the one-center $\mathbf{G}^{SS,SS}$ -type integrals that scale linear with system size. In the same sample case this took 2 min per iteration. The error in the calculated atomization energies relative to the unapproximated result was only 1.2 kcal/mol (0.03%). This proves that the errors in the positron levels did not influence the electronic energy significantly.

Still, it would be desirable to devise an efficient scheme in which also the positron levels attain the proper energies (slightly below $-2c^2 = -3.7 \cdot 10^4$ au). The key ingredient is the approximate representation of the interaction with the large component wave functions in the extensive small component basis set. Errors are tolerable if they are small compared to $-2c^2$. For the Coulomb interaction it should under these conditions be possible to express the large component density in a limited set of fit functions that can be used for efficient evaluation of multicenter interactions. The exchange operator applies only at very short range because it depends on the overlap with the very local small component part of the wave function. For this term the unmodified one-center approximation is probably sufficient. It is trivial to introduce more sophisticated schemes in which all interactions within a certain domain are calculated exactly and interactions outside this domain are modeled. This can be included in the implementation of linear scaling techniques^{48,49} to approximate also long range “nonrelativistic” interactions that involve the $\mathbf{G}^{LL,LL}$ integrals.

When the problem with the additional relativistic two-electron integrals is fully solved and implemented into the available production (four-component) codes, it becomes interesting to investigate the possible bottlenecks due to the large and complex Fock matrices that show up in the algorithms. One possibility would be to use a matrix formulation of the elimination of the small-component (ESC) procedure,^{50–52} but this gives in the limit the nonrelativistic scaling. More appropriate are perhaps schemes⁵³ in which explicit formation and diagonalization of the Fock matrix is completely avoided, because they also tackle the large component or nonrelativistic cubic scaling problem.

Index Transformations

No-pair electron correlation treatments require transformation of matrices from the atomic primitive basis to the molecular orbital basis. Because two-index transformation occur also in the Hartree–Fock step I concentrate on the four-index transformation of the two-electron interaction integrals g^{Coulomb} .

In this step it is appropriate to represent the spin coordinates of the electrons via quaternion algebra and go from a four-component complex to a two-component quaternion form.⁵⁴ This means that the matrix of MO coefficients is quaternion, and can be represented by four real matrices $c_{\mu p}^{X,\Lambda_1}$ that each multiply a different quaternion unit e_Λ ($\Lambda = 0, 3$). Multiplication of quaternion units that refer to the same electron follows the rules for quaternion multiplication, whereas multiplication of units that belong to different electrons should be regarded as a direct product. This means that one can define one quaternion matrix for each electron

$$B_{\mu\nu,pq}^{XX,\Lambda_{12}} e_{\Lambda_{12}} = \sum_{\Lambda_1=0}^3 \sum_{\Lambda_2=0}^3 c_{\mu p}^{X,\Lambda_1} c_{\nu q}^{X,\Lambda_2} e_{\Lambda_1}^* e_{\Lambda_2}^* \quad (10)$$

to give

$$G_{pqrs}^{\Lambda_{12}\Lambda_{34}} = \sum_X \sum_Y \sum_{\mu,\nu} \sum_{\kappa,\lambda} B_{\mu\nu,pq}^{XX,\Lambda_{12}} G_{\mu\nu\kappa\lambda}^{XXYY} B_{\kappa\lambda,rs}^{YY,\Lambda_{34}} \times (\Lambda_{12}, \Lambda_{34} = 0, 1, 2, 3). \quad (11)$$

as the formula to be evaluated in the four-index transformation of the two-electron interaction operator. N_L (N_S) is the number of large (small) components basisfunctions. The quaternion algebra now takes care of spin integration, as it can be shown that the 16 real numbers contained in $G_{pqrs}^{\Lambda_{12}\Lambda_{34}}$ correspond to integrals in the complex 4-spinor representation through the relations (overbars indicate Kramers partners)

$$\begin{aligned} G_{pqrs} &= (G_{pqrs}^{00} - G_{pqrs}^{11}) + i(G_{pqrs}^{01} + G_{pqrs}^{10}) \\ G_{p\bar{q}\bar{r}s} &= (G_{pqrs}^{00} + G_{pqrs}^{11}) + i(G_{pqrs}^{10} - G_{pqrs}^{01}) \\ G_{p\bar{q}r\bar{s}} &= (G_{pqrs}^{22} - G_{pqrs}^{33}) + i(G_{pqrs}^{23} + G_{pqrs}^{32}) \\ G_{p\bar{q}rs} &= (-G_{pqrs}^{22} - G_{pqrs}^{33}) + i(G_{pqrs}^{23} - G_{pqrs}^{32}) \\ G_{pq\bar{r}\bar{s}} &= (G_{pqrs}^{02} + G_{pqrs}^{13}) + i(G_{pqrs}^{03} + G_{pqrs}^{12}) \\ G_{p\bar{q}\bar{r}s} &= (-G_{pqrs}^{02} - G_{pqrs}^{13}) + i(G_{pqrs}^{03} - G_{pqrs}^{12}) \\ G_{p\bar{q}rs} &= (G_{pqrs}^{20} - G_{pqrs}^{31}) + i(G_{pqrs}^{30} + G_{pqrs}^{21}) \\ G_{p\bar{q}rs} &= (-G_{pqrs}^{20} - G_{pqrs}^{31}) + i(G_{pqrs}^{30} - G_{pqrs}^{21}) \end{aligned} \quad (12)$$

In the nonrelativistic Lévy–Leblond formalism the transformation becomes

$$G_{pqrs}^{\text{Lévy-Leblond}} = \sum_{\mu,\nu} \sum_{\kappa,\lambda}^{N_L} B_{\mu\nu,pq}^{LL,0} G_{\mu\nu\kappa\lambda}^{LLLL} B_{\kappa\lambda,rs}^{LL,0} \quad (13)$$

while the spinfree modified Dirac equation gives

$$G_{pqrs}^{\text{spinfree}} = \sum_X \sum_Y^{L,S} \sum_{\mu,\nu}^{N_X} \sum_{\kappa,\lambda}^{N_Y} B_{\mu\nu,pq}^{XX,0} G_{\mu\nu\kappa\lambda}^{XXYY} B_{\kappa\lambda,rs}^{YY,0} \quad (14)$$

All equations were written in terms of **B**-matrices to make the difference between the formalisms transparent. In an actual implementation⁵⁵ it is better to work directly with the original coefficient **C**-matrices to obtain n^5 -order scaling. Starting from a non-relativistic reference point one could say that introduction of scalar relativistic effects requires transformation of the $\mathbf{G}^{SS,LL}$, $\mathbf{G}^{LL,SS}$, and $\mathbf{G}^{SS,SS}$ integrals using a real coefficient matrix, while subsequent introduction of spin-orbit coupling effects makes the algebra quaternionic and adds an additional multiplicative scaling factor of 16. The latter scaling factor is the hardest to overcome, as small couplings can already strongly mix near-degenerate orbitals and make the $\mathbf{B}^{LL,\Lambda}$ matrices nonnegligible. This situation applies also to two-component methods if a spin-orbit coupling term is included at the Hartree–Fock stage.⁵⁶ These methods give the transformation

$$G_{pqrs}^{\text{Two-spinor } \Lambda_{12}\Lambda_{34}} = \sum_{\mu,\nu} \sum_{\kappa,\lambda}^{N_L} B_{\mu\nu,pq}^{LL,\Lambda_{12}} G_{\mu\nu\kappa\lambda}^{LLLL} B_{\kappa\lambda,rs}^{LL,\Lambda_{34}} \times (\Lambda_{12}, \Lambda_{34} = 0, 1, 2, 3) \quad (15)$$

Computational efficiency in the full transformation (11) can thus be increased by reducing the work associated with the evaluation of either the spin-orbit or the scalar relativistic relativistic contributions. If the orbitals are weakly affected by spin-orbit coupling it may be possible to neglect the quaternion imaginary parts of the **B**-matrices and use transformation (14). This approximation is especially valid if the active space consists of only valence molecular orbitals because the larger splittings of the inner occupied orbitals can be taken into account in the generation of a frozen core operator.

For scalar relativistic effects similar arguments as in the discussion of the Hartree–Fock step apply. The small-component contribution to the spinor overlap densities is either very small or localized around the nuclei. This again allows introduction of one-center approximations in this step. Upon neglecting all $\mathbf{G}^{SS,SS}$ integrals the most time consuming steps of the four-index transformation become the evaluation of the $\mathbf{G}^{LL,SS}$ contribution to the half-transformed integrals

$$G_{\mu\nu,rs}^{LL,\Lambda_{34}} \leftarrow \sum_{\kappa,\lambda}^{N_S} G_{\mu\nu\kappa\lambda}^{LLSS} B_{\kappa\lambda,rs}^{SS,\Lambda_{34}} \quad (16)$$

and the transformation of the $\mathbf{G}^{SS,LL}$ half-transformed integrals to the final integrals

$$G_{pqrs}^{\Lambda_{12}\Lambda_{34}} \leftarrow \sum_{\mu,\nu}^{N_S} B_{\mu\nu,pq}^{SS,\Lambda_{12}} G_{\mu\nu,rs}^{SS,\Lambda_{34}} \quad (17)$$

Keeping only one-center contributions in the \mathbf{B}^{SS} -matrices reduces the work in these steps by a factor N_S .

Electron Correlation Methods

A bottleneck in the current implementation of relativistic CC methods is the reorganization of transformed two-electron integrals. The matrix elements V_{RS}^{PQ} of the second quantized Hamiltonian are related to the transformed integrals of eq. (12) by $V_{PQ}^{RS} = G_{PRQS} - G_{PRSQ}$ where the capital indices indicate that P runs over both partners $\{\bar{p}\}$ and $\{p\}$ of the Kramers pairs. The difference in index sequence gives rise to conflicting ordering requirements for the index transformation and MBPT or CC steps. This is especially apparent when implementing parallel algorithms for distributed systems. Choosing the optimal algorithm for the index transformation gives integral batches where the last two indices qs in G_{prqs} and the quaternion unit Λ_{34} are fixed, while the optimal storage mode for the CC calculations requires batches that are sorted on the index pair RS . This problem is also known in the nonrelativistic coupled cluster theory, but can be solved by incorporating the antisymmetrization and sorting step in the index transformation. Adoption of this solution the relativistic case is, however, difficult because antisymmetrization is nontrivial in the quaternion representation of the integrals. The present parallel implementation in RELCCSD³⁹ gathers all integrals G_{PRQS} on a master node to create and subsequently distribute the V_{PQ}^{RS} integrals over the (local) disks on the other nodes. This makes the implementation restricted to cases where fast access by one processor to the complete set of transformed integrals is possible. In practice, this does not present a real computational bottleneck because the storage requirements scale as n^4 , while the operation count increases as n^6 or n^7 with the number of active orbitals. Bypassing storage requirements via use of integral direct-coupled cluster techniques⁵⁷ is less effective in the relativistic case because the difference in size between the active orbital set and the AO basis set is so large. This is due to both the presence of the small component basis set (but approximation techniques could be used as described in the previous section) and the typical application to complexes with heavy elements in which only valence electrons are correlated and a significant number of core orbitals is kept frozen.

A special situation occurs when a spinfree relativistic formalism is adopted because this allows use of nonrelativistic electron correlation implementations. An interface between the nonrelativistic LUCIA program⁵⁸ and the relativistic DIRAC code⁴⁰ was recently completed⁵⁹ to enable large multireference CI calculations in situations where single reference methods like CC and MBPT are not suited. A four-component generalization of the mean-field spin-orbit approach⁶⁰ can hereby be used to include spin-orbit effects in an efficient way. This method will include the two-electron spin-orbit interactions with occupied orbitals in an

Table 1. Comparison of CPU Times (in Seconds on a 450 Mhz Pentium III) for CCSD(T)-Calculations in Different Point Groups.

Point group	Total number of 4-spinor integrals	CPU time per CCSD it	CPU time (T)-correction
C_1^*	31,560,165	426.9	1678.1
C_1	11,908,813 (3)	16.4 (26)	102.3 (16)
C_{2v}^*	15,781,479	22.9	168.3
C_{2v}	3,346,165 (5)	2.2 (10)	14.0 (12)
D_{2h}^*	7,892,111	7.1	55.0
D_{2h}	1,666,267 (5)	0.9 (8)	6.4 (9)

The relativistic calculations indicated by an asterix, the numbers in parentheses give the ratio between the full and the spinfree calculations. The test calculation concerns the F_2 molecule in the molecular 4-spinor basis (116 functions). The 1s spinors were frozen.

effective one-electron operator Z_Q^P . Because the formation of Z has the same computational scaling as the formation of one Fock matrix, this additional effort is negligible compared to the formation of V . The resulting mean field operator can then be used in the two-component version of the LUCIA program.⁶¹ Another route was taken to make spin-free coupled-cluster calculations more efficient. Here, the relativistic RELCCSD program was slightly modified to use the nonrelativistic symmetry that is appropriate for spin-free calculations. This could be done by replacing the double group multiplication tables for the irreps spanned by the spinors by multiplication tables formed as direct products of the spatial and spin multiplication tables. Because all symmetry information is provided to the program by specification of these table and the applicable algebra (real or complex), this automatically lead to a reduction of computational effort.

In the sample case of the fluorine molecule⁶² we find that this transition to the spinfree formalism gives a speedup factors of 8 to 26 in the coupled-cluster calculations (Table 1). An optimized nonrelativistic code that goes beyond the spinorbital formalism and uses a spinrestricted form could, of course, also be interfaced, and may give even higher factors. It is, however, questionable whether pursuing this purely spinfree route is worth the effort, because more efficient ways of treating scalar relativistic effects are available.

Application Fields

I will try to outline the range of possible applications of the methods sketched above. As such an overview is biased by one's own interests, it is inevitable that some areas will be overlooked, and that the importance of others may be overstated.

Diatomic Molecules

Applications of four-component methods often concerned benchmark applications for the standard spectroscopic parameters (r_e , D_e , ω_e , and $\omega_e x_e$) of diatomic molecules. This kind of applications will be of interest for some time to come as the availability of more powerful computers and parallel implementations allows the use of

larger basis sets. The necessary use of moderate or sometimes even modest size basis sets in the past made it essential to compare the outcome of four-component relativistic and other calculations in (approximately) the same basis set. Only in this manner was it possible to extract reliable information on the importance of (higher order) relativistic effects that were missing in the one- or two-component methods. Calculations in large basis sets that give essentially basis set free results can fill a gap in this area and make such comparisons more rigorous.

A different application that has emerged recently is the accurate calculation of core properties that require a relativistic description to get a qualitatively correct answers. One example is the computation of electric field gradients at the nuclei as a means to obtain accurate nuclear quadrupole moments. Here, one needs a high-level treatment of electron (core) correlation, picture change and spin-orbit effects to obtain stable results. Because the four-component approaches do not suffer from picture change errors and have spin-orbit coupling effects included automatically, one needs only concentrate on the electron correlation problem. This makes the application costly but straightforward. Another example is calculation of parity-violating energy shifts that take a much simpler form in the relativistic framework^{63,64} than in the nonrelativistic framework.⁶⁵

Implementation of analytic derivative techniques at the coupled cluster level of theory would be beneficial to increase the numerical stability and computational efficiency for both properties. Such techniques are essential when looking at magnetic properties because introduction of a finite magnetic field will break the Kramers' degeneracy that is used in the restricted Hartree-Fock, DFT, or MCSCF procedures. Calculation of absolute shielding constants for NMR nuclei is probably the most important application once an implementation of second derivatives of correlated methods becomes available. Here, one also sees some interesting aspects⁶⁶ and possible breakdowns of the no-pair approach due to the magnetic field dependence [for simplicity neglected in eqs. (2) and (3)] of the relation between the large and small components. This area still requires significant implementation and method development work.

Small Molecules

Structure and energetics of small molecules (say up to 7 atoms) form another domain in which four-component methods can be applied directly when the techniques discussed in the preceding sections become available in a user-friendly manner. This class would include both molecules that contain heavy elements ($Z > 72$), for which a relativistic description is needed to get qualitatively correct results, and light molecules for which a very high accuracy is desired.^{16,67,68} The advantage of a four-component approach is the black-box nature of inclusion of relativistic effects. In contrast to, for example, a spin-orbit CI-type method, it is not necessary to make an *a priori* assessment of the spin-orbit coupling between potential energy surfaces, because all significant interactions will be included automatically. Given the rapid increase of available computational resources, this may be an area that becomes more prominent in the next decade. Limiting factors are at the moment the lack of systematically optimized relativistic

basis sets suitable for high-level correlation treatments and the still experimental nature of most implementations of the methods.

Typical application examples are on the one end high-accuracy studies of atmospheric or interstellar species, and on the other end studies of actinide complexes in which the usual chemical accuracy is sufficient.

Large molecules

Structure and energetics of larger molecules form an area in which it is not so much the relativistic treatment that gives rise to computational problems, but rather the n^5 or higher order scaling of conventional *ab initio* correlation methods. In this area only the four-component relativistic density functional method may become a competitive method in the foreseeable future. This requires that the rather simple problem of representing the large component potential in the small component basis set is solved (see the section on Hartree–Fock theory above). Improvement of the density functionals by using true relativistic density functionals is probably less urgent, because at the accuracy level that can be anticipated other sources of error will prevail. For the same reason it is hard to predict whether the additional work compared to the more common ECP, Douglas–Kroll–Heß or (two-component) ZORA methods will pay off in significant improvement of reliability and accuracy. Like the case of *ab initio* methods for small molecules, it may be that the black-box nature of inclusion of relativistic effects is worth the effort. For these type of systems it is pertinent to have reliable and efficient structure optimization tools available, some of which will become available in a new release of the DIRAC program.⁴⁰

Conclusions

Although the direct use of the Dirac equation is still far from mainstream quantum chemistry, important progress has been made in the last decade. The key problem to be solved to remove the remaining computational bottlenecks is the approximate yet accurate representation of the Coulomb and exchange potential in the small component basis set. Ways to overcome this difficulty at the Hartree–Fock (or DFT) level of theory and in index transformation procedures have been outlined. The central physical idea behind these approaches is the observation that it is only necessary to provide an accurate description in the vicinity of the (heavy) nuclei of the molecule under consideration. This makes application of one-center approximations feasible.

If the progress seen in the last decade continues at the same rate during the next decade I expect to see more use of Dirac based methods by nonspecialists at the end of this first decade of the 21st century.

Acknowledgments

Many of the ideas presented in this article resulted from fruitful and enjoyable discussion with other members of the small but active community of four-component method developers. The

author would especially like to thank Trond Saue, Ken Dyall, and Hans-Jørgen Jensen for influencing his thoughts in this area.

References

1. Pyykkö, P. *Chem Rev* 1988, 88, 563.
2. Hay, P. J.; Wadt, W. R. *J Chem Phys* 1985, 82, 299.
3. Dolg, M. *Mol Phys* 1996, 88, 1645.
4. Seijo, L.; Barandiaran, Z.; Harguindey, E. *J Chem Phys* 2001, 114, 118.
5. Hess, B. A. *Phys Rev A* 1986, 33, 3742.
6. van Lenthe, E.; Baerends, E. J.; Snijders, J. G. *J Chem Phys* 1993, 99, 4597.
7. Visscher, L.; Visser, O.; Aerts, P. J. C.; Merenga, H.; Nieuwpoort, W. C. *Comp Phys Commun* 1994, 81, 120.
8. Saue, T.; Faegri, K.; Helgaker, T.; Gropen, O. *Mol Phys* 1997, 91, 937.
9. Quiney, H. M.; Skaane, H.; Grant, I. P. *Chem Phys Lett* 1998, 290, 473.
10. Grant, I. P.; Quiney, H. M. *Adv Mol Phys* 1988, 23, 37.
11. Dirac, P. A. M. *Proc R Soc Lond A* 1928, 118, 351.
12. Dirac, P. A. M. *Proc R Soc Lond A* 1928, 117, 610.
13. Gaunt, J. A. *Proc R Soc* 1929, A124, 163.
14. Breit, G. *Phys Rev* 1929, 34, 553.
15. Labzowsky, L.; Goidenko, I.; Tokman, M.; Pyykkö, P. *Phys Rev A* 1999, 59, 2707.
16. Pyykkö, P.; Dyall, K. G.; Csazar, A. G.; Tarczay, G.; Polyansky, O. L.; Tennyson, J. *Phys Rev A* 2001, 6302, 4502.
17. Liberman, D.; Waber, J. T.; Cromer, D. T. *Phys Rev A* 1965, 137, 27.
18. Engel, E.; Dreizler, R. M. In *Density Functional Theory II; Topics in Current Chemistry*; Springer Verlag: Berlin, 1996; Vol. 118, p. 1.
19. Liu, W. J.; Hong, G. Y.; Dai, D. D.; Li, L. M.; Dolg, M. *Theor Chem Acc* 1997, 96, 75.
20. Varga, S.; Fricke, B.; Nakamatsu, H.; Mukoyama, T.; Anton, J.; Geschke, D.; Heitmann, A.; Engel, E.; Bastug, T. *J Chem Phys* 2000, 112, 3499.
21. Liu, W. J.; van Wüllen, C. *J Chem Phys* 2000, 113, 2506.
22. Varga, S.; Fricke, B.; Nakamatsu, H.; Mukoyama, T.; Anton, J.; Geschke, D.; Heitmann, H.; Engel, E.; Bastug, T. *J Chem Phys* 2000, 113, 2508.
23. Chang, C.; Pélissier, M.; Durand, P. *Phys Scr* 1986, 34, 394.
24. Philipsen, P. H. T.; van Lenthe, E.; Snijders, J. G.; Baerends, E. J. *Phys Rev B* 1997, 56, 13556.
25. van Lenthe, E.; Ehlers, A.; Baerends, E. J. *J Chem Phys* 1999, 110, 8943.
26. van Wüllen, C. *J Chem Phys* 1998, 109, 392.
27. van Lenthe, E.; Baerends, E. J.; Snijders, J. G. *J Chem Phys* 1996, 105, 2373.
28. Barysz, M.; Sadlej, A. J.; Snijders, J. G. *Int J Quantum Chem* 1997, 65, 225.
29. Barysz, M. *J Chem Phys* 2000, 113, 4003.
30. Goldman, S. P. *Phys Rev A* 1985, 31, 3541.
31. Thyssen, J. Ph.D. thesis, University of Southern Denmark, 2001.
32. Aa. Jensen, H. J.; Dyall, K. G.; Saue, T.; Faegri, K. *J Chem Phys* 1996, 104, 4083.
33. Talman, J. D. *Phys Rev Lett* 1986, 57, 1091.
34. Paldus, J. In *Relativistic and Electron Correlation Effects in Molecules and Solids*; Malli, G. L., Ed.; Plenum: New York, 1994, p. 207, vol. 318.
35. Esser, M. *Int J Quantum Chem* 1984, 26, 313.
36. Visscher, L.; Saue, T.; Nieuwpoort, W. C.; Faegri, K., Jr.; Gropen, O. *J Chem Phys* 1993, 99, 6704.
37. Dyall, K. G. *Chem Phys Lett* 1994, 224, 186.

38. Eliav, E.; Kaldor, U.; Ishikawa, Y. *Chem Phys Lett* 1994, 222, 82.
39. Visscher, L.; Lee, T. J.; Dyall, K. G. *J Chem Phys* 1996, 105, 8769.
40. Saue, T.; Bakken, V.; Enevoldsen, T.; Helgaker, T.; Aa. Jensen, H. J.; Laerdahl, J. K.; Ruud, K.; Thyssen, J.; Visscher, L. Dirac, a relativistic *ab initio* electronic structure program, Release 3.2 (2000); <http://dirac.chem.sdu.dk>.
41. Quiney, H. M.; Skaane, H.; Grant, I. P. *J Phys B* 1997, 30, L829.
42. Yanai, T.; Nakajima, T.; Ishikawa, Y.; Hirao, K. *J Chem Phys* 2001, 114, 6526.
43. Visscher, L.; Saue, T. *J Chem Phys* 2000, 113, 3996.
44. Lévy-Leblond, M. *Commun Math Phys* 1967, 6, 286.
45. Visscher, L. *Theor Chem Acc* 1997, 98, 68.
46. de Jong, G. Th.; Visscher, L. 2001, to appear. Results of a similar approximation scheme have also been presented by J. K. Pedersen and H. J. Aa. Jensen at the REHE 2001 meeting in Kerkrade, 2001.
47. Dunning, T. H., Jr. *J Chem Phys* 1989, 90, 1007.
48. White, C. A.; Johnson, B.; Gill, P.; Head-Gordon, M. *Chem Phys Lett* 1994, 230, 8.
49. Fonseca Guerra, C.; Snijders, J. G.; te Velde, G.; Baerends, E. J. *Theor Chem Acc* 1998, 99, 391.
50. Dyall, K. G. *J Chem Phys* 1997, 106, 9618.
51. Dyall, K. G. *J Chem Phys* 1998, 109, 4201.
52. Dyall, K. G.; Enevoldsen, T. *J Chem Phys* 1999, 111, 10000.
53. Helgaker, T.; Larsen, H.; Olsen, J.; Jørgensen, P. *Chem Phys Lett* 2000, 327, 397.
54. Saue, T.; Aa. Jensen, H. J. *J Chem Phys* 1999, 111, 6211.
55. Visscher, L.; Laerdahl, J. K.; Saue, T., unpublished work.
56. Ilias, M.; Kellö, V.; Visscher, L.; Schimmelpfennig, B. *J Chem Phys*, 2001.
57. Koch, H.; Aa. Jensen, H. J.; Jørgensen, P.; Helgaker, T.; Scuseria, G. E.; Schaefer, H. F., III. *J Chem Phys* 1990, 92, 4924.
58. Olsen, J.; Jørgensen, P.; Simons, J. *Chem Phys Lett* 1990, 169, 463.
59. Fleig, T.; Visscher, L., to appear.
60. Hess, B. A.; Marian, C. M.; Wahlgren, U.; Gropen, O. *Chem Phys Lett* 1996, 251, 365.
61. Fleig, T.; Olsen, J.; Marian, C. M. *J Chem Phys* 2001, 114, 4775.
62. Visscher, L. *Chem Phys Lett* 1996, 253, 20.
63. Thyssen, J.; Laerdahl, J. K.; Schwerdtfeger, P. *Phys Rev Lett* 2000, 85, 3105.
64. Laerdahl, J. K.; Schwerdtfeger, P.; Quiney, H. M. *Phys Rev Lett* 2000, 84, 3811.
65. Berger, R.; Quack, M. *J Chem Phys* 2000, 112, 3148.
66. Aucar, G. A.; Saue, T.; Visscher, L.; Aa. Jensen, H. J. *J Chem Phys* 1999, 110, 6208.
67. Partridge, H.; Bauschlicher, C. W.; Visscher, L. *Chem Phys Lett* 1995, 246, 33.
68. Kedziora, G. S.; Pople, J. A.; Ratner, M. A.; Redfern, P. C.; Curtiss, L. A. *J Chem Phys* 2001, 115, 718.