Quasirelativistic theory. II. Theory at matrix level

Wenjian Liu and Werner Kutzelnigg

Citation: J. Chem. Phys. **126**, 114107 (2007); doi: 10.1063/1.2710258
View online: http://dx.doi.org/10.1063/1.2710258
View Table of Contents: http://jcp.aip.org/resource/1/JCPSA6/v126/i11
Published by the AIP Publishing LLC.

Quasirelativistic theory. II. Theory at matrix level

Wenjian Liu
Beijing National Laboratory for Molecular Sciences, Institute of Theoretical and Computational Chemistry, Peking University, Beijing 100871, People's Republic of China; State Key Laboratory of Rare Earth Materials Chemistry and Applications, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, People's Republic of China; and Center for Computational Science and Engineering, Peking University, Beijing 100871, People's Republic of China

Werner Kutzelnigg
Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, D-44780 Bochum, Germany

(Received 28 November 2006; accepted 25 January 2007; published online 19 March 2007)

The Dirac operator in a matrix representation in a kinetically balanced basis is transformed to the matrix representation of a quasirelativistic Hamiltonian that has the same electronic eigenstates as the original Dirac matrix (but no positronic eigenstates). This transformation involves a matrix **X**, for which an exact identity is derived and which can be constructed either in a noniterative way or by various iteration schemes, not requiring an expansion parameter. Both linearly convergent and quadratically convergent iteration schemes are discussed and compared numerically. The authors present three rather different schemes, for each of which even in unfavorable cases convergence is reached within three or four iterations, for all electronic eigenstates of the Dirac operator. The authors present the theory both in terms of a non-Hermitian and a Hermitian quasirelativistic Hamiltonian. Quasirelativistic approaches at the matrix level known from the literature are critically analyzed in the frame of the general theory. © *2007 American Institute of Physics*. [DOI: 10.1063/1.2710258]

I. INTRODUCTION

A theory in terms of four-component spinors (bispinors) is called *fully relativistic*. Actually, such a theory is only necessary if one wants to describe *both electrons and positrons*. A theory in terms of only two-component spinors (or simply *spinors*) is referred to as *quasirelativistic*. Such a theory is *always possible* if one is only interested in electrons. This is easily illustrated for one-particle systems.

Let the bispinor $\psi=(\varphi,\chi)$, built up from the upper component spinor φ and the lower component spinor χ, satisfy the Dirac equation

$$D\psi=(E+mc^2)\psi, \quad D=\beta mc^2+c\boldsymbol{\alpha}\mathbf{p}+V \tag{1}$$

(with D the Dirac operator). For any *electronic* state φ and χ are related as

$$\chi=X\varphi, \tag{2}$$

with X solution of the implicit equation[1,2]

$$2mc^2X=c\boldsymbol{\sigma}\mathbf{p}-[X,V]-cX\boldsymbol{\sigma}\mathbf{p}X. \tag{3}$$

An electronic state is fully characterized by the spinor φ. Actually, φ is an eigenfunction of the non-Hermitian operator L,[3]

$$L=V+c\boldsymbol{\sigma}\cdot\mathbf{p}X, \tag{4}$$

$$L\varphi=E\varphi, \tag{5}$$

or alternatively of the Hermitian operator $\bar{\mathbf{L}}$ with a nonunit metric,[4]

$$\bar{L}=\tfrac{1}{2}\{(1+X^\dagger X)L+L^\dagger(1+X^\dagger X)\}, \tag{6}$$

$$\bar{L}\varphi=E(1+X^\dagger X)\varphi. \tag{7}$$

Finally, one can describe a one-electron state by the Foldy-Wouthuysen (FW) spinor ϕ related to φ as

$$\phi=(1+X^\dagger X)^{1/2}\varphi. \tag{8}$$

This satisfies[4,5] the Hermitian eigenvalue equation with a unit metric

$$L^+\phi=E\phi, \tag{9}$$

$$L^+=\tfrac{1}{2}\{(1+X^\dagger X)^{1/2}L(1+X^\dagger X)^{-1/2}+(1+X^\dagger X)^{-1/2}L^\dagger(1+X^\dagger X)^{1/2}\}. \tag{10}$$

In Paper I of this series[6] the quasirelativistic theory at *operator level* has been reviewed critically. One of the messages has been that an exact quasirelativistic theory, with an X that satisfies Eq. (2), is only possible if one describes the *weak singularities* of φ at the position of the point nuclei correctly. This condition is even necessary in order to get an *upper bound* to the exact electronic ground state.

Another observation of Paper I has been that some methods in current use are actually *not* based on a transformation from fully relativistic to quasirelativistic theory at operator level, although they are often presented in this way, but rather on constructing first *a matrix representation* of the Dirac operator and a subsequent transformation at the matrix level. It is the aim of the present paper to consider this transformation at the matrix level from a general point of view. It

will turn out that this procedure is even the preferred one, by which most of the problems that arise for a transformation at the operator level are avoided. While it is usually *not possible* to find a transformation of the Dirac operator to an exact quasirelativistic operator, the corresponding *matrix transformation* can be performed *to any desired accuracy*.

A short Communication on this topic has been published earlier.[7] A comment by Filatov[8] on relations between our approach[7] and an earlier paper Dyall[9] on the *normalized elimination of the small component* (NESC) has been published back to back with the response of the present authors.[10] A message has been that the NESC approach can be regarded as a *precursor* to our method, but certainly not as equivalent to it.

II. MATRIX REPRESENTATION *A PRIORI* VERSUS *A POSTERIORI*

We are exclusively concerned with approximation methods based on the *expansion in a basis*; i.e., we willingly discard methods where the wave function is represented on a grid in real space.

To construct the matrix representation of a quasirelativistic Hamiltonian in a basis, one has two choices:

(a) to transform first the *Dirac operator* (or Dirac-Coulomb operator, etc.) to a quasirelativistic operator and to construct then the matrix representation of the latter, or
(b) to expand first the Dirac operator in a basis, and then transform the matrix representation of the Dirac operator to a quasirelativistic matrix.

We have seen[6] that (a) is generally not feasible and that even if it were, it would not be recommended since it requires a basis which is able to describe the *weakly singular* behavior of the wave function at the *position of a nucleus* correctly. The choice of a *regular basis* becomes highly problematic, especially in the context of a perturbation expansion of the quasirelativistic operator.

In (b) the only problem is the appropriate choice of the basis for the expansion of the Dirac operator. If one chooses the basis inappropriately, one runs into the problem of the so-called *variational collapse*[3,11] i.e., one finds physically meaningless eigenvalues in the *forbidden region* between the electronic ground state and the onset of the negative-energy continuum. The origin of this is that for a poorly chosen basis one fails to describe the kinetic energy correctly, even in the nonrelativistic limit (nrl). Various procedures to avoid the variational collapse have been proposed, among which the so-called *kinetic balance* has become standard.[12] Imposing the kinetic balance (see Sec. III), one achieves that both the nrl and the leading relativistic correction are correct and the errors of higher order corrections should decrease with increasing basis size.

The choice of a basis for the expansion of the Dirac operator is the *only approximation* that we introduce. The subsequent transformation to a quasirelativistic Hamiltonian matrix can be done numerically to any desired precision. In particular, we do not introduce any *ad hoc completeness insertions*.

Among the quasirelativistic methods in current use, the Douglas-Kroll-Hess (DKH) method[13,14] has always been implemented as a matrix transformation, including its modification by Barysz *et al.*[15] *Direct perturbation theory* (DPT) has been formulated and applied in two versions, one at the operator level[16] for exactly solvable problems and one [stationary direct perturbation theory (S-DPT)] at the matrix level[2] for numerical solutions. The *regular approximation*[17] (RA) has been formulated and implemented as an operator transformation with subsequent basis expansion. Only recently, a variant of RA at the matrix level has been studied.[18] The NESC (Ref. 9) has been a matrix theory from the very beginning. Very recently, a noniterative construction of a quasirelativistic Hamiltonian matrix has been implemented by Jensen and co-workers[19] in the DIRAC program package.

III. MATRIX REPRESENTATION OF THE DIRAC OPERATOR IN A KINETICALLY BALANCED BASIS

We start from the representation of the Dirac operator in a *finite regular kinetically balanced basis*, and we want to apply a transformation to it such that the result is a quasirelativistic Hamiltonian, again in a basis representation. The eigenstates of the new quasi-relativistic Hamiltonian will be *the same as those of the matrix representation of the Dirac Hamiltonian*. Unlike the FW transformation operator the transformation matrix can be constructed exactly, and it is not necessary to consider a specific expansion parameter. In the matrix theory a matrix $\mathbf{X}$ will play a similar role as the operator X at the operator level. The restriction of X to *bound states*, of which we have made some use in the first part,[6] has no obvious counterpart for $\mathbf{X}$.

Let us expand the upper and lower components φ and χ of the bispinor ψ in a given kinetically balanced basis, constructed from the m-dimensional regular basis $\{g_\mu\}$ for the upper component. We express the relation between χ and φ as[7]

$$\chi_k = \tilde{X}\varphi_k \tag{11}$$

for all state labels k, with $\tilde{X}$ the same for all m states expandable in the given basis. This $\tilde{X}$ cannot be identical with the X that satisfies Eq. (2) because the exact χ and φ cannot be expanded in a regular basis, but $\tilde{X}$ can be regarded as an approximation of X. We shall only need the matrix representation $\mathbf{X}$ [Eq. (16)] rather than $\tilde{X}$ itself.

$$\varphi = \sum_\mu a_\mu g_\mu, \tag{12}$$

$$\chi = \sum_\mu b_\mu f_\mu = \sum_\mu b_\mu \boldsymbol{\sigma}\mathbf{p} g_\mu = \tilde{X}\varphi = \sum_\mu a_\mu \tilde{X} g_\mu. \tag{13}$$

The scalar product from the left with g_ν or $\boldsymbol{\sigma}\mathbf{p}g_\nu$ is

$$\langle g_\nu|\varphi\rangle = \sum_\mu a_\mu \langle g_\nu|g_\mu\rangle, \tag{14}$$

$$\langle g_\nu|\boldsymbol{\sigma}\mathbf{p}|\chi\rangle = \sum_\mu b_\mu 2m\langle g_\nu|T|g_\mu\rangle = \langle g_\nu|\boldsymbol{\sigma}\mathbf{p}\tilde{X}|\varphi\rangle$$
$$= \sum_\mu a_\mu\langle g_\nu|\boldsymbol{\sigma}\mathbf{p}\tilde{X}|g_\mu\rangle. \quad (15)$$

We define the matrices **S**, **V**, **T**, **X**, and **U** through their matrix elements,

$$S_{\nu\mu} = \langle g_\nu|g_\mu\rangle, \quad V_{\nu\mu} = \langle g_\nu|V|g_\mu\rangle, \quad T_{\nu\mu} = \langle g_\nu|T|g_\mu\rangle,$$
$$X_{\nu\mu} = \langle g_\nu|\boldsymbol{\sigma}\mathbf{p}\tilde{X}|g_\mu\rangle; U_{\nu\mu} = \langle g_\nu|\boldsymbol{\sigma}\mathbf{p}V\boldsymbol{\sigma}\mathbf{p}|g_\mu\rangle, \quad (16)$$

and represent φ as well as χ by the vectors **a** and **b** of their coefficients a_μ and b_μ. We can then write Eq. (15) as

$$\mathbf{Tb} = \frac{1}{2m}\mathbf{Xa}, \quad \mathbf{b} = \frac{1}{2m}\mathbf{T}^{-1}\mathbf{Xa}, \quad (17)$$

which relates the expansion coefficients of χ and φ.

The expectation value of the Dirac operator for an electronic eigenstate is[4]

$$\langle D - mc^2\rangle = \frac{\langle\psi|D - mc^2|\psi\rangle}{\langle\psi|\psi\rangle}$$
$$= \frac{\langle\varphi|V|\varphi\rangle + 2c\ \mathrm{Re}\langle\chi|\boldsymbol{\sigma}\mathbf{p}|\varphi\rangle - 2mc^2\langle\chi|\chi\rangle + \langle\chi|V|\chi\rangle}{\langle\varphi|\varphi\rangle + \langle\chi|\chi\rangle}$$
$$= \frac{\mathbf{a}^\dagger\mathbf{Va} + 2mc\mathbf{a}^\dagger\mathbf{Tb} + 2mc\mathbf{b}^\dagger\mathbf{Ta} + \mathbf{b}^\dagger\mathbf{Ub} - 4m^2c^2\mathbf{b}^\dagger\mathbf{Tb}}{\mathbf{a}^\dagger\mathbf{Sa} + 2m\mathbf{b}^\dagger\mathbf{Tb}}. \quad (18)$$

Conditions for stationarity with respect to variations of **a** and **b** are, respectively,

$$0 = \mathbf{Va} + 2mc\mathbf{Tb} - E\mathbf{Sa}, \quad (19)$$

$$0 = 2mc\mathbf{Ta} + \mathbf{Ub} - 4m^2c^2\mathbf{Tb} - 2mE\mathbf{Tb}. \quad (20)$$

We get the same equations if we start from the Dirac equation in block form, insert the expansions (12) and (13), and take the scalar product from the left of the first equation with g_μ and the second one with f_μ. So these are the equations that we try to satisfy in a matrix representation of the Dirac equation.

We can, of course, write Eqs. (19) and (20) as an eigenvalue equation with a double dimension and a nonunit metric,

$$\begin{pmatrix} \mathbf{V} & 2mc\mathbf{T} \\ 2mc\mathbf{T} & \mathbf{U} - 4m^2c^2\mathbf{T} \end{pmatrix}\begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix} = E\begin{pmatrix} \mathbf{S} & 0 \\ 0 & 2m\mathbf{T} \end{pmatrix}\begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix}, \quad (21)$$

which is recognized as the *modified Dirac equation* of Dyall,[20] which was first proposed by one of the present authors.[3] Note that the modified Dirac equation results here automatically as a *condition for stationarity* of the energy expectation value and is not introduced in an *ad hoc* way.

IV. THE KEY RELATION AT THE MATRIX LEVEL

Let us now eliminate **b** from Eqs. (19) and (20) via Eq. (17), but let us explicitly take care of state labels k,

$$0 = \mathbf{Va}_k + c\mathbf{Xa}_k - E_k\mathbf{Sa}_k, \quad (22)$$

$$0 = 2mc\mathbf{Ta}_k + \frac{1}{2m}\mathbf{UT}^{-1}\mathbf{Xa}_k - 2mc^2\mathbf{Xa}_k - E_k\mathbf{Xa}_k. \quad (23)$$

We have two equations for the determination of **X**, $\mathbf{a}_k$, and E_k.

Let us premultiply Eq. (22) by $-\mathbf{XS}^{-1}$ and add this to Eq. (23). The result is

$$0 = \left(-\mathbf{XS}^{-1}\mathbf{V} - c\mathbf{XS}^{-1}\mathbf{X} + 2mc\mathbf{T} + \frac{1}{2m}\mathbf{UT}^{-1}\mathbf{X} - 2mc^2\mathbf{X}\right)\mathbf{a}_k. \quad (24)$$

This is an equation independent of E_k, which must hold for all $\mathbf{a}_k$, implying the matrix equation

$$\mathbf{X} = c^{-1}\mathbf{T} + \frac{1}{4m^2c^2}\mathbf{UT}^{-1}\mathbf{X} - \frac{1}{2mc^2}\mathbf{XS}^{-1}\mathbf{V} - \frac{1}{2mc}\mathbf{XS}^{-1}\mathbf{X}. \quad (25)$$

Having constructed **X** from Eq. (25), we have to build the quasi-relativistic Hamiltonian, from which we get the E_k and $\mathbf{a}_k$. There are various possibilities, the simplest of which starts from Eq. (22).

We define the *non-Hermitian* quasirelativistic Hamiltonian matrix as

$$\mathbf{L} = \mathbf{V} + c\mathbf{X} \quad (26)$$

such that $\mathbf{a}_k$ satisfies the eigenvalue equation

$$\mathbf{La}_k = E_k\mathbf{Sa}_k. \quad (27)$$

Other choices of quasirelativistic Hamiltonians will be discussed in Sec. V

In view of Eq. (26), we can write Eq. (25) alternatively as

$$\mathbf{X} = c^{-1}\mathbf{T} + \frac{1}{4m^2c^2}\mathbf{UT}^{-1}\mathbf{X} - \frac{1}{2mc^2}\mathbf{XS}^{-1}\mathbf{L} \quad (28)$$

and require that this is solved together with Eq. (26). **X** is uniquely defined if we require Eq. (24) for all eigenstates. While **S**, **T**, and **U** are constants, **X** and **L** change from iteration to iteration. Spin-orbit effects are entirely in the factor **U**. So the dimension of **U** is twice that of **S** and **T**, and of the iteration start for **X** and **L**. A spin-free theory is obtained if we replace **U** by its spin-free counterpart. Then, all operators become real, while they are complex otherwise. The nrl is obviously obtained if one neglects the terms of $O(c^{-2})$ on the right-hand side of Eq. (28).

$$\mathbf{X}_{\mathrm{nr}} = c^{-1}\mathbf{T}, \quad \mathbf{L}_{\mathrm{nr}} = \mathbf{V} + \mathbf{T} = \mathbf{H}_0. \quad (29)$$

V. HERMITIAN QUASIRELATIVISTIC HAMILTONIANS

Having constructed **X** independently of how we have achieved this, we can get the φ or the **a** from the non-Hermitian eigenvalue problem [Eq. (27)] or alternatively

from a Hermitian one [Eq. (33)] with a nonunit metric. From $\mathbf{a}$ we get $\mathbf{b}$, which represents χ, via Eq. (17).

Right eigenvectors $\mathbf{a}_k$ and $\mathbf{a}_l$ to different eigenvalues are not orthogonal. From the orthonormality of the full Dirac bispinors, we get

$$\delta_{\mu\nu} = \mathbf{a}_\mu^\dagger \tilde{\mathbf{S}} \mathbf{a}_\nu, \quad \tilde{\mathbf{S}} = \mathbf{S} + \frac{1}{2m}\mathbf{X}^\dagger \mathbf{T}^{-1}\mathbf{X}. \tag{30}$$

To formulate the *Hermitian* eigenvalue problem, we start from the energy expectation value [Eq. (18)] and use Eq. (17) in order to eliminate $\mathbf{b}$ in favor of $\mathbf{X}$

$$E = \frac{\mathbf{a}^\dagger \tilde{\mathbf{L}} \mathbf{a}}{\mathbf{a}^\dagger \tilde{\mathbf{S}} \mathbf{a}}, \tag{31}$$

$$\tilde{\mathbf{L}} = \mathbf{V} + c\mathbf{X} + c\mathbf{X}^\dagger + \frac{1}{4m^2}\mathbf{X}^\dagger \mathbf{T}^{-1}\mathbf{U}\mathbf{T}^{-1}\mathbf{X} - c^2\mathbf{X}^\dagger \mathbf{T}^{-1}\mathbf{X}. \tag{32}$$

This is the expectation value of a Hermitian operator, but with a nonunit metric. If we make this stationary with respect to the variation of $\mathbf{a}$, we get the condition

$$\tilde{\mathbf{L}}\mathbf{a} = E\tilde{\mathbf{S}}\mathbf{a}. \tag{33}$$

This Hermitian eigenvalue problem is much more complicated than, but equivalent to, the non-Hermitian one [Eq. (27)]. The eigenvector $\mathbf{a}$ still represents the upper-component spinor φ.

Three more reformulations of $\tilde{\mathbf{L}}$ are possible. They are equivalent to Eq. (32), provided that $\mathbf{X}$ satisfies the exact relation (28). These are

$$\mathbf{L}' = \tilde{\mathbf{S}}\mathbf{S}^{-1}\mathbf{L}, \tag{34}$$

$$\mathbf{L}'^\dagger = \mathbf{L}\mathbf{S}^{-1}\tilde{\mathbf{S}}, \tag{35}$$

$$\bar{\mathbf{L}} = \tfrac{1}{2}(\mathbf{L}' + \mathbf{L}'^\dagger). \tag{36}$$

It is trivial that Eq. (27) implies that $\mathbf{L}'\mathbf{a} = \tilde{\mathbf{S}}\mathbf{a}$ and vice versa.

To show the equivalence of $\mathbf{L}'$ and $\tilde{\mathbf{L}}$, we start from Eq. (34) and assume Eq. (28),

$$\begin{aligned}\mathbf{L}' &= \left(\mathbf{S} + \frac{1}{2m}\mathbf{X}^\dagger \mathbf{T}^{-1}\mathbf{X}\right)\mathbf{S}^{-1}\mathbf{L} = \mathbf{L} + \frac{1}{2m}\mathbf{X}^\dagger \mathbf{T}^{-1}(\mathbf{X}\mathbf{S}^{-1}\mathbf{L}) \\ &= \mathbf{L} + \frac{1}{2m}\mathbf{X}^\dagger \mathbf{T}^{-1}\left(2mc\mathbf{T} - 2mc^2\mathbf{X} + \frac{1}{2m}\mathbf{U}\mathbf{T}^{-1}\mathbf{X}\right) = \mathbf{V} \\ &\quad + c\mathbf{X} + c\mathbf{X}^\dagger - c^2\mathbf{X}^\dagger \mathbf{T}^{-1}\mathbf{X} + \frac{1}{4m^2}\mathbf{X}^\dagger \mathbf{T}^{-1}\mathbf{U}\mathbf{T}^{-1}\mathbf{X} = \tilde{\mathbf{L}}.\end{aligned} \tag{37}$$

This is obviously Hermitian and agrees with Eq. (32), but only so if $\mathbf{X}$ satisfies Eq. (28). If this is not the case, the symmetrized form (36) is preferable to Eq. (34).

The most robust expression is obviously Eq. (32) since it is valid even if $\mathbf{X}$ is not exact. The energy will always converge faster if it is evaluated in terms of $\tilde{\mathbf{L}}$ rather than $\mathbf{L}'$ or $\bar{\mathbf{L}}$.

The equivalence of $\mathbf{L}'$ or $\tilde{\mathbf{L}}$, under the condition that Eq. (28) holds, implies the following interesting identity, first explicitly formulated by Filatov,[8] starting from an earlier paper by Dyall,[9]

$$\mathbf{S}^{-1}\mathbf{L} = \tilde{\mathbf{S}}^{-1}\tilde{\mathbf{L}}. \tag{38}$$

This can be regarded as a condition equivalent to Eq. (28).

Finally, we can care for the matrix counterpart of the FW spinor [Eq. (8)], represented by the vector

$$\mathbf{d} = \mathbf{S}^{-1/2}\tilde{\mathbf{S}}^{1/2}\mathbf{a}, \tag{39}$$

such that

$$\langle \phi_k | \phi_l \rangle = \mathbf{d}_k^\dagger \mathbf{S}\mathbf{d}_l = \mathbf{a}_k^\dagger \tilde{\mathbf{S}}\mathbf{a}_l = \delta_{kl} \tag{40}$$

and that $\mathbf{d}$ is solution of the eigenvalue problem

$$\mathbf{L}^+\mathbf{d} = \frac{1}{2}\{\mathbf{S}^{1/2}\tilde{\mathbf{S}}^{1/2}\mathbf{S}^{-1}\mathbf{L}\tilde{\mathbf{S}}^{-1/2}\mathbf{S}^{1/2} + \text{H.c.}\}\mathbf{d} = E\mathbf{S}\mathbf{d}. \tag{41}$$

This is the matrix analog of the FW transformation, restricted to the electronic part. The full transformation will be considered in the next section. If Eq. (38) holds, the expression

$$\tilde{\mathbf{L}}^+ = \mathbf{S}^{1/2}\tilde{\mathbf{S}}^{-1/2}\tilde{\mathbf{L}}\tilde{\mathbf{S}}^{-1/2}\mathbf{S}^{1/2} \tag{42}$$

is equivalent to $\mathbf{L}^+$. This $\tilde{\mathbf{L}}^+$ is related to $\mathbf{L}^+$ in the same way as $\tilde{\mathbf{L}}$ is related to $\bar{\mathbf{L}}$.

VI. THE MATRIX ANALOG OF THE FW TRANSFORMATION

The most convenient formulation of the FW transformation at the operator level is in terms of a sequence of two transformations, one that achieves block diagonalization and another one that reestablishes the normalization. The first of these is

$$\begin{aligned}L_T &= \begin{pmatrix} 1 & X^\dagger \\ -X & 1 \end{pmatrix}\begin{pmatrix} V & c\boldsymbol{\sigma}\mathbf{p} \\ c\boldsymbol{\sigma}\mathbf{p} & -2mc^2 + V \end{pmatrix}\begin{pmatrix} 1 & -X^\dagger \\ X & 1 \end{pmatrix} \\ &= \begin{pmatrix} L_T^{11} & L_T^{12} \\ L_T^{21} & L_T^{22} \end{pmatrix},\end{aligned} \tag{43}$$

$$L_T^{11} = V + c\boldsymbol{\sigma}\mathbf{p}X + cX^\dagger\boldsymbol{\sigma}\mathbf{p} + X^\dagger(V - 2mc^2)X, \tag{44}$$

$$L_T^{12} = -VX^\dagger + c\boldsymbol{\sigma}\mathbf{p} - cX^\dagger\boldsymbol{\sigma}\mathbf{p}X^\dagger + X^\dagger V - 2mc^2X^\dagger, \tag{45}$$

$$L_T^{21} = -XV - cX\boldsymbol{\sigma}\mathbf{p}X + c\boldsymbol{\sigma}\mathbf{p} + VX - 2mc^2X, \tag{46}$$

$$L_T^{22} = XVX^\dagger - cX\boldsymbol{\sigma}\mathbf{p} - c\boldsymbol{\sigma}\mathbf{p}X^\dagger + V - 2mc^2. \tag{47}$$

Block diagonalization is achieved if X satisfies Eq. (3). The second transformation is

$$L_{\text{FW}} = W_T^\dagger L_T W_T, \tag{48}$$

$$W_T = \begin{pmatrix} (1+X^\dagger X)^{-1/2} & 0 \\ 0 & (1+XX^\dagger)^{-1/2} \end{pmatrix}. \tag{49}$$

One arrives at the corresponding transformation at the matrix level in the following way:

$$\mathbf{L}_T = \begin{pmatrix} 1 & \mathbf{x}^\dagger \\ -\mathbf{x} & 1 \end{pmatrix} \begin{pmatrix} \mathbf{v} & c\boldsymbol{\pi}^\dagger \\ c\boldsymbol{\pi} & -2mc^2 + \mathbf{u} \end{pmatrix} \begin{pmatrix} 1 & -\mathbf{x}^\dagger \\ \mathbf{x} & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{L}_\mathbf{T}^{11} & \mathbf{L}_\mathbf{T}^{12} \\ \mathbf{L}_\mathbf{T}^{21} & \mathbf{L}_\mathbf{T}^{22} \end{pmatrix}, \tag{50}$$

$$\mathbf{L}_\mathbf{T}^{11} = \mathbf{v} + c\boldsymbol{\pi}^\dagger\mathbf{x} + c\mathbf{x}^\dagger\boldsymbol{\pi} + \mathbf{x}^\dagger\mathbf{u}x - 2mc^2\mathbf{x}^\dagger\mathbf{x}, \tag{51}$$

$$\mathbf{L}_\mathbf{T}^{12} = -\mathbf{v}\mathbf{x}^\dagger + c\boldsymbol{\pi}^\dagger - c\mathbf{x}^\dagger\boldsymbol{\pi}\mathbf{x}^\dagger + \mathbf{x}^\dagger\mathbf{u} - 2mc^2\mathbf{x}^\dagger, \tag{52}$$

$$\mathbf{L}_\mathbf{T}^{21} = -\mathbf{x}\mathbf{v} - c\mathbf{x}\boldsymbol{\pi}^\dagger\mathbf{x} + c\boldsymbol{\pi} + \mathbf{u}\mathbf{x} - 2mc^2\mathbf{x}, \tag{53}$$

$$\mathbf{L}_\mathbf{T}^{22} = \mathbf{x}\mathbf{v}\mathbf{x}^\dagger - c\mathbf{x}\boldsymbol{\pi}^\dagger - c\boldsymbol{\pi}\mathbf{x}^\dagger + \mathbf{u} - 2mc^2, \tag{54}$$

where the matrix elements of $\mathbf{v}$, $\mathbf{u}$, π, and $\mathbf{x}$ are given in terms of a basis $\{g_\mu\}$ for the upper ("large") and $\{f_\mu\}$ for the lower ("small") component, which are each assumed to be an *orthonormal* set. The basis sets $\{g_\mu\}$ and $\{f_\mu\}$ need not have the same dimension, so $\mathbf{x}$ and π are not necessarily square matrices,

$$v_{\mu,\nu} = \langle g_\mu | V | g_\nu \rangle, \quad u_{\mu,\nu} = \langle f_\mu | V | f_\nu \rangle, \tag{55}$$

$$\pi_{\mu,\nu} = \langle f_\mu | \boldsymbol{\sigma}\mathbf{p} | g_\nu \rangle, \quad x_{\mu,\nu} = \langle f_\mu | \tilde{X} | g_\nu \rangle. \tag{56}$$

The matrix $\mathbf{x}$ is determined by the condition that $\mathbf{L}_\mathbf{T}^{12} = \mathbf{L}_\mathbf{T}^{21} = 0$.

For the second transformation we get

$$\mathbf{W}_\mathbf{T}^{11} = (1+\mathbf{x}^\dagger\mathbf{x})^{-1/2}, \tag{57}$$

$$\mathbf{W}_\mathbf{T}^{22} = (1+\mathbf{x}\mathbf{x}^\dagger)^{-1/2}, \tag{58}$$

$$\mathbf{W}_\mathbf{T}^{12} = \mathbf{W}_\mathbf{T}^{21} = 0, \tag{59}$$

such that finally

$$\mathbf{L}_\mathbf{FW}^{11} = \mathbf{W}_\mathbf{T}^{11}\mathbf{L}_\mathbf{T}^{11}\mathbf{W}_\mathbf{T}^{11}, \tag{60}$$

$$\mathbf{L}_\mathbf{FW}^{22} = \mathbf{W}_\mathbf{T}^{22}\mathbf{L}_\mathbf{T}^{22}\mathbf{W}_\mathbf{T}^{22}, \tag{61}$$

$$\mathbf{L}_\mathbf{FW}^{12} = \mathbf{L}_\mathbf{FW}^{21} = 0. \tag{62}$$

If the basis sets for two-component spinors are *not orthonormal* but are associated with overlap integrals,

$$s_{\mu,\nu} = \langle g_\mu | g_\nu \rangle, \quad t_{\mu,\nu} = \langle f_\mu | f_\nu \rangle, \tag{63}$$

the transformation becomes

$$\mathbf{L}_T = \begin{pmatrix} \mathbf{s} & \mathbf{x}^\dagger \\ -\mathbf{x} & \mathbf{t} \end{pmatrix} \begin{pmatrix} \mathbf{s}^{-1} & 0 \\ 0 & \mathbf{t}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{v} & c\boldsymbol{\pi}^\dagger \\ c\boldsymbol{\pi} & -2mc^2\mathbf{t} + \mathbf{u} \end{pmatrix} \times \begin{pmatrix} \mathbf{s}^{-1} & 0 \\ 0 & \mathbf{t}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{s} & -\mathbf{x}^\dagger \\ \mathbf{x} & \mathbf{t} \end{pmatrix} \tag{64}$$

and the blocks of the transformed matrix are

$$\mathbf{L}_\mathbf{T}^{11} = \mathbf{v} + c\boldsymbol{\pi}^\dagger\mathbf{t}^{-1}\mathbf{x} + c\mathbf{x}^\dagger\mathbf{t}^{-1}\boldsymbol{\pi} + \mathbf{x}^\dagger\mathbf{t}^{-1}\mathbf{u}\mathbf{t}^{-1}\mathbf{x} - 2mc^2\mathbf{x}^\dagger\mathbf{t}^{-1}\mathbf{x}, \tag{65}$$

$$\mathbf{L}_\mathbf{T}^{12} = -\mathbf{v}\mathbf{s}^{-1}\mathbf{x}^\dagger + c\boldsymbol{\pi}^\dagger - c\mathbf{x}^\dagger\mathbf{t}^{-1}\boldsymbol{\pi}\mathbf{s}^{-1}\mathbf{x}^\dagger + \mathbf{x}^\dagger\mathbf{t}^{-1}\mathbf{u} - 2mc^2\mathbf{x}^\dagger, \tag{66}$$

$$\mathbf{L}_\mathbf{T}^{21} = -\mathbf{x}\mathbf{s}^{-1}\mathbf{v} - c\mathbf{x}\mathbf{s}^{-1}\boldsymbol{\pi}^\dagger\mathbf{t}^{-1}\mathbf{x} + c\boldsymbol{\pi} + \mathbf{u}\mathbf{t}^{-1}\mathbf{x} - 2mc^2\mathbf{x}, \tag{67}$$

$$\mathbf{L}_\mathbf{T}^{22} = \mathbf{x}\mathbf{s}^{-1}\mathbf{v}\mathbf{s}^{-1}\mathbf{x}^\dagger - c\mathbf{x}\mathbf{s}^{-1}\boldsymbol{\pi}^\dagger - c\boldsymbol{\pi}\mathbf{s}^{-1}\mathbf{x}^\dagger + \mathbf{u} - 2mc^2\mathbf{t}. \tag{68}$$

To construct the second transformation we note that

$$\begin{pmatrix} \mathbf{s} & \mathbf{x}^\dagger \\ -\mathbf{x} & \mathbf{t} \end{pmatrix} \begin{pmatrix} \mathbf{s}^{-1} & 0 \\ 0 & \mathbf{t}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{s} & -\mathbf{x}^\dagger \\ \mathbf{x} & \mathbf{t} \end{pmatrix} = \begin{pmatrix} \mathbf{s}+\mathbf{x}^\dagger\mathbf{t}^{-1}\mathbf{x} & 0 \\ 0 & \mathbf{t}+\mathbf{x}\mathbf{s}^{-1}\mathbf{x}^\dagger \end{pmatrix}. \tag{69}$$

The full transformation is unitary if

$$\mathbf{W}_\mathbf{T}^{11\dagger}\mathbf{s}^{-1}(\mathbf{s}+\mathbf{x}^\dagger\mathbf{t}^{-1}\mathbf{x})\mathbf{s}^{-1}\mathbf{W}_\mathbf{T}^{11} = \mathbf{s}, \tag{70}$$

$$\mathbf{W}_\mathbf{T}^{22\dagger}\mathbf{t}^{-1}(\mathbf{t}+\mathbf{x}^\dagger\mathbf{s}^{-1}\mathbf{x})\mathbf{t}^{-1}\mathbf{W}_\mathbf{T}^{22} = \mathbf{t}, \tag{71}$$

$$\mathbf{W}_\mathbf{T}^{12} = \mathbf{W}_\mathbf{T}^{21} = 0. \tag{72}$$

This is achieved if

$$\mathbf{W}_\mathbf{T}^{11} = \mathbf{s}(\mathbf{s}+\mathbf{x}^\dagger\mathbf{t}^{-1}\mathbf{x})^{-1/2}\mathbf{s}^{1/2}, \tag{73}$$

$$\mathbf{W}_\mathbf{T}^{22} = \mathbf{t}(\mathbf{t}+\mathbf{x}\mathbf{s}^{-1}\mathbf{x}^\dagger)^{-1/2}\mathbf{t}^{1/2}, \tag{74}$$

$$\mathbf{W}_\mathbf{T}^{12} = \mathbf{W}_\mathbf{T}^{21} = 0 \tag{75}$$

such that, finally,

$$\mathbf{L}_\mathbf{FW}^{11} = \mathbf{W}_\mathbf{T}^{11\dagger}\mathbf{s}^{-1}\mathbf{L}_\mathbf{T}^{11}\mathbf{s}^{-1}\mathbf{W}_\mathbf{T}^{11}, \tag{76}$$

$$\mathbf{L}_\mathbf{FW}^{22} = \mathbf{W}_\mathbf{T}^{22}\mathbf{t}^{-1}\mathbf{L}_\mathbf{T}^{22}\mathbf{t}^{-1}\mathbf{W}_\mathbf{T}^{22\dagger}, \tag{77}$$

$$\mathbf{L}_\mathbf{FW}^{12} = \mathbf{L}_\mathbf{FW}^{21} = 0. \tag{78}$$

The only approximation that we make is in the choice of the basis. The transformation does not cause any problems. There are no singularities as in the FW transformation at the operator level.[6] The eigenvalues of the matrix representation of the Dirac operator are exactly the same as the eigenvalues of the transformed matrix. One must, of course, choose the basis, such that the variational collapse[3,11] is avoided. This is conveniently done with a kinetically balanced basis as defined in Sec. III with the overlap matrices $\mathbf{s}=\mathbf{S}$ for g and $\mathbf{t}=2m\mathbf{T}$ for f. We get, noting that $\boldsymbol{\pi}=2m\mathbf{T}$,

$$\mathbf{L}_\mathbf{T}^{11} = \mathbf{V} + c\mathbf{X} + c\mathbf{X}^\dagger + \frac{1}{4m^2}\mathbf{X}^\dagger\mathbf{T}^{-1}\mathbf{U}\mathbf{T}^{-1}\mathbf{X} - c^2\mathbf{X}^\dagger\mathbf{T}^{-1}\mathbf{X}, \tag{79}$$

$$\mathbf{L}_\mathbf{T}^{12} = -\mathbf{V}\mathbf{S}^{-1}\mathbf{X}^\dagger + 2mc\mathbf{T} - c\mathbf{X}^\dagger\mathbf{S}^{-1}\mathbf{X}^\dagger + \frac{1}{2m}\mathbf{X}^\dagger\mathbf{T}^{-1}\mathbf{U} - 2mc^2\mathbf{X}^\dagger, \tag{80}$$

$$\mathbf{L}_{\mathbf{T}}^{21} = -\mathbf{X}\mathbf{S}^{-1}\mathbf{V} - c\mathbf{X}\mathbf{S}^{-1}\mathbf{X} + 2mc\mathbf{T} + \frac{1}{2m}\mathbf{U}\mathbf{T}^{-1}\mathbf{X} - 2mc^2\mathbf{X}, \tag{81}$$

$$\mathbf{L}_{\mathbf{T}}^{22} = \mathbf{X}\mathbf{S}^{-1}\mathbf{V}\mathbf{S}^{-1}\mathbf{X}^{\dagger} - 2mc\mathbf{X}\mathbf{S}^{-1}\mathbf{T} - 2mc\mathbf{T}\mathbf{S}^{-1}\mathbf{X}^{\dagger} + \mathbf{U} - 4m^2c^2\mathbf{T}, \tag{82}$$

$$\mathbf{W}_{\mathbf{T}}^{11} = \mathbf{S}\left(\mathbf{S} + \frac{1}{2m}\mathbf{X}^{\dagger}\mathbf{T}^{-1}\mathbf{X}\right)^{-1/2}\mathbf{S}^{1/2}, \tag{83}$$

$$\mathbf{W}_{\mathbf{T}}^{22} = (2m\mathbf{T})(2m\mathbf{T} + \mathbf{X}\mathbf{S}^{-1}\mathbf{X}^{\dagger})^{-1/2}(2m\mathbf{T})^{1/2}, \tag{84}$$

$$\mathbf{W}_{\mathbf{T}}^{12} = \mathbf{W}_{\mathbf{T}}^{21} = 0. \tag{85}$$

If $\mathbf{X}$ satisfies Eq. (25), the transformed matrix is block diagonal.

The block [Eq. (76)] with Eqs. (79) and (83) is identical with Eq. (42). The derivation given ins Sec. V has been simpler, but the present one has the advantage that we also get a complementary quasirelastivistic Hamiltonian for positrons.

The block $\mathbf{L}_{\mathbf{T}}^{22}$, which represents positronic states, will, however, be a poor approximation for the latter because in the *counterpart of the kinetic balance for positronic states* the role of the f and g basis sets is inverted. For a better description of positronic states one will take basis sets, in terms of which $\mathbf{t}=\mathbf{S}$ and $\mathbf{s}=2m\mathbf{T}$. One must further interchange $\mathbf{V}$ with $\mathbf{U}$ and replace $\boldsymbol{\pi}$ and $\mathbf{X}$ by their Hermitian adjoints.

$$\mathbf{L}_{\mathbf{T}}^{22} = \mathbf{V} - c\mathbf{X} - c\mathbf{X}^{\dagger} + \frac{1}{4m^2}\mathbf{X}^{\dagger}\mathbf{T}^{-1}\mathbf{U}\mathbf{T}^{-1}\mathbf{X} - 2mc^2\mathbf{S}, \tag{86}$$

$$\mathbf{L}_{\mathbf{T}}^{21} = \mathbf{V}\mathbf{S}^{-1}\mathbf{X}^{\dagger} + 2mc\mathbf{T} - c\mathbf{X}^{\dagger}\mathbf{S}^{-1}\mathbf{X}^{\dagger} - \frac{1}{2m}\mathbf{X}^{\dagger}\mathbf{T}^{-1}\mathbf{U} - 2mc^2\mathbf{X}^{\dagger}, \tag{87}$$

$$\mathbf{L}_{\mathbf{T}}^{12} = \mathbf{X}\mathbf{S}^{-1}\mathbf{V} - c\mathbf{X}\mathbf{S}^{-1}\mathbf{X} + 2mc\mathbf{T} - \frac{1}{2m}\mathbf{U}\mathbf{T}^{-1}\mathbf{X} - 2mc^2\mathbf{X}, \tag{88}$$

$$\mathbf{L}_{\mathbf{T}}^{11} = \mathbf{X}\mathbf{S}^{-1}\mathbf{V}\mathbf{S}^{-1}\mathbf{X}^{\dagger} + 2mc\mathbf{X}\mathbf{S}^{-1}\mathbf{T} + 2mc\mathbf{T}\mathbf{S}^{-1}\mathbf{X}^{\dagger} + \mathbf{U} - 2mc^2\mathbf{X}\mathbf{S}^{-1}\mathbf{X}^{\dagger}, \tag{89}$$

$$\mathbf{W}_{\mathbf{T}}^{22} = \mathbf{S}\left(\mathbf{S} + \frac{1}{2m}\mathbf{X}^{\dagger}\mathbf{T}^{-1}\mathbf{X}\right)^{-1/2}\mathbf{S}^{1/2}, \tag{90}$$

$$\mathbf{W}_{\mathbf{T}}^{11} = (2m\mathbf{T})(2m\mathbf{T} + \mathbf{X}\mathbf{S}^{-1}\mathbf{X}^{\dagger})^{-1/2}(2m\mathbf{T})^{1/2}, \tag{91}$$

$$\mathbf{W}_{\mathbf{T}}^{12} = \mathbf{W}_{\mathbf{T}}^{21} = 0. \tag{92}$$

The matrix $\mathbf{X}$ that leads to block diagonalization in this context differs by a factor of -1 from that appropriate for electrons.

VII. CONSTRUCTION OF THE MATRIX X

We have to solve Eq. (25) for $\mathbf{X}$. This is the matrix counterpart of the key relation (3). Equation (25) is a quadratic equation for the unknown matrix $\mathbf{X}$. The solution of this equation is entirely a problem of numerical mathematics. Physical pictures or expansion parameters are rather irrelevant in this context. The main options for a solution are *by iteration*, either linearly or quadratically.

It is shown in the Appendix that a *linear* matrix equation [like Eq. (25) without the last term on the right-hand side] can be solved in a *noniterative way*, i.e., in terms of elementary matrix operations such as multiplication, inversion, and diagonalization (even if for these operations iterative algorithms are used). A quadratic matrix equation requires an iterative approach, either a linearly or a quadratically convergent one.

We can construct $\mathbf{X}$ either as a solution of the quadratic matrix equation [Eq. (25)] or that of the coupled system of pseudolinear equations [Eqs. (28) and (26)].

A. Simple linear iteration

The formally simplest iterative solution of Eq. (28) is by means of the algorithm

$$\mathbf{X}^{(n+1)} = c^{-1}\mathbf{T} + \frac{1}{4m^2c^2}\mathbf{U}\mathbf{T}^{-1}\mathbf{X}^{(n)} - \frac{1}{2mc^2}\mathbf{X}^{(n)}\mathbf{S}^{-1}\mathbf{L}^{(n)}, \tag{93}$$

$$\mathbf{L}^{(n)} = \mathbf{V} + c\mathbf{X}^{(n)}. \tag{94}$$

It is inspired by DPT, and the start iteration is the nrl [Eq. (29)].

One should like to have all terms linear in $\mathbf{X}$ on the left-hand side. This will be achieved in the following subsection. It is simple to shift *two* of the *three* terms linear in $\mathbf{X}$ to the left-hand side. This leads us to the following two possibilities:

$$\mathbf{X}^{(n+1)}\left(1 + \frac{1}{2mc^2}\mathbf{S}^{-1}\mathbf{L}^{(n)}\right) = c^{-1}\mathbf{T} + \frac{1}{4m^2c^2}\mathbf{U}\mathbf{T}^{-1}\mathbf{X}^{(n)}, \tag{95}$$

$$\mathbf{X}^{(n+1)} = \left(c^{-1}\mathbf{T} + \frac{1}{4m^2c^2}\mathbf{U}\mathbf{T}^{-1}\mathbf{X}^{(n)}\right)\left[1 + \frac{1}{2mc^2}\mathbf{S}^{-1}\mathbf{L}^{(n)}\right]^{-1} \tag{96}$$

or

$$\left(1 - \frac{1}{4m^2c^2}\mathbf{U}\mathbf{T}^{-1}\right)\mathbf{X}^{(n+1)} = c^{-1}\mathbf{T} - \frac{1}{2mc^2}\mathbf{S}^{-1}\mathbf{L}^{(n)}, \tag{97}$$

$$\mathbf{X}^{(n+1)} = \left[1 - \frac{1}{4m^2c^2}\mathbf{U}\mathbf{T}^{-1}\right]^{-1}\left[c^{-1}\mathbf{T} - \frac{1}{2mc^2}\mathbf{S}^{-1}\mathbf{L}^{(n)}\right]. \tag{98}$$

In the iteration start one neglects $\mathbf{L}^{(0)}$. Scheme (98) is in the spirit of the RA,[17] and the start iteration is the *zeroth-order regular approximation* (ZORA) in the matrix representation (see Sec. IX E).

For all three schemes, the rate of convergence for $\mathbf{X}$ depends on that for $\mathbf{L}$. We know that $\tilde{\mathbf{L}}$ converges faster than $\mathbf{L}$. This suggests the use of relation (38) in order to eliminate $\mathbf{L}$ in favor of $\tilde{\mathbf{L}}$. We hence replace Eq. (93) by

$$\mathbf{X}^{(n+1)} = c^{-1}\mathbf{T} + \frac{1}{4m^2c^2}\mathbf{U}\mathbf{T}^{-1}\mathbf{X}^{(n)} - \frac{1}{2mc^2}\mathbf{X}^{(n)}(\tilde{\mathbf{S}}^{(n)})^{-1}\tilde{\mathbf{L}}^{(n)}. \tag{99}$$

This equation is also found in Filatov's Comment[8] as his interpretion of an iteration scheme implicit in Dyall's paper.[9]

For the other two simple linear iterations we replace Eq. (96) by

$$\mathbf{X}^{(n+1)} = \left(c^{-1}\mathbf{T} + \frac{1}{4m^2c^2}\mathbf{U}\mathbf{T}^{-1}\mathbf{X}^{(n)}\right) \times \left[1 + \frac{1}{2mc^2}(\tilde{\mathbf{S}}^{(n)})^{-1}\tilde{\mathbf{L}}^{(n)}\right]^{-1} \tag{100}$$

and Eq. (98) by

$$\mathbf{X}^{(n+1)} = \left(1 - \frac{1}{4m^2c^2}\mathbf{U}\mathbf{T}^{-1}\right)^{-1}\left[c^{-1}\mathbf{T} - \frac{1}{2mc^2}(\tilde{\mathbf{S}}^{(n)})^{-1}\tilde{\mathbf{L}}^{(n)}\right]. \tag{101}$$

The start iteration for scheme (101) is the *infinite-order regular approximation*[21] (IORA) in the matrix representation (see Sec. IX E).

B. Refined linear iteration

Equation (28) can be solved for $\mathbf{X}$ in a single step if one switches to a basis, in which either $\mathbf{L}$ or $\mathbf{U}$ is diagonal. An iteration is necessary, nevertheless, because we have to repeat the same procedure with an updated $\mathbf{L}$.

Let us first choose a basis $\mathbf{a}_k$ in which $\mathbf{L}$ is diagonal, in the sense that

$$\mathbf{L}^{(n)}\mathbf{a}_\mu^{(n)} = E_\mu^{(n)}\mathbf{S}\mathbf{a}_\mu^{(n)}. \tag{102}$$

We then try to solve

$$\left(1 - \frac{1}{4m^2c^2}\mathbf{U}\mathbf{T}^{-1} + \frac{1}{2mc^2}E_\mu^{(n)}\right)\mathbf{X}^{(n+1)}\mathbf{a}_\mu^{(n)} = c^{-1}\mathbf{T}\mathbf{a}_\mu^{(n)}. \tag{103}$$

This is nothing but Eq. (23), which could also be written as Eq. (20), which characterizes the elimination of the small component (at the matrix level). There are two alternative strategies to solve Eq. (103). One is to construct the vectors

$$\mathbf{k}_\mu^{(n)} = \left\{c - \frac{1}{4m^2c}\mathbf{U}\mathbf{T}^{-1} + \frac{1}{2mc}E_\mu^{(n)}\right\}^{-1}\mathbf{T}\mathbf{a}_\mu^{(n)} \tag{104}$$

from $E_\mu^{(n)}$ and $\mathbf{a}_\mu^{(n)}$ of the previous iteration and then to obtain $\mathbf{X}^{(n+1)}$ from the set of equations

$$\mathbf{X}^{(n+1)}\mathbf{a}_\mu^{(n)} = \mathbf{k}_\mu^{(n)} \tag{105}$$

combined to the matrix equation

$$\mathbf{K}^{(n+1)} = \mathbf{X}^{(n+1)}\mathbf{A}^{(n)}, \quad \mathbf{X}^{(n+1)} = \mathbf{K}^{(n)}\{\mathbf{A}^{(n)}\}^{-1}, \tag{106}$$

with the vectors $\mathbf{k}_\mu^{(n)}$ combined to a matrix $\mathbf{K}^{(n)}$ and the vectors $\mathbf{a}_\mu^{(n)}$ combined to a matrix $\mathbf{A}^{(n)}$. From $\mathbf{X}^{(n+1)}$ we get $\mathbf{L}^{(n+1)}$, and from this the $E_p^{(n+1)}$ and $\mathbf{a}_\mu^{(n+1)}$.

An alternative strategy, more in the spirit of the traditional elimination of the small component (ESC), is to solve the eigenvalue problem with an energy-dependent $\mathbf{L}(E)$

$$\mathbf{L}(E_\mu^{(n)})\mathbf{a}_\mu^{(n+1)} = E_\mu^{(n+1)}\mathbf{S}\mathbf{a}_\mu^{(n+1)}, \tag{107}$$

$$\mathbf{L}(E_\mu^{(n)}) = \mathbf{V} + \left\{1 - \frac{1}{4m^2c^2}\mathbf{U}\mathbf{T}^{-1} + \frac{1}{2mc^2}E_\mu^{(n)}\right\}^{-1}\mathbf{T} \tag{108}$$

iteratively for each μ separately and to construct the vectors $\mathbf{k}$ and from them the matrix $\mathbf{X}$ only when convergence has been reached for all μ. This is very tedious and has actually been considered a decade ago by van Lenthe *et al.*[22] We are therefore not pursuing this possibility.

In a basis, in which $\mathbf{U}$ is diagonal, we proceed in the following way. We choose a basis h_μ, such that

$$\mathbf{U}\mathbf{h}_\mu = U_\mu\mathbf{T}\mathbf{h}_\mu, \tag{109}$$

which implies (since $\mathbf{U}$ is Hermitian)

$$\mathbf{h}_\mu^\dagger\mathbf{U} = U_\mu\mathbf{h}_\mu^\dagger\mathbf{T}. \tag{110}$$

We then get

$$\mathbf{h}_\mu^\dagger\mathbf{X}^{(n+1)}\left(1 - \frac{1}{4m^2c^2}U_\mu + \frac{1}{2mc^2}\mathbf{S}^{-1}\mathbf{L}^{(n)}\right) = c^{-1}\mathbf{h}_\mu^\dagger\mathbf{T} \tag{111}$$

or the Hermitian conjugate

$$\left(c - \frac{1}{4m^2c}U_\mu + \frac{1}{2mc}\mathbf{L}^{\dagger(n)}\mathbf{S}^{-1}\right)\mathbf{X}^{\dagger(n+1)}\mathbf{h}_\mu = \mathbf{T}\mathbf{h}_\mu. \tag{112}$$

The corresponding iterative scheme is somewhat simpler than that based on Eq. (102) since the vectors $\mathbf{h}_\mu$ and the eigenvalues U_μ do not change during the iterations, so only $\mathbf{X}$ and $\mathbf{L}$ have to be updated. We construct

$$\mathbf{g}_\mu^{(n+1)} = \mathbf{X}^{\dagger(n+1)}\mathbf{h}_\mu = \left(c - \frac{1}{4m^2c}U_\mu + \frac{1}{2mc}\mathbf{L}^{(n)}\mathbf{S}^{-1}\right)^{-1}\mathbf{T}\mathbf{h}_\mu. \tag{113}$$

Combining $\mathbf{g}_{\mu^{(n+1)}}$ and $\mathbf{h}_\mu$ to matrices $\mathbf{G}^{(n+1)}$ and $\mathbf{H}$, respectively, we get

$$\mathbf{X}^{\dagger(n+1)} = \mathbf{G}^{(n+1)}\mathbf{H}^{-1}, \tag{114}$$

from which we obtain $\mathbf{X}^{(n+1)}$ as the Hermitian conjugate.

For either choice of Eq. (103) or (111), the rate of convergence for $\mathbf{X}$ depends on that for $\mathbf{L}$. We know that $\tilde{\mathbf{L}}$ converges faster than $\mathbf{L}$. This suggests again the use of relation (38) in order to eliminate $\mathbf{L}$ in favor of $\tilde{\mathbf{L}}$. We hence replace Eq. (102) by

$$\tilde{\mathbf{L}}^{(n)}\mathbf{a}_\mu^{(n)} = E_\mu^{(n)}\tilde{\mathbf{S}}\mathbf{a}_\mu^{(n)}. \tag{115}$$

This relates the approach based on Eq. (103) to the NESC of Dyall.[9]

We must further replace Eqs. (111) and (112) by

$$\mathbf{h}_\mu^\dagger \mathbf{X}^{(n+1)} \left(1 - \frac{1}{4m^2c^2} U_\mu + \frac{1}{2mc^2} \tilde{\mathbf{S}}^{-1} \tilde{\mathbf{L}}^{(n)}\right) = c^{-1} \mathbf{h}_\mu^\dagger \mathbf{T}, \tag{116}$$

$$\left(c - \frac{1}{4m^2c} U_\mu + \frac{1}{2mc} \tilde{\mathbf{L}}^{\dagger(n)} \tilde{\mathbf{S}}^{-1}\right) \mathbf{X}^{\dagger(n+1)} \mathbf{h}_\mu = \mathbf{T} \mathbf{h}_\mu. \tag{117}$$

C. Iterative solution of the quadratic equation for X

Let us now consider the quadratic equation Eq. (25) and solve this directly for **X**. We do this iteratively in a way inspired by the DKH approximation[4,7,13] (see Sec. IX A). To define this we multiply **V** (and consequently **U**) formally by the expansion parameter λ,

$$\mathbf{X} + (2mc)^{-1} \mathbf{X} \mathbf{S}^{-1} \mathbf{X} = c^{-1} \mathbf{T} - \lambda \{(2mc^2)^{-1} \mathbf{X} \mathbf{S}^{-1} \mathbf{V} - (4m^2c^2)^{-1} \mathbf{U} \mathbf{T}^{-1} \mathbf{X}\}. \tag{118}$$

The zeroth-order equation in λ for **X** them becomes

$$\mathbf{X}^{(0)} + (2mc)^{-1} \mathbf{X}^{(0)} \mathbf{S}^{-1} \mathbf{X}^{(0)} = c^{-1} \mathbf{T}. \tag{119}$$

Since $\mathbf{X}^{(0)}$ and $\mathbf{X}^{(0)} \mathbf{S}^{-1} \mathbf{X}^{(0)}$ have the same eigenvectors $\mathbf{d}_k$, these must also be eigenvectors of **T**,

$$\mathbf{T} \mathbf{d}_k = t_k \mathbf{S} \mathbf{d}_k, \tag{120}$$

$$\mathbf{X}^{(0)} \mathbf{d}_k = x_k \mathbf{S} \mathbf{d}_k, \tag{121}$$

$$\mathbf{X}^{(0)} \mathbf{S}^{-1} \mathbf{X}^{(0)} \mathbf{d}_k = x_k \mathbf{X}^{(0)} \mathbf{d}_k = (x_k)^2 \mathbf{S} \mathbf{d}_k, \tag{122}$$

$$x_k = -mc + \sqrt{m^2c^2 + 2mt_k}. \tag{123}$$

So, we can construct $\mathbf{X}^{(0)}$ from the eigenstates of **T**. In the next iterations one replaces **T** by

$$\mathbf{T}' = \mathbf{T} - (2mc)^{-1} \mathbf{X} \mathbf{S}^{-1} \mathbf{V} + (4m^2c)^{-1} \mathbf{U} \mathbf{T}^{-1} \mathbf{X} \tag{124}$$

and proceeds as before, being aware that $\mathbf{T}'$ is non-Hermitian and has complex eigenvalues.

The *other solution* of the quadratic equation that we discard actually leads to an **X** for *positronic* states. See Paper I Ref. 6 for the analog relation at the operator level.

A variant of this scheme is based on the decomposition of **X** into its Hermitian and anti-Hermitian parts $\mathbf{X}_H$ and $\mathbf{X}_A$. One gets from Eq. (25) a coupled system of equations for $\mathbf{X}_H$ and $\mathbf{X}_A$. An advantage of this approach is that the counterpart of $\mathbf{T}'$ [Eq. (124)] is either Hermitian or anti-Hermitian, so that the corresponding eigenstates are more easily obtained than for an arbitrary matrix.

D. Iteration with quadratic convergence

Equation (25), which we want to solve, is of the general type discussed in some detail in the Appendix.

$$\mathbf{F}(\mathbf{X}) = \mathbf{f}(\mathbf{X}) - \mathbf{X} = 0. \tag{125}$$

We start from a guess $\mathbf{X}^{(n)}$ for **X** and write

$$\mathbf{X}^{(n+1)} = \mathbf{X}^{(n)} + \mathbf{\Delta}. \tag{126}$$

Then, **Δ** must satisfy the linear matrix equation

$$\mathbf{F}(\mathbf{X}^{(n)}) + \left(\frac{1}{4m^2c^2} \mathbf{U} \mathbf{T}^{-1} - \frac{1}{2mc} \mathbf{X}^{(n)} \mathbf{S}^{-1}\right) \mathbf{\Delta} - \mathbf{\Delta} \frac{1}{2mc^2} \mathbf{S}^{-1} \mathbf{L}^{(n)} - \mathbf{\Delta} = 0. \tag{127}$$

We let this act on the eigenvectors $\mathbf{a}_\mu^{(n)}$ of $\mathbf{L}^{(n)}$ with metric **S** and eigenvalues $E_\mu^{(n)}$,

$$0 = \mathbf{F}(\mathbf{X}^{(n)}) \mathbf{a}_\mu^{(n)} + \left[\left(\frac{1}{4m^2c^2} \mathbf{U} \mathbf{T}^{-1} - \frac{1}{2mc} \mathbf{X}^{(n)} \mathbf{S}^{-1} - \frac{1}{2mc^2} E_\mu^{(n)}\right) - 1\right] \mathbf{\Delta} \mathbf{a}_\mu^{(n)}. \tag{128}$$

From this we get the vectors

$$\mathbf{d}_\mu = \mathbf{\Delta} \mathbf{a}_\mu. \tag{129}$$

This scheme is of *regula falsi* type and converges quadratically—if it converges. One must hence take care of having a reasonable starting iteration. This Newton-Raphson (NR) update can be combined with any of the iteration schemes just discussed.

E. Noniterative construction of X

There is finally a way to construct **X** and **L** *without any iteration*, namely, by first solving the double-dimension eigenvalue problem [Eq. (21)] and then combining the relations (17) for all states $\mathbf{a}_k, \mathbf{b}_k$ to the matrix equation

$$\mathbf{B} = \frac{1}{2m} \mathbf{T}^{-1} \mathbf{X} \mathbf{A}, \quad \mathbf{X} = 2m \mathbf{T} \mathbf{B} \mathbf{A}^{-1}. \tag{130}$$

The eigenvalue problem [Eq. (21)] has twice as many eigenstates as its quasirelativistic counterpart, and one must discard the eigenstates corresponding to positrons. Since we have expanded *D* in a basis that is especially appropriate for electronic states, positronic ones may be represented poorly, and there may occur some kind of *inverse variational collapse*; i.e., positronic states may get eigenvalues above the positronic continuum limit $-2mc^2$.

One may object to the just-discussed procedure that one wants to construct a quasirelativistic Hamiltonian *in order to avoid* the diagonalization of the Dirac operator, but there are situations where this construction makes sense, nevertheless, e.g., in a theory of many-electron systems in order to formulate the *external-field no-pair projection*.

This noniterative scheme has probably been first studied by Dyall,[9] but was not popularized and hardly noticed by the community of relativistic quantum chemistry. It has recently been rediscovered and implemented by Iliaš *et al.*[19]

VIII. NUMERICAL RESULTS

The results of our numerical calculations are summarized in Table I, where the numbers of iterations sufficient to obtain convergence in the sense of a rather severe criterion are collected.

In the first six columns the results for the *simple linear iteration* schemes of Sec. VII A are collected. These schemes only converge for small *Z* and small basis sets. The replace-

TABLE I. Number of iterations, for various schemes, necessary to converge the ten lowest electronic energies for atomic one-electron ions with charge Z to ten significant digits. The results with the Newton-Raphson update are in parentheses.

		Based on Eq. No.											
Z	basis	(93)	(99)	(96)	(100)	(98)	(101)	(102),(106)	(115),(106)	(113),(111)	(115),(116)	(119),(124)	(139)
10	10*s*5*p*	8	4	6	3	10	5	6	3	6	3	4	3
		(3)	(3)	(div)	(3)	(div)	(3)	(3)	(3)	(3)	(3)	(3)	(3)
	24*s*16*p*	div	div	div	3	div	6	div	3	div	3	5	3
		(div)	(3)	(div)	(3)	(div)	(div)	(10)	(3)	(10)	(3)	(3)	(3)
	32*s*29*p*	div	div	div	11	div	div	div	4	div	4	5	3
		(div)	(3)	(div)	(div)	(div)	(div)	(div)	(3)	(div)	(3)	(3)	(3)
20	10*s*5*p*	div	5	20	4	div	7	25	3	25	3	5	4
		(4)	(3)	(div)	(3)	(5)	(4)	(4)	(3)	(4)	(3)	(3)	(3)
	32*s*29*p*	div	div	div	11	div	div	div	4	div	4	5	6
		(div)	(3)	(div)	(div)	(div)	(div)	(div)	(3)	(div)	(3)	(3)	(3)
80	10*s*5*p*	div	div	div	10	div	div	div	7	div	7	13	12
		(div)	(6)	(div)	(4)	(div)	(div)	(6)	(4)	(6)	(4)	(4)	(4)
	32*s*29*p*	div	div	div	div	div	div	div	7	div	7	23	div
		(div)	(div)	(div)	(div)	(div)	(div)	(div)	(4)	(div)	(4)	(4)	(div)
	50*s*	div	div	div	div	div	div	div	div	div	div	div	div
		(div)	(div)	(div)	(div)	(div)	(div)	(div)	(4)	(div)	(4)	(4)	(div)

ment of $\mathbf{S}^{-1}\mathbf{L}$ by $\tilde{\mathbf{S}}^{-1}\tilde{\mathbf{L}}$ leads in some cases to an improvement, but even the Newton-Raphson update (Sec. VII D) is of little help. There is no advantage of the iteration schemes (98) and (101) inspired by the regular approximation over that in the spirit of direct perturbation theory [Eqs. (93) and (99)]. All these simple linear iteration schemes appear to be rather useless.

The improved linear iteration schemes of Sec. VII B are definitely better, especially in the variants in which $\mathbf{S}^{-1}\mathbf{L}$ is replaced by $\tilde{\mathbf{S}}^{-1}\tilde{\mathbf{L}}$. In these variants, combined with the NR update, *convergence in three or four iterations has been achieved in all studied cases*. These schemes had not yet been considered in our short Communication.[7] The scheme based on Eqs. (104) and (115) has some relation to the NESC method of Dyall,[9] but the other one is definitely new.

A very robust scheme, already studied in our short paper,[7] appears to be that of Sec. VII C, inspired by the DKH transformation, in which a set of quadratic equations is solved at every iteration.[7] This scheme converges in nearly all cases (one exception), even without a NR update, but with this update convergence never requires more than four iterations. This scheme is perfectly competitive with the other two just-discussed ones, provided that one uses the NR update. Without this, the convergence is a little slower.

We have also included an iteration scheme related to the work of Dyall and Filatov, which is based on Eq. (139) (see Sec. IX C). This scheme appears to converge fast for small Z, but for large Z it diverges, except for rather small basis sets. After the completion of the present work a paper by Filatov and Dyall, published online,[23] on this very iteration scheme and variants of it, came to our knowledge. We shall comment on it in Sec. IX C.

We have to modify one of the messages of our short paper.[7] There, we have concluded that a reliable convergence is only possible for the DKH inspired approach. Now, we know at least two completely different iteration schemes that perform similarly well.

We conclude from this numerical study that there is *no single way* towards a fast iterative construction of a quasirelativistic Hamiltonian. One way (Sec. VII C) is via the free-particle FW transformation at the matrix level, like in the DKH transformation, but with a much faster convergence than in the traditional DKH transformation. Another one, based on Eq. (103) is related to the elimination of the small component. A third one based on the diagonalization of matrix $\mathbf{U}$ [Eq. (109)] performs similarly well. For the two latter schemes use of the identity (38), probably first stressed by Filatov, is essential. A Newton-Raphson update *is recommended* in all cases. In appreciating the excellent performance of improved iterative schemes, one should not forget that even the noniterative construction of the quasirelativistic Hamiltonian at the matrix level (Sec. VII E) is hardly more time consuming than any of the iteration schemes.

IX. RELATION TO KNOWN APPROACHES

A. The Douglas-Kroll-Hess (DKH) method

In the Douglas-Kroll-Hess (DKH) method one searches for a unitary matrix $\mathbf{W}$ that transforms the matrix representation $\mathbf{D}$ of the Dirac operator (see Sec. VI) in such a way that in a first step the matrix representation $\mathbf{D}_0$ of the Dirac operator for a *free particle* is transformed to a block diagonal form. This is achieved by the matrix representation $\mathbf{W}_0$ of the *free-particle Foldy-Wouthuysen transformation*, which is easily performed. It requires only the diagonalization [Eq. (120)] of the matrix $\mathbf{T}$ and the solution of a quadratic equation for each eigenvalue. Transforming $\mathbf{D}$ by $\mathbf{W}_0$, an off-diagonal block remains it can be removed by a sequence of further transformations. This sequence corresponds to an ex-

pansion in powers of the coupling strength (i.e., in powers of a formal perturbation parameter multiplied to the external potential V).

A compact formulation of the leading order of the DKH transformation can be found in the following way. Let us multiply Eq. (119) from both sides by $\mathbf{S}^{-1/2}$ and let us define

$$\mathbf{t} = \mathbf{S}^{-1/2}\mathbf{T}\mathbf{S}^{-1/2}, \quad \mathbf{x} = \mathbf{S}^{-1/2}\mathbf{X}_0\mathbf{S}^{-1/2},$$

$$\mathbf{v} = \mathbf{S}^{-1/2}\mathbf{V}\mathbf{S}^{-1/2}, \quad \mathbf{u} = \mathbf{S}^{-1/2}\mathbf{U}\mathbf{S}^{-1/2}, \tag{131}$$

$$\mathbf{l} = \mathbf{S}^{-1/2}\mathbf{L}_0\mathbf{S}^{-1/2}, \quad \tilde{\mathbf{l}} = \mathbf{S}^{-1/2}\tilde{\mathbf{L}}_0\mathbf{S}^{-1/2}, \quad \tilde{\mathbf{s}} = \mathbf{S}^{-1/2}\tilde{\mathbf{S}}_0\mathbf{S}^{-1/2}.$$

Then, Eq (119) becomes

$$\mathbf{x}^2 + 2mc\mathbf{x} - 2m\mathbf{t} = 0, \tag{132}$$

with $\mathbf{x}$ and $\mathbf{t}$ commuting, such that

$$\mathbf{x} = mc\left(\pm\sqrt{1 + \frac{2\mathbf{t}}{mc^2}} - 1\right), \tag{133}$$

$$\mathbf{l} = \mathbf{v} + c\mathbf{x}, \tag{134}$$

$$\tilde{\mathbf{s}} = 1 + \frac{1}{2m}\mathbf{x}^2\mathbf{t}^{-1}, \tag{135}$$

$$\tilde{\mathbf{l}} = c\tilde{\mathbf{s}}\mathbf{x} + \mathbf{v} + \mathbf{x}\mathbf{t}^{-1}\mathbf{u}\mathbf{t}^{-1}\mathbf{x}, \tag{136}$$

$$\mathbf{l}^+ = c\mathbf{x} + \tilde{\mathbf{s}}^{-1/2}\mathbf{v}\tilde{\mathbf{s}}^{-1/2} + \tilde{\mathbf{s}}^{-1/2}\mathbf{x}\mathbf{t}^{-1}\mathbf{u}\mathbf{t}^{-1}\mathbf{x}\tilde{\mathbf{s}}^{-1/2}. \tag{137}$$

While the leading order of the DKH transformation[13] is quite straightforward, higher orders become increasingly complicated.[14] Although the expansion in powers of the coupling strength *appears to converge in all practical cases*, this convergence is rather slow. The *leading-order* DKH transformation has been the basis of one of the most popular *approximative* quasirelativistic theories,[13] while the generalization to higher orders has been considered only relatively recently,[14] and is hardly competitive with the method outlined in this paper.

The essential difference between the DKH scheme and our method is that we formulate the theory in terms of matrix $\mathbf{X}$, which satisfies much simpler equations than $\mathbf{W}$, and that we do not care to construct $\mathbf{W}$ from $\mathbf{X}$ before the iteration scheme has converged. We further do *not necessarily* use the free-particle Foldy-Wouthuysen transformation as the first step, although in one of our possible options this is actually done. We can further accelerate the convergence by a NR update.

So, our scheme is significantly simpler and more efficient than that of DKH, and there is *no increase of complexity* with the increase of order. In Table II and Fig. 1 the performance of our iterative scheme, in the variant inspired by DKH, is compared with that of the genuine DKH expansion. For the electronic ground state of Hg^{79+} in a 50s basis, the DKH method leads at 14th order to a result with seven figure accuracy. Our method, without NR update, leads to the same accuracy after six iterations, while with NR update we obtain a nine-figure accuracy after only three iterations and an eight-figure accuracy after two iterations.

TABLE II. The electronic ground state of Hg^{79+} obtained with 50 Gaussian s functions (Ref. 14, Ch. van Wuellen) DKH-type iteration scheme based on Eqs. (116) and (119), in comparison with various orders of the Douglas-Kroll-Hess (DKH) transformation. NRU: Newton-Raphson update.

Iteration	Without NRU[a]	With NRU	DKH[b]	DKH[c]
1	−3686.448 75	−3686.448 75	−3686.448 75	−3686.447 46
2	−3532.931 72	−3532.192 20	−3532.324 89	−3523.324 90
3	−3532.325 18	−3532.192 14	−3533.119 62	−3533.119 58
4	−3532.204 98	−3532.192 14	−3531.708 61	−3531.708 57
5	−3532.194 98	−3532.192 14	−3532.461 52	−3532.462 50
6	−3532.192 79		−3532.101 37	−3532.101 20
7	−3532.192 34			−3532.225 33
8	−3532.192 22			−3532.177 00
9	−3532.192 19			−3532.200 16
10	−3532.192 18			−3532.188 44
11	−3532.192 17			−3532.193 69
12	−3532.192 17			−3532.191 26
13	−3532.192 19			−3532.192 49
14	−3532.192 17			−3532.191 84
15	−3532.192 67			
16	−3532.192 17			
17	−3532.192 35			
DEQ	−3532.192 13[d]	−3532.192 13[d]	−3532.192 13[d]	−3532.192 15[e]

[a]Oscillation starts from the 15th iteration due to numerical instability.
[b]Reference 14 (Ch. van Wuellen).
[c]Reference 14 (Reiher and Wolf). 75 Gaussian s functions were used therein.
[d]Four-component Dirac energy value with the same basis set from Ref. 14 (Ch. van Wuellen).
[e]Analytic Dirac energy value.

B. The method of Barysz, Sadlej, and Snijders (BSS)

The method of Barysz, Sadlej, and Snijders (BSS),[15] in its more recent formulation also known as the "infinite-order two-component" method, is closely related to DKH. Like in DKH the first step is the transformation of $\mathbf{D}$ with the free-particle Foldy-Wouthuysen transformation $\mathbf{W}_0$ at the matrix level. The remaining transformation $\mathbf{W}'$, such that $\mathbf{W} = \mathbf{W}_0\mathbf{S}^{-1}\mathbf{W}'$, is, however, not formulated as a sequence of unitary transformations, but as a single transformation, expressed in terms of matrix $\mathbf{R}$, from which $\mathbf{W}'$ is constructed in an iterative way, like in our scheme where the full transformation $\mathbf{W}$ is constructed from $\mathbf{X}$. So the BSS method is somewhere in between DKH and our scheme. Like the DKH method it appears to converge in all cases of interest, though the convergence is often rather slow. A NR update should also be possible within the BSS scheme, but has not yet been tried. An advantage with respect to high-order DKH is that (like in our scheme) there is no increased complexity with increasing order. However, as already mentioned, the way via the free-particle Foldy-Wouthuysen transformation, common to DKH and BSS, is only *one of several options* in our approach.

C. The method of Dyall and Filatov

When we refer to the method of Dyall and Filatov, we mean a method that is *implicit* in a paper by Dyall,[9] and that has later been analyzed by Filatov[8] and interpreted as a precursor to our method.

The relation between this method and ours has recently

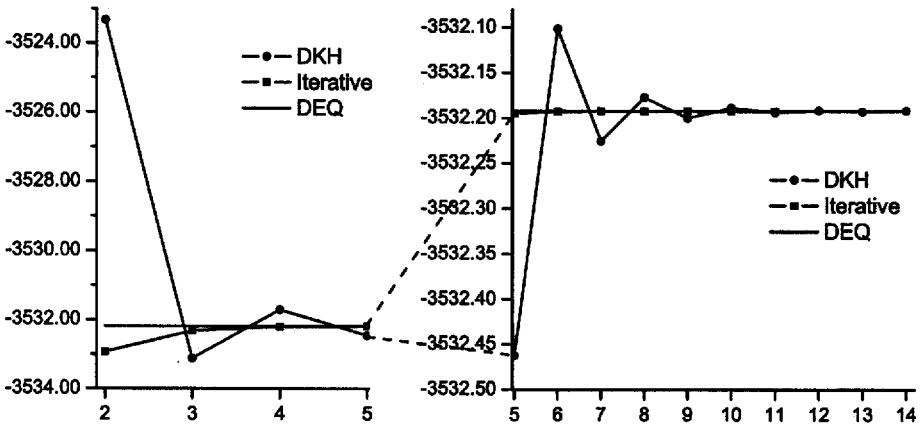

FIG. 1. The ground state energy of Hg^{79+} from the iteratively constructed quasirelativistic Hamiltonian and for different orders of the DKH Hamiltonian. DEQ: Four-component Dirac energy.

been discussed,[8,10] so we need not go into details. Translated to the nomenclature used in the present paper, Dyall[9] has realized that there is an operator $\mathbf{X}$, which relates the vectors $\mathbf{a}$ and $\mathbf{b}$ in the sense of Eq. (17), and has implicitly derived that Eqs. (27) and (33) must be simultaneously satisfied for all $\mathbf{a}_k$, with $\mathbf{L}$ defined by Eq. (26) and $\tilde{\mathbf{L}}$ by Eq. (32). One can actually use Eqs. (27) and (33) as a condition for the construction of $\mathbf{L}$,[10] although this has not explicitly been suggested by Dyall. The two equations can be satisfied simultaneously only if Eq. (38) holds. From

$$\mathbf{L} = c\mathbf{X} + \mathbf{V} = \mathbf{S}(\tilde{\mathbf{S}})^{-1}\tilde{\mathbf{L}}, \tag{138}$$

we get an iteration scheme for $\mathbf{X}$.

$$\mathbf{X} = c^{-1}[\mathbf{S}(\tilde{\mathbf{S}})^{-1}\tilde{\mathbf{L}} - \mathbf{V}]. \tag{139}$$

Although Dyall's paper[9] is, in some sense, a precursor to our work,[7] the method discussed in this section has been formulated explicitly only after our paper has been published. Very recently, numerical calculations with this method have been published,[23] with two interesting messages:

(1) As it stands, the iteration scheme based on Eqs. (139) and (33) fails to converge for large Z (typically, $Z=92$) and realistic basis sets, which include steep basis functions.

(2) Convergence can be enforced for a *subset of all bound eigenstates* if one does not iterate Eq. (139), but rather

$$\mathbf{X} = c^{-1}[\mathbf{S}(\tilde{\mathbf{S}})^{-1}\tilde{\tilde{\mathbf{L}}} - \mathbf{V}], \tag{140}$$

where $\tilde{\tilde{\mathbf{L}}}$ is an operator with the same eigenvectors $\mathbf{a}_k$ as $\tilde{\mathbf{L}}$, but with the same eigenvalues E_k *only for* E_k *below an imposed threshold* ϵ^{up}, while the E_k above the threshold are replaced either by zero or by $(1+E_k/\epsilon^{\mathrm{up}})^{-1}$.

With this *projection to a subset of the bound eigenstates*, convergence for these subsets can be achieved within three or four iterations. Such a projection is *not in the spirit* of the methods that we have studied and advocated here, where convergence for *all positive-energy eigenstates* of the matrix representation of the Dirac operator is achieved.

The method advocated in Ref. 23 does not give any information on the states outside the projected subset, and it cannot be used for the construction of the quasirelativistic Hamiltonian $\tilde{\mathbf{L}}$ or $\tilde{\mathbf{L}}$ for all electronic states.

D. Stationary direct perturbation theory

In Paper I of this series[6] we have included a discussion of DPT (Ref. 16) and its quasidegenerate version.[5,24] The matrix counterpart of this theory at the operator level is S-DPT.[2] One might therefore expect some comments on S-DPT in the context of quasirelativistic Hamiltonians at the matrix level. We have, however, decided to postpone this to a separate paper. The reason is that the relations are less close than one might have guessed.

In order to formulate S-DPT, one must first choose a particular state, or a set of states, which define a *model space*, before one starts to solve the eigenvalue equation for the matrix representation of the Dirac operator in an expansion in the *natural* perturbation parameter c^{-1}. This is not possible via an expansion of the matrices $\mathbf{X}$, $\mathbf{L}$, and $\tilde{\mathbf{L}}$ in powers of c^{-1} because these matrices do not allow such an expansion. This expansion converges only if it is limited to a subset of eigenstates of $\mathbf{D}$. In quasidegenerate S-DPT one constructs an effective Hamiltonian matrix $\mathbf{K}$ that acts in the model space, which is a subspace of that of all electronic states (representable in the given basis), while matrix $\mathbf{L}$, which plays a central role in the present paper, acts in the *entire space* of electronic states. This difference requires rather different strategies.

E. "The regular approximation" at the matrix level

Before the construction of a quasirelativistic Hamiltonian matrix equivalent to the matrix representation of the Dirac operator was known, quasirelativistic theory was dominated by two competing approximative quasirelativistic Hamiltonians. One of these was the leading order of the DKH transformation, already mentioned in Sec. IX A, the other the so-called regular approximation (RA), specifically the zeroth-order regular approximation (ZORA).[17] While the implementation of DKH has been done in a matrix representation from the very beginning, the original ZORA has been a theory in configuration space, i.e., at the operator level.

The motivation for ZORA has been an analysis of the derivation of the c^{-1} expansion from the elimination of the small component. The expression

$$\chi = c(V - 2mc^2 - E)^{-1}\varphi, \quad V = -Z/r \tag{141}$$

is analytic in c^{-1} only for $r > Z/2mc^2$, while the expansion in powers of $(-Z/r - 2mc^2)^{-1}$ does not have this problem. However, this expansion does not correspond to a physically meaningful expansion parameter, and it also leads to complicated integrals. Therefore, only the leading approximation, ZORA, has played a significant role. It is simply based on replacing Eq. (141) by

$$\chi_{\mathrm{ZORA}} = -c(V - 2mc^2)^{-1}\boldsymbol{\sigma}\mathbf{p}\varphi_{\mathrm{ZORA}}. \tag{142}$$

Since we have neglected $|E|$ with respect to $2mc^2$, it is obvious that this can be a decent approximation only for states for which $|E| \ll 2mc^2$, but certainly not for the entire spectrum.

One can also apply the same strategy in a theory starting from the matrix representation of the Dirac operator. Then, instead of Eq. (142), one has

$$\begin{aligned}\mathbf{b}_{\mathrm{ZORA}} &= \frac{1}{2m}\mathbf{T}^{-1}\mathbf{X}_{\mathrm{ZORA}}\mathbf{a}_{\mathrm{ZORA}} \\ &= 2mc\left(\mathbf{T} - \frac{1}{4m^2c^2}\mathbf{U}\right)^{-1}\mathbf{T}\mathbf{a}_{\mathrm{ZORA}},\end{aligned} \tag{143}$$

which corresponds to

$$\mathbf{X}_{\mathrm{ZORA}} = c^{-1}\left(\mathbf{T}^{-1} - \frac{1}{4m^2c^2}\mathbf{T}^{-1}\mathbf{U}\mathbf{T}^{-1}\right)^{-1}, \tag{144}$$

$$\mathbf{L}_{\mathrm{ZORA}} = \mathbf{V} + c\mathbf{X}_{\mathrm{ZORA}} = \mathbf{V} + \left(\mathbf{T}^{-1} - \frac{1}{4m^2c^2}\mathbf{T}^{-1}\mathbf{U}\mathbf{T}^{-1}\right)^{-1}. \tag{145}$$

This is the first iteration of Eq. (98). The approximate matrix $\mathbf{L}_{\mathrm{ZORA}}$ happens to be Hermitian, although its exact counterpart $\mathbf{L}$ is usually non-Hermitian.

Of course, at the matrix level, the motivation for the introduction of the RA is lost to a large extent. The singularity of V at $r=0$ does not matter in a matrix representation. This is probably why the matrix RA has been considered systematically only rather recently.[18]

An improved energy is obtained if we use the approximate $\mathbf{X}_{\mathrm{ZORA}}$ to construct $\tilde{\mathbf{L}}_{\mathrm{IORA}}$ as defined by Eq. (32) and to solve the eigenvalue [Eq. (33)]. For historical reasons this approach is known as (the matrix version of) the infinite-order regular approximation (IORA).[21] The name is misleading because at variance with what it suggests, it is *not exact* (unlike, e.g., infinite-order DKH).

It is somewhat curious that $\tilde{\mathbf{L}}_{\mathrm{IORA}}$ is identical with $\mathbf{L}_{\mathrm{ZORA}}$,

$$\begin{aligned}\tilde{\mathbf{L}}_{\mathrm{IORA}} &= \mathbf{V} + c\mathbf{X}_{\mathrm{ZORA}} + c\mathbf{X}_{\mathrm{ZORA}}^{\dagger} - c^2\mathbf{X}_{\mathrm{ZORA}}^{\dagger}\left(\mathbf{T}^{-1} - \frac{1}{4m^2c^2}\mathbf{T}^{-1}\mathbf{U}\mathbf{T}^{-1}\right)\mathbf{X}_{\mathrm{ZORA}} \\ &= \mathbf{V} + c\mathbf{X}_{\mathrm{ZORA}} + c\mathbf{X}_{\mathrm{ZORA}}^{\dagger} - c^2\mathbf{X}_{\mathrm{ZORA}}^{\dagger}(c^{-1}\mathbf{X}_{\mathrm{ZORA}}^{-1})\mathbf{X}_{\mathrm{ZORA}} \\ &= \mathbf{V} + c\mathbf{X}_{\mathrm{ZORA}} = \mathbf{L}_{\mathrm{ZORA}}.\end{aligned} \tag{146}$$

The difference between ZORA and IORA in a matrix representation is only in the metric which is $\mathbf{S}$ in ZORA and $\tilde{\mathbf{S}}_{\mathrm{ZORA}}$ in IORA.

$$\begin{aligned}\tilde{\mathbf{S}}_{\mathrm{IORA}} &= \mathbf{S} + \frac{1}{2m}\mathbf{X}_{\mathrm{ZORA}}^{\dagger}\mathbf{T}^{-1}\mathbf{X}_{\mathrm{ZORA}} \\ &= \mathbf{S} + \frac{1}{2mc^2}\mathbf{T}\left(\mathbf{T} - \frac{1}{2mc^2}\mathbf{U} + \frac{1}{16m^2c^4}\mathbf{U}\mathbf{T}^{-1}\mathbf{U}\right)^{-1}\mathbf{T}.\end{aligned} \tag{147}$$

ZORA is the first approximation to the UESC, while IORA is the first approximation to the NESC of Dyall.[9] One drawback of the regular approximation in general is that the results depend on the gauge of the scalar potential, i.e., on the origin of the energy scale. As to possibilities of how to avoid this and/or more details, see the papers by Filatov and Cremer.

X. CONCLUSIONS

The problem of constructing a quasirelativistic Hamiltonian that is equivalent (for electronic states) to the original Dirac Hamiltonian can be regarded as solved. This construction is possible in an iterative way, starting from the matrix representation of the Dirac operator in a kinetically balanced basis, but cannot be achieved at the operator level. The problems that arise at the operator level have been discussed in Paper I of this series.[6]

The transformation from the matrix representation of the Dirac operator to a quasirelativistic Hamiltonian matrix involves a matrix $\mathbf{X}$ that is constructed iteratively. There are essentially three different and more or less equivalent strategies for the construction of $\mathbf{X}$. One of these, already presented in our preliminary Communication,[7] tries the iterative solution of a quadratic matrix equation for $\mathbf{X}$ and passes via the transformation of the free-particle Dirac operator to the quasirelativistic form, and involves the solution of a set of quadratic equations in every iteration. We call this "DKH-type" because it is an essential ingredient of the DKH transformation to start with a free-particle Foldy-Wouthuysen transformation a matrix level.

The two other efficient schemes, first presented in this paper, are based on the iterative solution of a coupled set of two pseudolinear matrix equations in terms of $\mathbf{X}$ and the quasirelativistic Hamiltonian $\mathbf{L}$, and are characterized by two ingredients. One is that the pseudolinear equation for $\mathbf{X}$ is solved noniteratively and that iterations are only necessary because there are two coupled matrix equations and $\mathbf{L}$ has to be updated. The other ingredient is that the iteration is formulated in terms of the Hermitian quasirelativistic Hamil-

tonian $\tilde{\mathbf{L}}$ rather than the non-Hermitian $\mathbf{L}$, which appears originally. Among the two schemes the one formulated in terms of the eigenvectors of the *perturbation matrix* $\mathbf{U}$ appears preferable over that in terms of the eigenstates of $\tilde{\mathbf{L}}$ because $\mathbf{U}$ does not change during the iterations. These two new schemes converge even faster than the DKH-type scheme, if no Newton-Raphson update is added. With such an update, which is highly recommended, all three schemes converge (for all eigenstates) within three or four iterations, even for lage Z and "difficult" basis sets (involving very steep basis functions).

It is definite that there is no obligation to pass via the quasirelativistic Hamiltonian for free particles (as has sometimes been claimed in the DKH or BSS communities). Iteration schemes inspired by DPT or the ZORA converge only in exceptionally well-behaved situations or if special tricks are applied.

Although we have presented very effective iterative schemes for the construction of $\mathbf{X}$ and $\tilde{\mathbf{L}}$, we should not forget that the noniterative construction via diagonalization of the matrix representation of the Dirac operator (see Sec. VII E) is hardly more demanding from a computational point of view compared to the iterative schemes favored here.

In this paper we have not discussed the matrix representation of direct perturbation theory, i.e., S-DPT. The reason is that the full quasirelativistic Hamiltonian matrix is not expandable in powers of c^{-1}. One must rather consider a limited subset of the eigenstates, for which an expansion in powers of c^{-1} is possible; i.e., one has to consider a quasirelativistic theory in terms of an *effective Hamiltonian*, rather than for the full quasirelativistic Hamiltonian.

The theory outlined in this paper and in the preceding paper[6] is a theory at the one-electron level; i.e., we have not discussed the role of the electron interaction. Closer to reality is, e. g., the relativistic Hartree-Fock theory, in which the bare-nuclear potential V is replaced by an effective potential that contains the Coulomb and exchange interactions with the other electrons. This aspect has been discussed in a recent paper.[25] In this paper the various Hermitian Hamiltonians $\bar{\mathbf{L}}$ and $\tilde{\mathbf{L}}$ with a nonunit metric and their Foldy-Wouthuysen analogs $\mathbf{L}^+$ and $\tilde{\mathbf{L}}^+$ were carefully examined. It is striking that with the *atomic approximation* for the molecular $\mathbf{X}$ matrix and a simple ansatz for correcting the two-electron picture change error, $\mathbf{L}^+$ requires only nonrelativistic-type molecular integrals and is thus an extremely efficient and accurate infinite-order quasirelativistic approach, making all existing quasirelativistic theories obsolete, both those of finite and of infinite order.

The following final comment is in order. The theory outlined here would be rather meaningless if we could not rely on the completeness of the kinetically balanced basis in the sense that by increasing the dimension of the basis we can reproduce the exact eigenvalues of the Dirac operator to any desired accuracy. This is supported by a lot of empirical evidence, but a proof has only been given quite recently.[26]

ACKNOWLEDGMENTS

This paper is dedicated to Volker Staemmler at the occasion of his 65th birthday. The work reported here has been supported by Deutsche Forschungsgemeinschaft (DFG) and Fonds der Chemie (FCh), as well as by the National Science Foundation of China (Project Nos. 20333020, 20573003, and 20625311) and by MOST of China (Project No. 2006CB601103). The authors acknowledge an exchange of correspondence with M. Filatov and K. Dyall, which culminated with the publication of Filatov's Comment (Ref. 8) on our short communication and our response to it (Ref. 10).

APPENDIX: SOLUTION OF LINEAR AND QUADRATIC MATRIX EQUATIONS

For the sake of simplicity we assume in this Appendix that all relevant matrices are normal, i.e., can be diagonalized by means of a unitary transformation. If this is not the case, some modifications are necessary; e.g., independent left and right eigenvectors must be considered, or one must only choose approaches that do not involve a diagonalization of matrices.

We want to solve

$$0 = \mathbf{F}(\mathbf{X}) = \mathbf{f}(\mathbf{X}) - \mathbf{X} = \mathbf{A} + \mathbf{B}\mathbf{X} + \mathbf{X}\mathbf{C} - \mathbf{X}. \tag{A1}$$

This is trivial if $\mathbf{B}=0$ or $\mathbf{C}=0$. Then,

$$\mathbf{X} = (1-\mathbf{B})^{-1}\mathbf{A} \quad \text{for } \mathbf{C}=0, \tag{A2}$$

$$\mathbf{X} = \mathbf{A}(1-\mathbf{C})^{-1} \quad \text{for } \mathbf{B}=0. \tag{A3}$$

Also, the case $\mathbf{C}=-\mathbf{B}$ or

$$0 = \mathbf{F}(\mathbf{X}) = \mathbf{A} + [\mathbf{B},\mathbf{X}] - \mathbf{X} \tag{A4}$$

is rather simple. In the basis, in which $\mathbf{B}$ is diagonal, we get

$$X_{\mu,\mu} = A_{\mu,\mu}, \tag{A5}$$

$$X_{\mu,\nu} = A_{\mu,\nu} + (B_{\mu,\mu} - B_{\nu,\nu})X_{\mu,\nu}, \quad \mu \neq \nu. \tag{A6}$$

In the case where $\mathbf{B}\neq 0$ and $\mathbf{C}\neq 0$ we can either choose the basis in which $\mathbf{B}$ is diagonal or that in which $\mathbf{C}$ is diagonal. Let us make the second choice and assume that

$$\mathbf{C}\mathbf{c}_\mu = \gamma_\mu \mathbf{c}_\mu. \tag{A7}$$

Let us define

$$\mathbf{b}_\mu = \mathbf{X}\mathbf{c}_\mu \tag{A8}$$

and consider

$$0 = \mathbf{F}(\mathbf{X})\mathbf{c}_\mu = (\mathbf{A} + [\mathbf{B} + \gamma_\mu - 1]\mathbf{X})\mathbf{c}_\mu, \tag{A9}$$

$$0 = [\mathbf{B} + \gamma_\mu - 1]^{-1}\mathbf{A}\mathbf{c}_\mu + \mathbf{b}_\mu. \tag{A10}$$

So, for each μ we can construct $\mathbf{b}_\mu$. Combining $\mathbf{c}_\mu$ and $\mathbf{b}_\mu$ to matrices $\mathbf{c}$ and $\mathbf{b}$, respectively, we can combine Eq. (A8) for all μ to a matrix equation

$$\mathbf{b} = \mathbf{X}\mathbf{c}, \quad \mathbf{X} = \mathbf{b}\mathbf{c}^{-1}, \tag{A11}$$

which allows us to obtain $\mathbf{X}$ in a somewhat indirect but noniterative way. It may sometimes be simpler to construct $\mathbf{X}$ iteratively by the following procedure:

$$\mathbf{X}^{(n+1)} = \mathbf{f}(\mathbf{X}^{(n)}). \tag{A12}$$

This requires that **B** and **C** are sufficiently small. The convergence is then linear; i.e., it goes like a geometrical series.

For a quadratic matrix equation a noniterative solution is usually not possible, but a Newton-Raphson (or rather *regula falsi*)-type approach is recommended.

We want to solve

$$0 = \mathbf{F}(\mathbf{X}) = \mathbf{f}(\mathbf{X}) - \mathbf{X} = \mathbf{A} + \mathbf{B}\mathbf{X} + \mathbf{X}\mathbf{C} + \mathbf{X}\mathbf{D}\mathbf{X} - \mathbf{X} \tag{A13}$$

with now a different meaning of $\mathbf{F}(\mathbf{X})$ and $\mathbf{f}(\mathbf{X})$. In a naive iteration scheme we construct

$$\mathbf{X}^{(n+1)} = \mathbf{f}(\mathbf{X}^{(n)}). \tag{A14}$$

In a NR-type approach we start from $\mathbf{F}^{(n)} = \mathbf{F}(\mathbf{X}^{(n)})$ and we want to solve

$$0 = \mathbf{F}(\mathbf{X}^{(n+1)}) = \mathbf{F}(\mathbf{X}^{(n)} + \mathbf{\Delta}) = \mathbf{F}^{(n)} + \mathbf{G}(\mathbf{\Delta}, \mathbf{X}^{(n)}), \tag{A15}$$

$$\mathbf{G}(\mathbf{\Delta}, \mathbf{X}) = (\mathbf{B} + \mathbf{X}\mathbf{D})\mathbf{\Delta} + \mathbf{\Delta}(\mathbf{C} + \mathbf{D}\mathbf{X}) - \mathbf{\Delta} \tag{A16}$$

for $\mathbf{\Delta}$. In **G** we ignore terms quadratic and of higher order in $\mathbf{\Delta}$, and so obtain a linear system of equations for $\mathbf{\Delta}$, to be solved as indicated above.

In the case in which we are interested, we have

$$\mathbf{F}(\mathbf{X}) = c^{-1}\mathbf{T} + \frac{1}{4m^2c^2}\mathbf{U}\mathbf{T}^{-1}\mathbf{X} - \frac{1}{2mc^2}\mathbf{X}\mathbf{S}^{-1}\mathbf{V} - \frac{1}{2mc}\mathbf{X}\mathbf{S}^{-1}\mathbf{X} - \mathbf{X}, \tag{A17}$$

$$\begin{aligned}\mathbf{G}(\mathbf{\Delta}, \mathbf{X}) &= \left(\frac{1}{4m^2c^2}\mathbf{U}\mathbf{T}^{-1} - \frac{1}{2mc}\mathbf{X}\mathbf{S}^{-1}\right)\mathbf{\Delta} - \mathbf{\Delta}\left(\frac{1}{2mc^2}\mathbf{S}^{-1}\mathbf{V} + \frac{1}{2mc}\mathbf{S}^{-1}\mathbf{X}\right) - \mathbf{\Delta} \\ &= \left(\frac{1}{4m^2c^2}\mathbf{U}\mathbf{T}^{-1} - \frac{1}{2mc}\mathbf{X}\mathbf{S}^{-1}\right)\mathbf{\Delta} - \mathbf{\Delta}\frac{1}{2mc^2}\mathbf{S}^{-1}\mathbf{L} - \mathbf{\Delta}.\end{aligned} \tag{A18}$$

We let this act on the eigenvectors $\mathbf{a}_\mu$ of **L** with metric **S** and eigenvalues E_μ,

$$0 = \mathbf{F}_n\mathbf{a}_\mu + \left(\frac{1}{4m^2c^2}\mathbf{U}\mathbf{T}^{-1} - \frac{1}{2mc}\mathbf{X}\mathbf{S}^{-1} + \frac{1}{2mc^2}E_\mu - 1\right)\mathbf{\Delta}\mathbf{a}_\mu. \tag{A19}$$

From this we get the vectors

$$\mathbf{d}_\mu = \mathbf{\Delta}\mathbf{a}_\mu. \tag{A20}$$

and from $\mathbf{a}_\mu$ and $\mathbf{d}_\mu$, finally, $\mathbf{\Delta}$, which we use to update **X**.

[1] J. L. Heully, I. Lindgren, E. Lindroth, S. Lundquist, and A. M. Mårtensen-Pendrill, J. Phys. B **19**, 2799 (1986).
[2] W. Kutzelnigg, Phys. Rev. A **54**, 1183 (1996).
[3] W. Kutzelnigg, Int. J. Quantum Chem. **25**, 107 (1984).
[4] W. Kutzelnigg, Chem. Phys. **225**, 203 (1997).
[5] W. Kutzelnigg, J. Chem. Phys. **110**, 8283 (1999).
[6] W. Kutzelnigg and W. Liu, Mol. Phys. **104**, 2225 (2006).
[7] W. Kutzelnigg and W. Liu, J. Chem. Phys. **123**, 241102 (2005).
[8] M. Filatov, J. Chem. Phys. **125**, 107101 (2006).
[9] K. G. Dyall, J. Chem. Phys. **106**, 9618 (1997).
[10] W. Kutzelnigg and W. Liu, J. Chem. Phys. **125**, 107102 (2006).
[11] W. H. E. Schwarz and H. Wallmeyer, Mol. Phys. **46**, 1045 (1982); W. H. E. Schwarz and E. Wechsel-Trakowski, Chem. Phys. Lett. **85**, 94 (1982).
[12] Y. S. Lee and A. P. McLean, J. Chem. Phys. **76**, 735 (1982); R. E. Stanton and S. J. Havrilak, *ibid.* **81**, 1910 (1984); P. J. C. Aerts and W. C. Nieuwpoort, Int. J. Quantum Chem., Quantum Chem. Symp. **19**, 267 (1986).
[13] B. A. Hess, Phys. Rev. A **33**, 3742 (1986); G. Jansen and B. A. Hess, *ibid.* **39**, 6061 (1989); R. Samzow, B. A. Hess, and G. Jansen, J. Chem. Phys. **96**, 1227 (1992); A. Wolf, M. Reiher, and B. Hess, in *Relativistic Electronic Structure Theory, Part I: Fundamentals*, Theoretical and Computational Chemistry Vol. 11, edited by P. Schwerdtfeger (Elsevier, Amsterdam, 2002).
[14] T. Nakajima and K. Hirao, J. Chem. Phys. **113**, 7786 (2000); M. Reiher and A. Wolf, *ibid.* **121**, 10945 (2004); Ch. van Wuellen, *ibid.* **120**, 7307 (2003).
[15] M. Barysz, A. J. Sadlej, and J. G. Snijders, Int. J. Quantum Chem. **65**, 225 (1997); M. Barysz, J. Chem. Phys. **114**, 9315 (2001); M. Barysz and A. J. Sadlej, *ibid.* **116**, 2696 (2002); D. Kedziera and M. Barysz, *ibid.* **121**, 6719 (2002); D. Kedziera and M. Barysz, Chem. Phys. Lett. **3093**, 521 (2004); D. Kedziera, J. Chem. Phys. **123**, 074109 (2005).
[16] A. Rutkowski, J. Phys. B **19**, 149 (1986); **19**, 3431 (1986); **19**, 3443 (1986); K. Jankowski and A. Rutkowski, Phys. Scr. **36**, 464 (1987); A. Rutkowski, K. Jankowski, and B. Mikielewicz, J. Phys. B **21**, L147 (1988); W. Kutzelnigg, Z. Phys. D: At., Mol. Clusters **11**, 15 (1989); in *Relativistic Electronic Structure Theory, Part I: Fundamentals*, Theoretical and Computational Chemistry Vol. 11 (Elsevier, Amsterdam, 2002).
[17] Ch. Chang, M. Pelissier, and M. Durand, Phys. Scr. **34**, 394 (1986); E. van Lenthe, E. J. Baerends, and J. G. Snijders, J. Chem. Phys. **99**, 4597 (1993).
[18] M. Filatov and D. Cremer, J. Chem. Phys. **119**, 11526 (2003); **122**, 044104 (2005); **122**, 064104 (2005).
[19] M. Iliaš and H. J. Aa. Jensen (unpublished); M. Iliaš, H. J. Aa. Jensen, V. Kellö, B. O. Roos, and M. Urban, Chem. Phys. Lett. **408**, 210 (2005); M. Iliaš and T. Saue, J. Chem. Phys. **126**, 064102 (2007).
[20] K. G. Dyall, J. Chem. Phys. **100**, 2118 (1994).
[21] K. G. Dyall and E. van Lenthe, J. Chem. Phys. **111**, 1366 (1999).
[22] E. van Lenthe, E. J. Baerends, and J. G. Snijders, J. Chem. Phys. **106**, 2373 (1996).
[23] M. Filatov and K. Dyall, Theor. Chem. Acc. **117**, 333 (2007).
[24] A. Rutkowski and W. H. E. Schwarz, J. Chem. Phys. **104**, 8546 (1996).
[25] W. Liu and D. Peng, J. Chem. Phys. **125**, 044102 (2006); **125**, 149901 (2006).
[26] W. Kutzelnigg, J. Chem. Phys. (submitted).