Exact decoupling of the Dirac Hamiltonian. I. General theory

Markus Reiher and Alexander Wolf

Additional information on J. Chem. Phys.

Exact decoupling of the Dirac Hamiltonian. I. General theory

Markus Reiher[a)] and Alexander Wolf[b)]

Lehrstuhl für Theoretische Chemie, Universität Bonn, Wegelerstrasse 12, D-53115 Bonn, Germany

Exact decoupling of positive- and negative-energy states in relativistic quantum chemistry is discussed in the framework of unitary transformation techniques. The obscure situation that each scheme of decoupling transformations relies on different, but very special parametrizations of the employed unitary matrices is critically analyzed. By applying the most general power series ansatz for the parametrization of the unitary matrices it is shown that all transformation protocols for decoupling the Dirac Hamiltonian have necessarily to start with an initial free-particle Foldy–Wouthuysen step. The purely numerical iteration scheme applying X-operator techniques to the Barysz–Sadlej–Snijders (BSS) Hamiltonian is compared to the analytical schemes of the Foldy–Wouthuysen (FW) and Douglas–Kroll–Hess (DKH) approaches. Relying on an illegal $1/c$ expansion of the Dirac Hamiltonian around the nonrelativistic limit, any higher-order FW transformation is in principle ill defined and doomed to fail, irrespective of the specific features of the external potential. It is shown that the DKH method is the only valid analytic unitary transformation scheme for the Dirac Hamiltonian. Its exact infinite-order version can be realized purely numerically by the BSS scheme, which is only able to yield matrix representations of the decoupled Hamiltonian but no analytic expressions for this operator. It is explained why a straightforward numerical iterative extension of the DKH procedure to arbitrary order employing matrix representations is not feasible within standard one-component electronic structure programs. A more sophisticated ansatz based on a symbolical evaluation of the DKH operators via a suitable parser routine is needed instead and introduced in Part II of this work. © *2004 American Institute of Physics.* [DOI: 10.1063/1.1768160]

I. INTRODUCTION

Over the last 30 years methods of relativistic quantum chemistry[1,2] have emerged as a well-established branch of theoretical chemistry. Due to a plethora of intrinsically "relativistic" molecules, such methods have become a widespread standard tool in modern electronic structure theory. Most compounds containing heavy-element atoms require a theoretical description based on the four-component Dirac equation to yield even qualitatively correct results. For highly accurate calculations of properties and energies relativistic methods also have to be applied even for molecules containing light, i.e., weakly charged nuclei—especially if spin–orbit coupling becomes decisive.

For one electron moving in an attractive external electric field the Dirac Hamiltonian is given in standard notation by[3]

$$H_D = c\,\boldsymbol{\alpha}\cdot\mathbf{p} + (\beta-1)mc^2 + V = \begin{pmatrix} V & c\,\boldsymbol{\sigma}\cdot\mathbf{p} \\ c\,\boldsymbol{\sigma}\cdot\mathbf{p} & V-2mc^2 \end{pmatrix}, \quad (1)$$

with $\boldsymbol{\sigma}$ being the familiar Pauli spin matrices. In order to get electronic binding energies from Eq. (1) directly comparable to the nonrelativistic Schrödinger theory the energy scale has been shifted by the rest energy mc^2 of the electron. Reflecting the (4×4) structure of the operators also the wave func-

tions are promoted to four-component spinors ϕ, whose upper (L) and lower (S) two components, ϕ^L and ϕ^S, respectively, may formally be related by a so-called (2×2) X operator,[4]

$$\phi^S = X\phi^L. \quad (2)$$

The external potential V is of Coulomb type, $V(\mathbf{r}) = -Ze^2/r$, or alternatively a potential derived from an extended nuclear charge distribution.[5]

Any first-quantized description of a many-electron system based on Eq. (1) is necessarily confined to a fixed number of particles, i.e., establishes always a no-pair theory. The possibility of creation and annihilation of electron-positron pairs requires the field-theoretical second-quantized framework of quantum electrodynamics (QED) and has to include interactions with the quantized electromagnetic field, i.e., with photons. For most situations of chemical interest, however, the threshold for pair-creation processes and excitations of the positronic degrees of freedom is far beyond the energy scale of the valence shell. It is thus a very good approximation to integrate them out at the very beginning and to neglect all QED corrections, restricting to a quasirelativistic formulation with a fixed number of particles. As a drawback of this unphysical first-quantized description negative-energy states show up, which are commonly assigned to positronic states, albeit positrons in the real world feature strictly positive energies. As a result of these pathologies it is perhaps best to follow Thaller and to adopt a pragmatic attitude towards the Dirac equation and "to consider it as a useful

[a)]Author to whom correspondence should be addressed. Electronic mail: reiher@thch.uni-bonn.de
[b)]Author to whom correspondence should be addressed. Electronic mail: awolf@thch.uni-bonn.de

model to be applied in situations, where the Schrödinger equation is already imprecise due to relativistic effects, but where the involved energies are too small to trigger pair-creation processes.'' [6]

Since the early days of relativistic quantum chemistry both numerical and basis-set methods being directly based on the four-component first-quantized Dirac framework have been applied to atomic and small molecular electronic structure calculations (see Refs. 7 and 8 and references therein for detailed reviews of this field). However, four-component methods are affected by two major difficulties. First, the proper representation of negative-energy states related to the lower components of the Dirac spinors requires substantial additional computational effort due to the requirements of kinetic balance.[9,10] The derivative operator occurring in the nonrelativistic limit of Eq. (2), $\phi^S \sim [(\boldsymbol{\sigma} \cdot \mathbf{p})/2mc]\phi^L$, enforces the small component basis to contain functions of higher angular momentum than the large component basis in order to represent both upper and lower parts of the Dirac Hamiltonian with equal quality for electronic solutions. This does not only increase the number of integrals over one- and two-electron operators significantly, but also enlarges the size of matrix representations. Second, the pure existence of small component basis sets in four-component approaches also gives rise to conceptual difficulties: Though negative-energy states are obtained in every self-consistent field (SCF) step, only the positive-energy (i.e., electronic) states are taken into account for the construction of the density matrix and the Fock operator. Thus, the vast majority of one-particle molecular spinors received in a four-component calculation is never used for the iterative setup of the SCF equations.

Therefore various approximate decoupling schemes have emerged over the last two decades in order to annihilate the contributions from the lower components and to restrict the theoretical description to the electronic, i.e., upper left part of the Dirac Hamiltonian. These schemes may be divided into elimination and transformation techniques, and we will only focus on the latter in this work. All transformation methods aim at the construction of suitable unitary transformations which decouple the Dirac Hamiltonian. Three main schemes have been suggested over the last 50 years: the Foldy–Wouthuysen (FW) transformations,[11] the Douglas–Kroll–Hess (DKH) approach,[12,13] and the analytical perturbative X-operator techniques.[14–16] Especially the DKH method has been introduced and developed to a viable relativistic tool of quantum chemistry by Hess and collaborators.[13,17–25] However, all these protocols have been truncated after a finite number of terms yielding thus only approximate quasirelativistic two- or one-component Hamiltonians with eigenvalues deviating more or less from the exact ones. It is thus highly desirable to follow the ideas of the general theory of effective Hamiltonians[26,27] and to devise quasirelativistic Hamiltonians, which describe only the electronic part of the spectrum—but with the same accuracy as provided by the original four-component methods.

It was not before 2002 that Barysz and Sadlej proposed the first infinite-order decoupling scheme for the Dirac Hamiltonian by applying a numerical iterative procedure in order to solve for the matrix representation of the X operator.[28] Since this procedure is based on ideas presented in an earlier paper by Barysz, Sadlej, and Snijders[14] we will refer to these numerical Hamiltonians as BSS Hamiltonians in this work. Inspired by their work, the purpose of this and the subsequent paper[29] is to present an analytical decoupling scheme based on the generalized Douglas-Kroll-Hess transformation[23] up to any predefined order in the external potential, which does not make any reference to the small component. The advantage of this procedure is that one obtains analytical closed-form expressions for the decoupled Hamiltonians. Furthermore one can analyze the contribution of each new order of the block-diagonal DKH Hamiltonian to energies or properties and thus establish rational criteria up to which order the decoupled Hamiltonians have to be expanded. Until now this analysis has only been possible up to fifth order in the external potential.[23] (After submission of this paper the first expressions for the sixth-order Hamiltonian have been given by van Wüllen.[30])

The organization of this paper is as follows. We start with a brief presentation of the general framework of unitary transformation techniques that decouple the Dirac Hamiltonian in order to introduce the basic notation. Starting from the most general ansatz for the unitary transformations the free-particle Foldy–Wouthuysen (fpFW) transformation is revisited in Sec. II B. This initial transformation may be followed by a sequence of suitably chosen further unitary transformations, which are commonly divided into a FW and a DKH branch depending on the chosen expansion parameter. Except for this choice, both transformation schemes are formally equivalent, which is demonstrated in this work by application of the most general parametrization of the unitary matrices. However, exceeding its inherent radius of convergence, it will be shown in Sec. II A. This FW expansion is completely ill defined and doomed to fail. It is demonstrated that this behavior is a general failure of any FW approach irrespective of the details of V. Only the DKH scheme, which avoids any expansion in $1/c$, yields regular and well-defined block-diagonal expressions for the Hamiltonian, which may be truncated after any order in V. Alternatively, the secondary unitary transformation following the fpFW step may also be performed purely numerically by an iterative scheme (BSS), whose discussion completes Sec. II. In Sec. III general properties of all decoupled Hamiltonians are discussed. In Sec. IV it is demonstrated why exact decoupling (i.e., up to any arbitrary order in the external potential) can in principle not be realized by a purely numerical recursion scheme within the stepwise DKH approach. This goal can only be achieved by symbolic operations which evaluate the Hamiltonian order by order in an analytical fashion. Such an algorithm will be presented in paper II.[29]

II. SURVEY OF UNITARY TRANSFORMATION METHODS

The Dirac Hamiltonian H_D as defined by Eq. (1) contains two even terms and one odd term,

$$H_D = (\beta - 1)mc^2 + V + c\,\boldsymbol{\alpha}\cdot\mathbf{p} \tag{3}$$

$$= \mathcal{E}_{[-2]} + \mathcal{E}_{[0]} + \mathcal{O}_{[-1]} \tag{4}$$

$$= \mathcal{E}_0 + \mathcal{E}_1 + \mathcal{O}_0, \tag{5}$$

where subscripts in brackets specify the order in $1/c$ and subscripts without brackets denote the order in the external potential V of the corresponding term. Note that each term of the Hamiltonian can be classified uniquely and equally well with respect to these two parameters, and it is purely a matter of personal taste which scheme one prefers. Even terms $(\mathcal{E})$ are block diagonal and commute with β, whereas odd terms $(\mathcal{O})$ are off-diagonal and anticommute with β.

If the odd term $c\,\boldsymbol{\alpha}\cdot\mathbf{p}$ were not present in the Hamiltonian, the upper and lower components of the Dirac spinor ϕ would already be decoupled and the Dirac equation could be solved independently for the two-spinors ϕ^L and ϕ^S. Therefore one aims at finding a suitable unitary transformation U which achieves this goal,

$$H_{bd} = U H_D U^\dagger = \begin{pmatrix} h_+ & 0 \\ 0 & h_- \end{pmatrix}. \tag{6}$$

Unfortunately no energy-independent closed-form solutions for the transformation U are known in general. However, U can always be expressed in terms of the X operator defined by Eq. (2) relating the large and small components of the Dirac spinor ϕ,[4] and its most general form is given by[22,23]

$$U = U(X) = \begin{pmatrix} (1+X^\dagger X)^{-1/2} & (1+X^\dagger X)^{-1/2}X^\dagger \\ -e^{i\varphi}(1+XX^\dagger)^{-1/2}X & e^{i\varphi}(1+XX^\dagger)^{-1/2} \end{pmatrix}, \tag{7}$$

where $\varphi \in [0, 2\pi]$ is an arbitrary phase factor. With this form of the unitary transformation the electronic part of the decoupled Hamiltonian may immediately be expressed as

$$h_+ = \frac{1}{\sqrt{1+X^\dagger X}} \{ V + c\,\boldsymbol{\sigma}\cdot\mathbf{p}X + X^\dagger c\,\boldsymbol{\sigma}\cdot\mathbf{p} + X^\dagger(V - 2mc^2)X \}$$

$$\times \frac{1}{\sqrt{1+X^\dagger X}}, \tag{8}$$

and is independent of the arbitrary phase φ. If the energy-independent X operator were known one could solve the simplified eigenvalue problem for h_+ instead of H_D which would no longer be plagued by negative-energy solutions. However, an energy-independent expression for the X operator is only known for very special potentials excluding the Coulomb potential.[31]

By exploitation of the Dirac equation it is very easy to give an energy-dependent expression for the operator X, relating only the large and small components of the eigenspinor with energy eigenvalue E: $X = X(E) = (E - V + 2mc^2)^{-1}c\,\boldsymbol{\sigma}\cdot\mathbf{p}$. As a consequence, each eigensolution of the Dirac equation has its own energy-dependent X operator, whereas there is only one energy-independent X operator valid for all infinitely many eigenstates of H_D. In the former case the electronic Hamiltonian h_+ is a very complicated function of its own eigenvalues and the eigenvalue equation will have only normalizable solutions if E is indeed a true

electronic eigenvalue of the Dirac Hamiltonian. Obviously, solving these eigenvalue problems for energy-dependent operators is a formidable task and at least as complicated as the original problem.

Therefore one has two options: either to decompose the overall unitary transformation U into a sequence of simpler unitary transformations,

$$U = \cdots U_4 U_3 U_2 U_1 U_0, \tag{9}$$

which can be parametrized such that the odd term of lowest order in a predetermined power-series expansion parameter is eliminated stepwise, or to restrict to a purely numerical solution for the operator X.

A. Requirements for the initial transformation U_0

Every transformation method aiming at decoupling of the Dirac Hamiltonian,

$$H_D = \mathcal{E}_{[-2]} + \mathcal{E}_{[0]}^{(0)} + \mathcal{O}_{[-1]}^{(0)} \tag{10}$$

$$= \mathcal{E}_0^{(0)} + \mathcal{E}_1^{(0)} + \mathcal{O}_0^{(0)}, \tag{11}$$

has necessarily to eliminate the odd term $\mathcal{O}_{[-1]}^{(0)} = \mathcal{O}_0^{(0)}$ in the first step. Note that we have chosen superscripts in parentheses here to indicate that these terms belong to the initial, i.e., untransformed Hamiltonian H_D and might thus not yet be the final terms with the corresponding order occurring in the decoupled Hamiltonian. Only the even term of order -2 according to the $1/c$ classification of Eq. (10) does already constitute the final expression appearing in the block-diagonal Hamiltonian and carries thus no superscript.

Since the odd term does not depend on the external potential it has to be removed by a unitary transformation independent of V being able to cancel the odd term $c\,\boldsymbol{\alpha}\cdot\mathbf{p}$, which is linear in both c and $\boldsymbol{\alpha}$. The most general ansatz for the initial transformation U_0 is thus to parametrize it as a power-series expansion (as it was done for all $U_{k>0}$ matrices in Ref. 23),

$$U_0 = U_0(W_{[0]}) = a_{0,0}\mathbf{1} + \sum_{k=1}^\infty a_{0,k} W_{[0]}^k, \tag{12}$$

where $W_{[0]}$ is an odd and anti-Hermitian operator, which is first order in $1/c$ but independent of V. If the operator $W_{[0]}$ were of zeroth or even lower order in $1/c$ it would introduce odd terms of lower than minus first order in the transformed Hamiltonian due to the presence of the term $\mathcal{E}_{[-2]}$, and if it were of second or even higher order in $1/c$, it could not account for eliminating $\mathcal{O}_{[-1]}^{(0)}$. Note that—due to the structure of the Dirac Hamiltonian, whose only odd term is $c\,\boldsymbol{\alpha}\cdot\mathbf{p}$—there is only little flexibility for the form of $W_{[0]}$: as a matter of fact it has to be proportional to $(\boldsymbol{\alpha}\cdot\mathbf{p})/c$.

The anti-Hermiticity requirement for the operator $W_{[0]}$ is purely due to technical reasons, since the adjoint unitary transformation $U_0^\dagger$ is then easily obtained employing the relation $(W_{[0]}^k)^\dagger = (-1)^k W_{[0]}^k$. The expansion coefficients $a_{0,k}$ have to satisfy the unitarity conditions explicitly given in Ref. 23; they are omitted here for brevity. We only recall that $a_{0,0} = \pm 1$, and all odd coefficients $(a_{0,1}, a_{0,3}, \ldots)$ can be chosen arbitrarily whereas all even coefficients

$(a_{0,2}, a_{0,4}, \dots)$ are uniquely determined by the lower odd coefficients. By inspection of the transformed Hamiltonian

$$H_1 = U_0 H_D U_0^\dagger = \mathcal{E}_{[-2]} + \mathcal{E}_{[0]} + \mathcal{E}_{[2]}$$
$$+ \underbrace{\mathcal{O}_{[-1]}^{(0)} + a_{0,0} a_{0,1} [W_{[0]}, \mathcal{E}_{[-2]}]}_{\mathcal{O}_{[-1]}^{(1)}}$$
$$+ \sum_{k=2}^{\infty} \mathcal{E}_{[2k]}^{(1)} + \sum_{k=1}^{\infty} \mathcal{O}_{[2k-1]}^{(1)}, \tag{13}$$

the condition, which one has to impose on $W_{[0]}$ in order to eliminate the lowest-order odd term occurring in H_1, is revealed,

$$\mathcal{O}_{[-1]}^{(1)} = \mathcal{O}_{[-1]}^{(0)} + a_{0,0} a_{0,1} [W_{[0]}, \mathcal{E}_{[-2]}] \overset{!}{=} 0. \tag{14}$$

Due to the very simple structure of $\mathcal{E}_{[-2]}$, which is only a constant independent of both $\mathbf{p}$ and V, inversion of the commutator is straightforward and yields immediately the odd and anti-Hermitian operator

$$W_{[0]} = \frac{a_{0,0}}{a_{0,1}} \beta \frac{c\,\boldsymbol{\alpha}\cdot\mathbf{p}}{2mc^2} = \frac{a_{0,0}}{a_{0,1}} \beta \frac{\boldsymbol{\alpha}\cdot\mathbf{p}}{2mc}, \tag{15}$$

which is indeed first order in $1/c$. Depending on the choice of the odd coefficients $a_{0,2k+1}$, we have thus found infinitely many different unitary parametrizations U_0 which decouple the Dirac Hamiltonian up to zeroth order in $1/c$. The resulting Hamiltonians H_1 consist of infinitely many terms which can all be given a definite order in both $1/c$ and V, e.g., the odd operator $\mathcal{O}_{[1]}^{(1)}$, which was obtained after the first ''(1)'' transformation is first order '[1]' in $1/c$ as well as first order '1' in V,

$$\mathcal{O}_{[1]}^{(1)} = \left[\beta \frac{\boldsymbol{\alpha}\cdot\mathbf{p}}{2mc}, V \right]. \tag{16}$$

Since U_0 is independent of the potential V, all the various Hamiltonians H_1, which are all given by a specific choice of the set of coefficients $a_{0,k}$, contain only terms of at most first order in V, albeit arbitrarily high orders in $1/c$. The leading terms of H_1 given by Eq. (13) are independent of the chosen parametrization of U_0 (i.e., independent of the specific choice of the expansion coefficients $a_{0,k}$) and are given by

$$\mathcal{E}_{[-2]} = (\beta - 1) mc^2, \tag{17}$$

$$\mathcal{E}_{[0]} = V + \beta \frac{\mathbf{p}^2}{2m}, \tag{18}$$

$$\mathcal{E}_{[2]} = -\beta \frac{\mathbf{p}^4}{8m^3c^2} - \frac{[\boldsymbol{\alpha}\cdot\mathbf{p}, [\boldsymbol{\alpha}\cdot\mathbf{p}, V]]}{8m^2c^2}. \tag{19}$$

One clearly recovers the rest energy $\mathcal{E}_{[-2]}$ and the nonrelativistic contributions $\mathcal{E}_{[0]}$ as the leading terms of the FW series given above. The expression for $\mathcal{E}_{[2]}$ can be cast into the more familiar form

$$\mathcal{E}_{[2]} = -\beta \frac{\mathbf{p}^4}{8m^3c^2} + \frac{\hbar^2}{8m^2c^2} (\Delta V)$$
$$+ \frac{\hbar}{4m^2c^2} \boldsymbol{\Sigma}\cdot[(\boldsymbol{\nabla}V)\times\mathbf{p}], \tag{20}$$

explicitly featuring the mass-velocity, the Darwin, and the spin-orbit coupling term, where Dirac's $\boldsymbol{\Sigma}$ matrices are just a four-component generalization of Pauli's spin matrices.

One might be tempted to consider all these infinitely many Hamiltonians H_1 as being equivalent, since they all seem to possess the same spectrum due to the unitary structure of the transformation U_0. However, this perception would be completely erroneous since the situation is more subtle. After having chosen a specific set of expansion coefficients $a_{0,k}$ in accordance with the unitarity conditions mentioned above, one must check the convergence of both the power-series ansatz for U_0 and the resulting series expansion of the Hamiltonian H_1. Especially since $W_{[0]}$ is linear in the momentum operator $\mathbf{p}$ it is by no means clear if these series expansions are valid for large momenta, i.e., if they are related to the Dirac Hamiltonian at all.

There is, however, one very special parametrization for the transformation U_0, which may be converted into a closed-form expression and which allows for a closed-form evaluation of the Hamiltonian H_1. Its expansion coefficients are given by

$$a_{0,0} = 1, \quad a_{0,1} = 1, \quad a_{0,2} = \tfrac{1}{2}, \quad a_{0,3} = \tfrac{3}{2}, \quad a_{0,4} = \tfrac{11}{8},$$
$$a_{0,5} = \tfrac{31}{8}, \quad a_{0,6} = \tfrac{69}{16}, \quad a_{0,7} = \tfrac{187}{16}, \quad a_{0,8} = \tfrac{1843}{128}, \tag{21}$$
$$a_{0,9} = \tfrac{4859}{128}, \cdots$$

and yield the familiar analytical closed-form expression for the traditional fpFW transformation,[11]

$$U_0 = \exp\left(\beta \frac{\boldsymbol{\alpha}\cdot\mathbf{p}}{2p} \arctan \frac{p}{mc} \right) = A_p (1 + \beta R_p), \tag{22}$$

with

$$A_p = \sqrt{\frac{E_p + mc^2}{2E_p}}, \quad E_p = \sqrt{\mathbf{p}^2 c^2 + m^2 c^4},$$
$$R_p = \frac{c\,\boldsymbol{\alpha}\cdot\mathbf{p}}{E_p + mc^2} = \boldsymbol{\alpha}\cdot\mathbf{P}_p. \tag{23}$$

Note that the coefficients defined by Eq. (21) satisfy the unitarity conditions given in Ref. 23 and do thus indeed constitute a unitary transformation. Accordingly, the fpFW Hamiltonian H_1 obtained by this specific transformation may then be expressed in closed form,

$$H_1 = \underbrace{\beta E_p - mc^2}_{\varepsilon_0} + \underbrace{A_p (V + R_p V R_p) A_p}_{\varepsilon_1} + \underbrace{\beta A_p [R_p, V] A_p}_{\mathcal{O}_1^{(1)}}, \tag{24}$$

and contains only well-defined expressions valid for all real values of the momentum p.

Only the specific choice for U_0 established by Eqs. (21) and (22) facilitates the opportunity of a closed-form evaluation of the Hamiltonian H_1 and, hence, abandons the neces-

J. Chem. Phys., Vol. 121, No. 5, 1 August 2004

Decoupling of the Dirac Hamiltonian   2041

sity of an expansion in $1/c$. The legitimacy of this expansion and hence the issue of higher-order FW transformations will be analyzed in the following section in detail. As a consequence of this choice for U_0, we arrive at a compact and concise expression for H_1, Eq. (24), where each constituent term can only be classified according to its order in V and no longer according to its order in $1/c$. Each term occurring in Eq. (24) contains arbitrarily high orders of $1/c$ as it is easily seen by a Taylor series expansion in this parameter. We would like to emphasize that our derivation of the initial step of any transformation scheme employs the most general ansatz possible. Although the final result is the familiar fpFW transformation being known and in use for decades, we stress that our analysis yields this result as a necessary consequence obtained from the most general viewpoint.

If there were no potential V present and the particle was moving freely, the Hamiltonian H_1 would already be block diagonal. In the presence of a potential, however, all odd terms occurring in the resulting Hamiltonian H_1 are at least first order in $1/c$ and exactly first order in V, in contrast to the original odd term of the Dirac Hamiltonian, which is proportional to c and independent of V, respectively. Even for nonvanishing potential V the importance of the odd blocks has thus been diminished. In the following we will analyze different strategies to further decouple the Hamiltonian H_1 in the presence of an external potential.

B. Pathologies of higher-order Foldy–Wouthuysen transformations

Following up the initial fpFW transformation U_0 one could try to establish a sequence of further unitary transformations U_i ($i=1,2,3,\dots$) which eliminate the respective lowest-order odd term in $1/c$ in each step. This procedure results in an expansion of the block-diagonal Hamiltonian H_{bd} in even terms of ascending order in $1/c$, and was first suggested in 1950 by Foldy and Wouthuysen.[11] We will thus strictly use the phrase FW transformation in this work to denote a $1/c$ expansion of H_{bd} rather than any arbitrary decoupling transformation of the Dirac Hamiltonian. The latter convention may be found in some papers and should not be confused with the notation of this paper. Though the $1/c$ FW expansion of the Hamiltonian is completely ill defined, as we will discuss in this section, this transformation protocol has been the subject of many investigations over the last decades and can still be found in many textbooks[32–35] and recent research papers.[36–40] Therefore, there is a need for a thorough discussion of the FW scheme in the context of a complete account on transformation techniques, which is pursued in this paper.

Within the FW framework, the decoupled Hamiltonian is formally given as a series of even terms of well-defined order in $1/c$,

$$H_{bd}=\begin{pmatrix} h_+ & 0 \\ 0 & h_- \end{pmatrix}$$

$$=\sum_{k=-1}^{\infty} \mathcal{E}_{[2k]}=\sum_{k=-1}^{\infty}\begin{pmatrix} \mathcal{E}_{[2k]+} & 0 \\ 0 & \mathcal{E}_{[2k]-} \end{pmatrix}, \tag{25}$$

which is achieved by a sequence of unitary transformations

$U_i=U_i(W_{[i]})$. In most presentations of the FW transformation the exponential function parametrization $U_{[i]}=\exp(W_{[i]})$ is applied. However, the specific choice of this parametrization does not matter at all, since one necessarily has to expand U_i into a power series in order to evaluate the Hamiltonian. This situation is completely analogous to the Douglas–Kroll–Hess protocol, where it was believed for many years that the specific square-root parametrization ($U_i=\sqrt{1+W_i^2}+W_i$) introduced by Douglas and Kroll[12] is decisive for the success of the method. However, in 2002 it was realized that the specific form of this parametrization does not matter at all and the most general unitary form leading to the generalized DKH approach was introduced.[22,23]

In order to guarantee the most general setup and for better comparison with the DKH method, the most general parametrization for the FW transformation is therefore employed, i.e., U_i is parametrized as a power-series expansion in an odd and anti-Hermitian operator $W_{[i]}$, which is of $(2i+1)$th order in $1/c$. After n transformation steps the intermediate, partially transformed Hamiltonian H_n has the following structure:

$$H_n=U_{n-1}H_{n-1}U_{n-1}^\dagger$$

$$=\sum_{k=-1}^{2n-1} \mathcal{E}_{[2k]}+\sum_{k=n}^{2n-1} \mathcal{O}_{[2k-1]}^{(n)}$$

$$+\sum_{k=2n}^{\infty} (\mathcal{E}_{[2k]}^{(n)}+\mathcal{O}_{[2k-1]}^{(n)}). \tag{26}$$

The next transformation U_n is determined by the odd and anti-Hermitian operator $W_{[n]}$, which has to be chosen as

$$W_{[n]}=\frac{a_{n,0}}{a_{n,1}}\beta \frac{\mathcal{O}_{[2n-1]}^{(n)}}{2mc^2} \tag{27}$$

in order to eliminate the odd term $\mathcal{O}_{[2n-1]}^{(n+1)}=\mathcal{O}_{[2n-1]}^{(n)}+a_{n,0}a_{n,1}[W_{[n]},\mathcal{E}_{[-2]}]$ of H_{n+1}.

Though this FW procedure could formally be repeated until exact decoupling seems to be achieved, the resulting even terms are highly singular and ill defined: they are not related to the original Dirac Hamiltonian, except of the leading nonrelativistic term $\mathcal{E}_{[0]}$ and to some extend the first relativistic correction $\mathcal{E}_{[2]}$, which may at least be evaluated perturbatively. The reason for this failure of the higher-order FW transformation is that it necessarily relies on an illegal, i.e., nonconvergent $1/c$ expansion of all terms occurring in the fpFW Hamiltonian H_1 defined by Eq. (24) in order to classify each term. Such a power-series expansion is, however, only permitted for analytic, i.e., holomorphic functions and must never be extended beyond a singular point. Since the square root occurring in the relativistic energy-momentum relation E_p of Eq. (23) possesses branching points at $x\equiv p/mc=\pm i$, any series expansion of E_p around the static nonrelativistic limit $x=0$ within the simply connected, sliced complex plane $\mathbb{C}_1$ is only related to the exact expression for E_p for nonultrarelativistic values of the momentum, i.e., $|x|<1$, as it is graphically illustrated in Fig. 1. Obviously, the alternative choice for the sliced complex
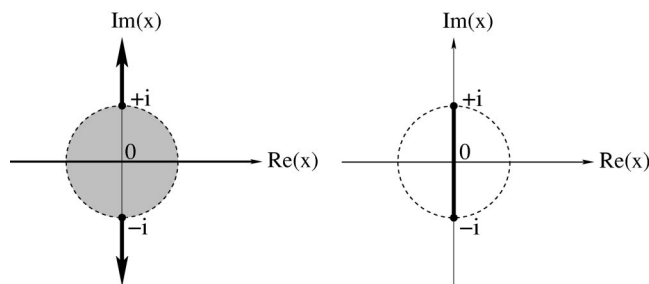
FIG. 1. Schematic representation of the sliced complex planes $\mathbb{C}_1$ (left) and $\mathbb{C}_2$ (right). After removal of suitable parts of the imaginary axis, indicated by thick lines, both regions are legitimate, single-valued domains for E_p excluding the branching points at $\pm i$. However, only within the shaded disk of radius 1 series expansions of E_p around $x=0$ are legitimate, and only the simply connected domain $\mathbb{C}_1$ contains the whole real axis, i.e., all physically relevant values for the momentum $x=p/mc$.

plane given in Fig. 1, $\mathbb{C}_2$, is not suited here, since it does not even contain the nonrelativistic case of vanishing momentum at $x=0$.

This is most easily seen by rewritting E_p as

$$E_p = mc^2 \sqrt{1+x^2} \equiv mc^2 \exp[\tfrac{1}{2}\ln(1+x^2)], \quad (28)$$

which clearly demonstrates that the branching singularities of the logarithm at $x=\pm i$ confine the domain of convergence of any power-series expansion of E_p to $|x|<1$. For larger momenta, the series expansion up to any arbitrary or even infinite order in p/mc does not represent the original function E_p. For these momenta, this series does not even converge at all, and the singular behavior of the series expansions of E_p, and hence of A_p and R_p, becomes thus the worse the more terms of the expansions are taken into account.

In Fig. 2 these pathologies are depicted graphically for E_p and A_p. The exact expressions are compared to the series expansions around $x=0$ which have been truncated after the term of $\mathcal{O}(x^{12})$. For small momenta $x<1$ the series expansions are excellent approximations to the exact closed-form expressions, whereas for larger momenta $x>1$ the truncated series expansions result in very different Hamiltonians with different spectra and eigensystems. Furthermore, the untruncated expansions are divergent for all $x>1$.

The intrinsic failure of the FW protocol is thus doubtlessly related to the illegal $1/c$ expansion of the kinetic term E_p, which does not bear any reference to the external potential V. However, in the literature the ill-defined behavior of

the FW transformation has sometimes erroneously been assigned to the singular behavior of the Coulomb potential near the nucleus, and even the existence of the correct nonrelativistic limit of the FW Hamiltonian is sometimes subject of dispute. Due to Eqs. (17) and (18) and the analysis given above the nonrelativistic limit $c \to \infty$, i.e., $x \to 0$ is obviously well defined and for positive-energy solutions given by the Schrödinger Hamiltonian $H_{\text{nr}} = \mathbf{p}^2/2m + V$.

The failure of the FW transformation does not impose a major fundamental problem on block-diagonalization procedures applied to the Dirac Hamiltonian. It rather demonstrates that neither these unitary transformations nor the Hamiltonian must be expanded naively in $1/c$, i.e., transgressing the domain of convergence of the resulting series. The initial transformation step discussed in Sec. II A has thus necessarily to be chosen as the closed-form, analytical fpFW transformation defined by Eq. (22).

C. Douglas–Kroll–Hess transformations

If an elegant expansion of the decoupled Hamiltonian similar to Eq. (25) is to be preserved for both analytical and numerical investigations, one necessarily has to classify each term of this expansion according to a new order parameter—the electron-nucleus interaction V, which is the only remaining possibility. The key feature of this expansion is that the closed-form expressions of the fpFW Hamiltonian H_1 given by Eq. (24) remain untouched during the whole transformation procedure, and the resulting block-diagonal Hamiltonian is well defined for all momenta $p \in \mathbb{R}_0^+$ and features exactly the same spectrum as the original Dirac Hamiltonian. This decoupling scheme has first been mentioned by Douglas and Kroll[12] in 1974 and—due to the pioneering work of Hess,[13,41] who first realized its efficiency and made quantum chemical implementations feasible—has become one of the most successful quasirelativistic methods over the last two decades. As a consequence of the above discussion, it is the only valid analytic expansion technique for the Dirac Hamiltonian.

The DKH procedure aims at decoupling of the Hamiltonian by a sequence of further unitary transformations following the initial fpFW step. The final Hamiltonian may then be written as

$$H_{bd} = \cdots U_2 U_1 H_1 U_1^\dagger U_2^\dagger \cdots = \begin{pmatrix} h_+ & 0 \\ 0 & h_- \end{pmatrix} = \sum_{k=0}^{\infty} \mathcal{E}_k, \quad (29)$$
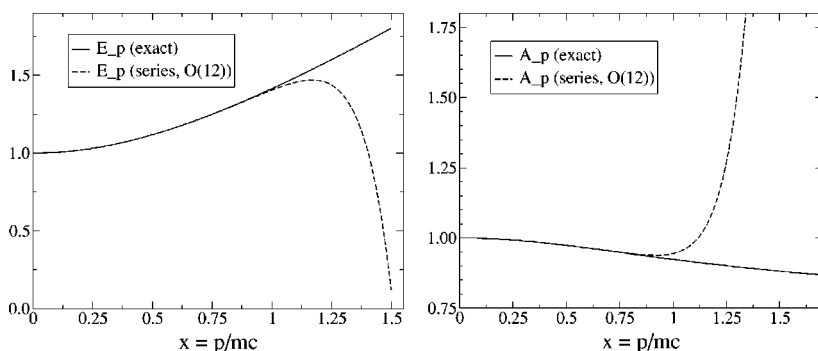

FIG. 2. Comparison of the exact expression for E_p (left) and A_p (right) with their series expansions around the nonrelativistic limit $x \equiv p/mc = 0$ up to $\mathcal{O}(x^{12})$. The series expansions do only represent the exact expressions for $x<1$. Beyond the branching point of the square root of E_p [cf., Eq. (23)] at $x=1$ the series expansions and hence the expanded Hamiltonians are completely ill defined.

where each term $\mathcal{E}_k$ comprises all contributions which are exactly of kth order in V.

We would like to recall that the specific form of the unitary matrices U_i is not decisive for the success of the DKH protocol.[23] Analyticity of these matrices in an *a priori* unknown odd and anti-Hermitian operator W_i, which has to be exactly of ith order in V in order to eliminate the lowest-order odd term is the only essential feature of U_i needed for the DKH approach. The determination of both W_i and the even terms $\mathcal{E}_k$ requires the transformations U_i to be expanded in a power series around $W_i=0$ in any case. Again, we therefore apply the most general parametrizations[22,23] for the transformations U_i in this work,

$$U_i = U_i(W_i) = a_{i,0}\mathbf{1} + \sum_{k=1}^{\infty} a_{i,k} W_i^k, \qquad (30)$$

with the coefficients $a_{i,k}$ satisfying the unitarity conditions mentioned earlier. All the different parametrizations for unitary matrices U_i occurring in the literature, e.g., the square-root parametrization $U_i = \sqrt{1+W_i^2} + W_i$ given by Douglas and Kroll,[12] are only special cases of this most general setup.[23]

Up to fourth order ($n=4$), the DKH Hamiltonians

$$H_{\text{DKH}n} = \sum_{k=0}^{4} \mathcal{E}_k + \sum_{k=5}^{n} \mathcal{E}_k^U \qquad (31)$$

are independent of the chosen parametrizations of the unitary transformations U_i. The fifth- and all higher-order terms, $\mathcal{E}_{k\geqslant5}^U$, depend on the coefficients $a_{i,k}$ of the general parametrization[23] (see Ref. 30 for numerical results obtained for different $\mathcal{E}_k^U$ operators with $k=5,6$). The spectrum and the eigenfunctions of the exact, i.e., untruncated and block-diagonal Hamiltonian $H_{bd}=H_{\text{DKH}\infty}$ do certainly not depend on the choice of the coefficients $a_{i,k}$, since all transformations applied have been unitary. We have thus found an infinite family of completely equivalent Hamiltonians, which do all describe the Dirac electron perfectly well and which are related via unitary transformations in Hilbert space. Given two different parametrizations of the transformations U_i ($i\geqslant1$), i.e., two different sets of expansion coefficients $a_{i,k}$, we could establish two different unitary transformations U and U' yielding two block-diagonal Hamiltonians H_{bd} and H_{bd}'. However, these two decoupled Hamiltonians are simply related to one another by the unitary transformation

$$H_{bd} = \underbrace{UU'^{\dagger}}_{S} H_{bd}' U'U^{\dagger} = SH_{bd}'S^{\dagger}. \qquad (32)$$

The DKH scheme is usually referred to as an expansion in the external potential V or—due to the specific form of the Coulomb potential $V(r)=-Ze^2/r$—an expansion in the coupling strength Ze^2. This might at first glance rise questions about the convergence of the DKH series, in sharp contrast to its excellent numerical performance. By inspection of the explicit expressions for each term $\mathcal{E}_k$ of the DKH Hamiltonians this rapidly convergent behavior is not surprising, since the true expansion parameter is the damped potential
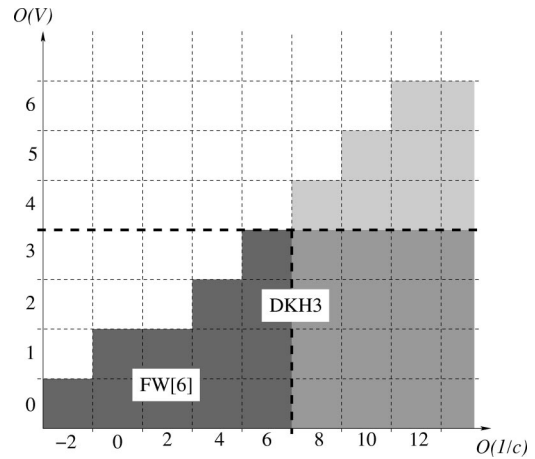

FIG. 3. Schematic representation of all terms occurring within the block-diagonal Hamiltonian H_{bd} of Eq. (6) (shaded fields). All terms have been expanded in $1/c$ in order to classify them uniquely according to their order in both V and $1/c$. Each order of a FW Hamiltonian, e.g., $H_{\text{FW}[6]}$, sums column by column (vertically) all terms up to $\mathcal{O}(1/c^6)$, whereas each order of a DKH Hamiltonian, e.g., $H_{\text{DKH}3}$ includes line by line summation over infinitely many terms. See text for further details.

$$\tilde{V}_{ij} = \frac{V_{ij}}{E_i+E_j} = \frac{V(\mathbf{p}_i,\mathbf{p}_j)}{\sqrt{\mathbf{p}_i^2c^2+m^2c^4}+\sqrt{\mathbf{p}_j^2c^2+m^2c^4}}, \qquad (33)$$

which is at least suppressed by a factor of $1/(2mc^2)$ as compared to the bare Coulomb potential V. Each order $\mathcal{E}_k$ ($k\geqslant1$) of the DKH Hamiltonian contains exactly $(k-1)$ factors of $\tilde{V}$. Since also the factor R_p of Eq. (23) is of leading order $1/c$, i.e., its lowest-order term occurring in a virtual series expansion in this parameter would be $\mathcal{O}(1/c)$, each term $\mathcal{E}_k$ is of leading order Z^{k-1}/c^{2k} (for $k\geqslant2$). Therefore, each higher-order term $\mathcal{E}_k$ of the Hamiltonian is formally suppressed by Z/c^2 as compared to the previous approximation $H_{\text{DKH}(k-1)}$, which is the basis for the rapid convergence properties of the DKH protocol. Furthermore, this analysis demonstrates that even the Coulomb singularity near a point-like nucleus does not cause any problems within the DKH scheme. It is strongly damped in the higher-order DKH terms rather than giving rise to nonintegrable and singular $1/r^k$ expressions.

The superior performance of the DKH scheme as compared to the ill-defined FW scheme is illustrated in Fig. 3. Each term occurring in the block-diagonal Hamiltonian H_{bd}—after virtual expansion in both $1/c$ and V—can be given a unique order in both $1/c$ and V, and is accordingly represented by one square in the $(1/c)$-V plane of Fig. 3. As it is illustrated for the sixth-order FW Hamiltonian $H_{\text{FW}[6]}$, it sums only very few terms (indicated by dark gray color) contributing to the exact block-diagonal Hamiltonian H_{bd}. The DKH scheme, however, sums all terms up to a given order in V, irrespective of their order in $1/c$. Thus, even the third-order Hamiltonian $H_{\text{DKH}3}$ contains all terms of the sixth-order FW Hamiltonian plus infinitely many additional terms (indicated in normal gray). This situation may be summarized by the statement that the DKH Hamiltonians contain "implicit partial summations up to infinite order in $1/c$." Only the terms indicated by light gray in Fig. 3 are not cov-

ered by the third-order DKH Hamiltonian, but might be taken into account by transition to higher-order approximations.

D. Iterative numerical solutions for *X*

The DKH scheme yields analytical closed-form expressions for each order in the expansion parameter $\tilde{V}$. Alternatively, a purely numerical procedure might be applied in order to derive a matrix representation of the final block-diagonal Hamiltonian. This approach was suggested quite recently by Barysz and Sadlej in a series of papers,[15,16,28] and makes extensive use of the X-operator formalism.

Following the mandatory initial fpFW transformation U_0, the sequence of subsequent unitary transformations U_i ($i \geqslant 1$) of Eq. (9) applied to the fpFW Hamiltonian H_1 is united to only one residual transformation step U_1, which is formally parametrized by an operator R, which constitutes the exact relationship between the upper and lower components of the fpFW spinor ϕ_1 for electronic (class I) solutions,[4] $\phi_1^S = R\phi_1^L$. The new symbol R has been used for this operator in order to distinguish it from the previously introduced X operator relating the small and large components of the original, i.e., untransformed Dirac spinor ϕ via Eq. (2). The matrix U_1 may then be expressed as

$$U_1 = \begin{pmatrix} (1+R^\dagger R)^{-1/2} & (1+R^\dagger R)^{-1/2}R^\dagger \\ -(1+RR^\dagger)^{-1/2}R & (1+RR^\dagger)^{-1/2} \end{pmatrix}, \qquad (34)$$

where the arbitrary phase of Eq. (7) has been fixed to zero. Similarly to Eq. (8) the requirement of vanishing off-diagonal blocks of H_{bd} leads to a condition imposed on the operator R,

$$R = [H_1^{SS}]^{-1}\{-H_1^{LS} + RH_1^{LL} + RH_1^{SL}R\}, \qquad (35)$$

where an obvious notation for the (2×2) components of H_1 has been introduced. After insertion of the explicit expressions for the components of the fpFW Hamiltonian H_1 given by Eq. (24), this equation may be converted to the form

$$\begin{aligned} E_p R + R E_p = {} & A_p[\boldsymbol{\sigma} \cdot \mathbf{P}_p, V]A_p + [A_p V A_p, R] \\ & + [A_p \boldsymbol{\sigma} \cdot \mathbf{P}_p V \boldsymbol{\sigma} \cdot \mathbf{P}_p A_p, R] \\ & + R A_p[\boldsymbol{\sigma} \cdot \mathbf{P}_p, V]A_p R. \end{aligned} \qquad (36)$$

For its numerical solution within a basis set approximation, it has first to be multiplied by the operator $P_p^{-1} \boldsymbol{\sigma} \cdot \mathbf{P}_p$ from the left in order to reduce it to computationally feasible form, where $P_p = |\mathbf{P}_p|$ is a scalar operator.[28] Subsequent introduction of the operator $Q = P_p^{-1} \boldsymbol{\sigma} \cdot \mathbf{P}_p R$ and frequent use of the relation $\boldsymbol{\sigma} \cdot \mathbf{P}_p \boldsymbol{\sigma} \cdot \mathbf{P}_p = P_p^2$ yields an equation for the (2×2) operator $Q = Q(R)$ analogous to Eq. (36). Afterwards, this equation can be solved by purely numerical iterative techniques and the matrix representation of the operator Q is obtained.[28] This result seems to be the best representation of the operator Q that can be achieved within a given basis and is only limited by machine accuracy. Note that all expressions used for the calculation of the matrix representation of the operator Q depend only on the squared momentum $\mathbf{p}^2$ rather than on the momentum variable itself. This is a very subtle point about the BSS approach and actually the key feature of its computational feasibility. No matrix representation of the original operator R can be obtained within a purely two-component transformation scheme.

Equation (36) is nonlinear, i.e., quadratic and, thus, bears the possibility of negative-energy solutions for the operator R and thus also for Q. The choice towards the positive-energy branch has to be implemented via the boundary conditions imposed on the numerical iterative technique. Essentially, Q and hence R have to be "small" operators with operator norms much smaller than unity.

Once the matrix representation of the operator Q is known, it can immediately be used to determine the desired matrix representation of the two-component Hamiltonian h_+ [analogous to Eq. (8)]. Numerical results obtained with this method for one-electron atoms are very encouraging:[28] For the first time Barysz and Sadlej could obtain the exact relativistic result, as it has so far only be accessible by four-component methods.

E. Comparison of FW, DKH, and BSS schemes

We conclude this section with a schematic comparison of the three different unitary decoupling schemes discussed so far. In Table I the most essential features of the FW, DKH, and BSS schemes are summarized and compared to each other. The historical FW scheme with its inherent $1/c$ expansions is completely ill defined and yields erroneous results for any order. The DKH approach avoids any expansions in $1/c$ and thus represents the only analytical, regular, and well-defined decoupling transformation for the Dirac Hamiltonian, which might be accomplished up to any desired order in V. The infinite-order Hamiltonian $H_{\text{DKH}\infty}$ is indeed completely decoupled and exact. All even terms $\mathcal{E}_k$ can be analyzed individually order by order such that the importance of any single order for relativistic effects in molecules might be estimated. Finally, the BSS Hamiltonian represents a purely numerical representation of the infinite-order DKH Hamiltonian, i.e., the matrix representations of all even terms $\mathcal{E}_k$ are explicitly summed up.

III. GENERAL PROPERTIES OF DECOUPLED DIRAC HAMILTONIANS

The decoupled Dirac Hamiltonian H_{bd} given by Eq. (6) features a very convenient mathematical structure, which allows for efficient and well-defined computational processing. These features do not depend on details of its derivation, but are common properties of all block-diagonal Hamiltonians. We shall briefly discuss the basic reasons of these salient features in this section.

Within any expansion of the block-diagonal Hamiltonian H_{bd}, all even terms $\mathcal{E}_k$ depend only quadratically on the momentum operator rather than on the linear operator $\mathbf{p}$ itself. The origin of this peculiarity lies, of course, in the structure of the original Dirac Hamiltonian, whose only odd component is the kinetic term $c\boldsymbol{\alpha} \cdot \mathbf{p}$, which is linear in $\mathbf{p}$. Since there is no other odd term available and since only the product of an even number of odd terms yields an even term, all terms contributing to H_{bd} do necessarily contain an even number of momentum operators, which can always be

TABLE I. Comparison of FW, DKH, and BSS decoupling transformation schemes. The FW procedure is completely ill defined and must not be used in any order in $1/c$. The regular DKH scheme yields analytic and well-defined expressions for any order in V, and the BSS protocol is a purely numerical realization of the infinite-order DKH scheme. See text for further explanations.

	Foldy–Wouthuysen (FW)	Douglas-Kroll-Hess (DKH)	Barysz–Sadlej–Snijders (BSS)
Initial transformation	$1/c$ expanded fpFW transformation U_0, see Eqs. (12) and (15)	Analytic, i.e., closed-form fpFW transformation $U_0 = A_p(1+\beta R_p)$, see Eq. (22)	
Methods of further decoupling	Sequence of subsequent unitary transformations U_i $(i=1,2,\dots)$		One-step diagonalization $U_1 = U_1(R)$, see Eq. (34)
Decoupling of different forms	$U_i = \sum_{k=0}^{\infty} a_{i,k} W_{[i]}^k$	$U_i = \sum_{k=0}^{\infty} a_{i,k} W_i^k$	$U_1 = \begin{pmatrix} \dfrac{1}{\sqrt{1+R^\dagger R}} & \dfrac{1}{\sqrt{1+R^\dagger R}} R^\dagger \\ \dfrac{-1}{\sqrt{1+RR^\dagger}} R & \dfrac{1}{\sqrt{1+RR^\dagger}} \end{pmatrix}$
Hamiltonian	Analytic (closed-form) expressions $h_+ = \sum_{k=-1}^{\infty} \mathcal{E}_{[2k]+}$	$h_+ = \sum_{k=0}^{\infty} \mathcal{E}_{k+}$	Numeric representation $h_+ = h_+(R)$
Formal expansion parameter	$1/c$	V	$\cdots$
True expansion parameter	$x = p/mc$	$\tilde{V}$	$\cdots$
Convergence	$\cdots$	Controlled via order parameter	Controlled by convergence of iterative solver
Finite-order approximation	Ill defined	Regular	$\cdots$
Formal infinite order	Ill defined	Exact, i.e., untruncated block-diagonal Hamiltonians are equivalent: $H_{bd} = H_{\mathrm{DKH}\infty} = H_{\mathrm{BSS}}$	
Infinite order in practice	$\cdots$	Matrix multiplications up to desired order	Iterative solution for operator R

grouped to p^2 or $\boldsymbol{\sigma}\cdot\mathbf{p} V \boldsymbol{\sigma}\cdot\mathbf{p}$ expressions. In this sense all even terms $\mathcal{E}_k$ are thus functions of p^2 rather than $\mathbf{p}$. This situation is graphically illustrated in Fig. 4.

As a consequence, application of computationally demanding large basis sets for the small components—due to the kinetic balance requirements—can be avoided for the evaluation of decoupled Hamiltonians, which is the key to their efficient implementation. No reference has to be made to odd terms with rectangular matrix representations being expensive to evaluate.

By restriction on the upper left block of the Hamiltonian, two-component methods do thus achieve enormous computational savings as compared to four-component methods, since not only the number of integrals to be evaluated is
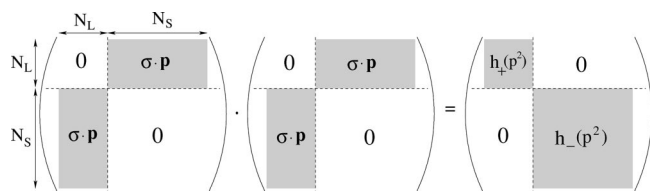

FIG. 4. Schematic diagram of the matrix representations of even and odd operators. Odd blocks depend always on the linear momentum variable and even blocks thus only on p^2. The upper left block h_+ of the Hamiltonian describes the positive-energy solutions and features a $(N_L \times N_L)$-dimensional matrix representation, whereas the negative-energy block h_- is much more demanding to compute.

decreased significantly, but also the size of the matrix representations of any operator encountered is reduced to a dimensionality of $(N_L \times N_L)$. As a matter of fact, the number of integrals and the size of the matrix representations within two-component frameworks is the same as for nonrelativistic approaches.

Besides these computational benefits two-component methods do also feature an essential conceptual advantage over four-component setups. Due to the block diagonalization of the Hamiltonian, no interactions between the positive-energy states representing electrons and the unphysical negative-energy states may occur. As a consequence, the electronic Hamiltonian h_+ is bounded from below by construction and no variational collapse can occur. Its spectrum comprises exactly the same positive-energy eigenvalues as the original Dirac Hamiltonian, i.e., no approximations are introduced by the block-diagonalization procedure. Furthermore, within two-component approaches scalar-relativistic and spin-dependent terms can always be separated using Dirac's relation

$$(\boldsymbol{\sigma}\cdot\mathbf{p})\omega(\boldsymbol{\sigma}\cdot\mathbf{p}) = \mathbf{p}\cdot\omega\mathbf{p} + i\boldsymbol{\sigma}\cdot(\mathbf{p}\times\omega\mathbf{p}), \qquad (37)$$

for Pauli spin matrices, where ω may be any scalar operator, e.g., the external potential V.

Due to the structure of the zeroth-order term $\mathcal{E}_0$ given by Eq. (24), which contains the square root of the differential operator $\mathbf{p}^2$, any decoupled Hamiltonian is highly nonlocal

in coordinate space. Since the Taylor expansion of E_p is completely ill defined (cf. Sec. II B), $\mathcal{E}_0$ has necessarily to be evaluated within a representation where the momentum operator $\mathbf{p}$ acts as a simple multiplicative operator. This may either be directly in momentum space or in the space spanned by the eigenfunctions of the nonrelativistic kinetic energy operator T, as it is done in the efficient implementation of the DKH transformation by Hess,[13] which has also been transferred to the BSS approach by Barysz and Sadlej.[28]

IV. EXACT DECOUPLING WITHIN THE DKH SCHEME

Inspired by the impressive success of the BSS scheme and its outstanding accuracy for one-electron systems,[28] it would be highly desirable to devise an extension of the DKH method up to arbitrarily high orders. For this purpose one could try to establish an iterative numerical scheme, which might be repeated until the exact result is obtained. The objective of this section is to demonstrate that such a purely numerical iterative DKH protocol cannot be realized within one- or two-component quantum chemistry packages in principle and one has thus to break new ground to achieve DKH transformations to arbitrary order.

Let us assume the intermediate Hamiltonian of the DKH procedure after n transformation steps were given. Due to the $(2n+1)$ rule[21,23] it may be written as

$$H_n = \sum_{k=0}^{2n-1} \mathcal{E}_k + \sum_{k=2n}^{\infty} \mathcal{E}_k^{(n)} + \sum_{k=n}^{\infty} \mathcal{O}_k^{(n)}, \tag{38}$$

where each term is classified according to its order in the external potential V and its even or odd structure. The next unitary transformation U_n is established by the odd operator W_n, which is of nth order in V, via a general power-series ansatz of the type given by Eq. (12) and yields

$$H_{n+1} = U_n H_n U_n^\dagger$$

$$= \sum_{k=0}^{2n+1} \mathcal{E}_k + \sum_{k=2n+2}^{\infty} \mathcal{E}_k^{(n+1)}$$

$$+ \underbrace{\mathcal{O}_n^{(n)} + a_{n,0} a_{n,1} [W_n, \mathcal{E}_0]}_{\mathcal{O}_n^{(n+1)}}$$

$$+ \sum_{k=n+1}^{\infty} \mathcal{O}_k^{(n+1)}. \tag{39}$$

W_n is determined uniquely by the requirement that it has to account for the elimination of the term $\mathcal{O}_n^{(n+1)}$. Being an nth order integral operator, this is guaranteed if the kernel of W_n is given by

$$W_n(p_0, p_1, \ldots, p_n) = \frac{a_{n,0}}{a_{n,1}} \beta \frac{\mathcal{O}_n^{(n)}(p_0, p_1, \ldots, p_n)}{E_{p_0} + E_{p_n}}. \tag{40}$$

If the matrix representation of this kernel were known, one could immediately employ it to evaluate all terms of the Hamiltonian H_{n+1}. This procedure could then be repeated until the resulting Hamiltonian is block diagonal up to the desired order in V.

It is, however, exactly this step which is not feasible within any two- or one-component implementation, since the matrix representations of odd operators are not calculated in standard quantum chemistry program packages. As discussed in Sec. III, their evaluation would inevitably require the introduction of basis sets analogously to those for the small components of the molecular spinors, which are approximately twice as large as the large component basis and have to contain exponents for higher angular momentum functions. This would not only blow up the computational costs of the calculations considerably, but even destroy the elegant framework of one- or two-component quasirelativistic methods, which could no longer be easily embedded in any nonrelativistic computer program.

The evaluation of higher-order DKH approximations does thus necessarily require new algorithmic principles, which accomplish the evaluation of the upper left block h_+ of the block-diagonal Hamiltonian H_{bd}, i.e., the even terms $\mathcal{E}_k$ without any reference to the small component. As shown above, such a new algorithmic scheme can in principle not work numerically, but has to determine the individual terms of the DKH Hamiltonian purely algebraicly, i.e., by a symbolical evaluation of the corresponding unitary transformations. Due to the increasing complexity and number of the higher-order terms, their determination, however, is only possible if the algebraic manipulations can be executed automatically by a suitable parser routine yielding analytic formulas for each order $\mathcal{E}_k$. Subsequently, this parser should be able to translate the resulting closed-form operator expressions into corresponding matrix multiplications arising in a basis set approach. Since the details of this procedure cannot be described in sufficient detail in this work, the construction and implementation of such a parser routine as well as its numerical performance will be presented in paper II of this series.[29]

V. CONCLUSION

The main objective of this work was to show that the DKH protocol is the only valid transformation technique for decoupling the Dirac Hamiltonian. For this purpose, all possible transformation schemes for the exact decoupling of the Dirac Hamiltonian have been investigated. By application of the most general ansatz for the unitary transformations we have shown that all schemes, independent of their special strategy, have necessarily to start with an initial fpFW step. Neither this initial nor any subsequent transformation must be expanded in $1/c$ around the nonrelativistic limit $p=0$, since any such series expansion is only valid for small momenta, i.e., $x=p/mc<1$. Meeting singularities within the complex plane at $x=\pm i$, these series expansions will consequently be ill defined for any larger value of the momentum.

Every series expansion of the block-diagonal Hamiltonian H_{bd} has thus necessarily to employ a different expansion parameter, and the damped external potential $\tilde{V}$ is the only acceptable choice. This naturally leads to the DKH protocol, i.e., the stepwise construction of the decoupled Hamiltonian by a sequence of unitary transformations, which for-

mally results in an expansion of this Hamiltonian in terms of ascending and well-defined orders in the potential V.

The main results of this study may be summarized as follows.

(a) The (well-known) expression of the fpFW transformation U_0 is rederived employing the framework of the generalized DKH transformation, i.e., the most general parametrization of unitary matrices.

(b) The uniqueness of the fpFW transformation is shown by a discussion of the radius of convergence of a Taylor series expansion of the analytical, block-diagonal, fpFW-transformed Dirac Hamiltonian H_1 [cf. Eq. (24)] in $1/c$. All other parametrizations of U_0, i.e., all other choices for the expansion coefficients $a_{0,k}$ yield ill-defined and singular expressions.

(c) The initial transformation U_0 of the DK procedure has necessarily and uniquely to be chosen as the fpFW transformation.

(d) The failure of higher-order FW transformations is shown to be related to the kinetic term $\mathcal{E}_0$ rather than being due to properties of the external potential V. This ill-defined behavior of higher-order FW transformations occurs for any choice of V and not only for Coulomb potentials. [Due to the $1/r$ singularity of the Coulomb potential V_C, the FW transformation of V_C does cause additional problems for $r \to 0$ as it is known for a long time (see, e.g., the work of Kutzelnigg[42,43]).]

(e) The DKH protocol is the unique transformation scheme for decoupling the Dirac Hamiltonian, and yields regular and well-defined expressions up to any arbitrary order in V.

(f) Consequently, the Barysz–Sadlej–Snijders approach can be discussed in this frame and represents a fully numerical variant of the infinite-order DKH scheme.

(g) The true expansion parameter of the DKH protocol is the damped potential $\widetilde{V}$, which is strongly suppressed by large energy denominators. This guarantees excellent convergence behavior of the DKH scheme for all values of the nuclear charge Z.

(h) It has been discussed, why the individual terms of the DKH Hamiltonians can only be determined purely algebraically, i.e., by symbolic evaluation of the unitary transformations.

We have thus demonstrated that the contributions of each individual order of the infinite-order DKH Hamiltonian are only to be analyzed if an automated analytical derivation of higher-order terms can be accomplished. This is due to the fact that absolutely no reference to a small component is to be made in a two-component framework. An algorithm for infinite-order DKH calculations (DKH∞) will be described in paper II.[29]

ACKNOWLEDGMENTS

[1] *Relativistic Electronic Structure Theory. Part 1. Fundamentals*, edited by P. Schwerdtfeger (Elsevier, Amsterdam, 2002).

[2] *Relativistic Effects in Heavy-Element Chemistry and Physics*, edited by B. A. Hess (Wiley, Chichester, 2003).

[3] P. A. M. Dirac, Proc. R. Soc., London, Ser. B **A117**, 610 (1928).

[4] J. L. Heully, I. Lindgren, E. Lindroth, S. Lundquist, and A. M. Mårtensson-Pendrill, J. Phys. B **19**, 2799 (1986).

[5] D. Andrae, Phys. Rev. B **336**, 413 (2000).

[6] B. Thaller, in Ref. 1, pp. 23–106.

[7] M. Reiher and J. Hinze, in Ref. 2, pp. 61–88.

[8] M. Reiher and B. A. Hess, in *Modern Methods and Algorithms of Quantum Chemistry*, NIC Series, Vol. 1 (John von Neumann Institute for Computing, Jülich, 2000), pp. 451–477.

[9] W. Kutzelnigg, Phys. Scr. **36**, 416 (1987).

[10] W. Kutzelnigg, Chem. Phys. **225**, 203 (1997).

[11] L. L. Foldy and S. A. Wouthuysen, Phys. Rev. **78**, 29 (1950).

[12] M. Douglas and N. M. Kroll, Ann. Phys. **82**, 89 (1974).

[13] B. A. Hess, Phys. Rev. A **33**, 3742 (1986).

[14] M. Barysz, A. J. Sadlej, and J. G. Snijders, Int. J. Quantum Chem. **65**, 225 (1997).

[15] M. Barysz and A. J. Sadlej, J. Mol. Struct.: THEOCHEM **573**, 181 (2001).

[16] M. Barysz, J. Chem. Phys. **114**, 9315 (2001).

[17] G. Jansen and B. A. Hess, Phys. Rev. A **39**, 6016 (1989).

[18] R. Samzow, B. A. Hess, and G. Jansen, J. Chem. Phys. **96**, 1227 (1992).

[19] N. J. M. Geipel and B. A. Hess, Chem. Phys. Lett. **273**, 62 (1997).

[20] B. A. Hess, Ber. Bunsenges. **101**, 1 (1997).

[21] T. Nakajima and K. Hirao, J. Chem. Phys. **113**, 7786 (2000).

[22] A. Wolf, M. Reiher, and B. A. Hess, in Ref. 1, pp. 622–663.

[23] A. Wolf, M. Reiher, and B. A. Hess, J. Chem. Phys. **117**, 9215 (2002).

[24] A. Wolf, M. Reiher, and B. A. Hess, in *Recent Advances in Relativistic Molecular Theory*, edited by K. Hirao and Y. Ishikawa (World Scientific, Singapore, 2004), pp. 137–190.

[25] A. Wolf, M. Reiher, and B. A. Hess, J. Chem. Phys. **120**, 8624 (2004).

[26] I. Lindgren and J. Morrison, *Atomic Many-Body Theory* (Springer, Berlin, 1982).

[27] Ph. Durand, Phys. Rev. A **28**, 3184 (1983).

[28] M. Barysz and A. J. Sadlej, J. Chem. Phys. **116**, 2696 (2002).

[29] M. Reiher and A. Wolf (unpublished).

[30] C. van Wüllen, J. Chem. Phys. **120**, 7307 (2004).

[31] A. G. Nikitin, J. Phys. A **31**, 3297 (1998).

[32] C. Itzykson and J.-B. Zuber, *Quantum Field Theory* (McGraw-Hill, New York, 1980).

[33] E. Merzbacher, *Quantum Mechanics*, 3rd ed. (Wiley, New York, 1998).

[34] H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms* (Plenum, New York, 1977).

[35] J. D. Bjorken and S. D. Drell, *Relativistic Quantum Mechanics* (McGraw-Hill, New York, 1964).

[36] K.-H. Yang, J. Phys. A **15**, 437 (1982).

[37] H. W. Fearing, G. I. Poulis, and S. Scherer, Nucl. Phys. A **570**, 657 (1994).

[38] H. W. Fearing, G. I. Poulis, and S. Scherer, Nucl. Phys. A **570**, 686 (1994).

[39] J. P. Costella and B. H. J. McKellar, Am. J. Phys. **63**, 1119 (1995).

[40] A. J. Silenko, J. Math. Phys. **44**, 2952 (2003).

[41] B. A. Hess, Phys. Rev. A **32**, 756 (1985).

[42] W. Kutzelnigg, Z. Phys. D: At., Mol. Clusters **11**, 15 (1989).

[43] W. Kutzelnigg, Z. Phys. D: At., Mol. Clusters **15**, 27 (1990).