REGULAR ARTICLE

Markus Reiher

Douglas–Kroll–Hess Theory: a relativistic electrons-only theory for chemistry

Abstract A unitary transformation allows to separate (block-diagonalize) the Dirac Hamiltonian into two parts: one part solely describes electrons, while the other gives rise to negative-energy states, which are the so-called positronic states. The block-diagonal form of the Hamiltonian no longer accounts for the coupling of both kinds of states. The positive-energy ('electrons-only') part can serve as a 'fully' relativistic electrons-only theory, which can be understood as a rigorous basis for chemistry. Recent developments of the Douglas–Kroll–Hess (DKH) method allowed to derive a sequence of expressions, which approximate this electrons-only Hamiltonian up to arbitrary-order. While all previous work focused on the numerical stability and accuracy of these arbitrary-order DKH Hamiltonians, conceptual issues and paradoxa of the method were mostly left aside. In this work, the conceptual side of DKH theory is revisited in order to identify essential aspects of the theory to be distinguished from purely computational consideration.

1 Introduction

This year, we celebrate the centenary of Einstein's theory of special relativity. The most important consequence of this theory is that any mathematical description of physical phenomena has to obey the principles of special relativity [1]: apart from the constant speed of light in any inertial frame of reference, the fundamental equations have to take the same form in any inertial frame of reference; in other words, they must be covariant under Lorentz transformations.

However, some nonrelativistic approaches to physical problems may be little affected by the so-called relativistic effects, i.e., their numerical results may not deviate significantly from a correct relativistic description. This, for instance, is the case with Newtonian mechanics, when applied

to the motion of objects in daily life. Nevertheless, many physical situations exist where deviations of a nonrelativistic description from numerical results of a relativistic formulation or from experiment cannot be neglected. In such cases, it is desirable to capture the numerical relativistic effect with a 'quasi-relativistic' theory, which is as consistent as possible with fundamental physical principles. It was realized about 30 years ago that such a situation arises in the chemistry of heavy elements (see [2–6]). For the calculation of the numerical relativistic effect in molecular systems, relativistic quantum chemistry [7–10] has been developed since then. Thus, 100 years after Einstein's annus mirabilis, relativistic quantum chemistry has matured to a well-developed theory whose principles shall be revisited in this work.

A fundamental theory for chemistry should be a relativistically correct quantum mechanical all-electron theory, which also enables us to do actual calculations. A truly relativistic theory for the many-electron systems in external fields, such as atoms and molecules, which fulfills these requirements has never been devised and it may be questioned whether it will be derived. Of course, physicists would object to this saying that quantum electrodynamics [11–13] represents such a theory. One may reply in various ways to such an objection but what is most important for chemistry is that actual calculations for arbitrary systems are unfeasible. Instead, a formulation based on Dirac's theory of a (single) electron [14–16] turned out to be efficient and feasible at the same time. In general, the deviation of results obtained in a first-quantized Dirac-based all-electron framework from quantum electrodynamics, or more precisely, from calculations which model certain quantum electrodynamical effects, turned out to be negligibly small [17].

This formulation uses Dirac's one-electron Hamiltonian,

$$h_D = c\boldsymbol{\alpha} \cdot \boldsymbol{p} + (\beta - 1)c^2 + V_{nuc}, \tag{1}$$

where the Dirac matrices may be represented as usual, i.e., $\boldsymbol{\alpha}$ is a 3-vector containing Pauli's spin matrices on the off-diagonal, β is a diagonal matrix with entries $(1, 1, -1, -1)$,

M. Reiher
Institut für Physikalische Chemie,
Friedrich-Schiller-Universität Jena,
Helmholtzweg 4, 07743 Jena, Germany
E-mail: markus.reiher@uni-jena.de

c is the speed of light, and V_{nuc} the external potential [see Eq. 3 below; Hartree atomic units are used throughout]. It is interesting to note that the explicit form of the matrices is not important. They do have to fulfill certain relations in order to make the basic equation of quantum mechanics, i.e., the equation of motion $h_D \psi = i \partial \psi / \partial t$, Lorentz covariant [18]. These relations require $\boldsymbol{\alpha}$ and β to be at least 4×4 matrices. Thus, the corresponding wave function is a 4-vector, a so-called 4-spinor. This 4-spinor ψ contains four functions (components) to characterize the quantum mechanical state of an electron. The lower two components ψ_S are usually small compared to the upper two components ψ_L for nuclear charge numbers less than say $Z = 100$ (both ψ_S and ψ_L, respectively, are two-component vectors: 2-spinors); for this reason the former are called the small and the latter the large components of the spinor. It is important to note that the off-diagonal terms of the three 4×4 matrices in $\boldsymbol{\alpha}$ couple the large and small components, respectively.

The Dirac Hamiltonian is now applied as a substitute for the Schrödinger one-electron operator in the electronic Hamiltonian for N electrons and M atomic nuclei in first-quantized form,

$$H_{\text{el,D}} = \sum_{i=1}^{N} h_{D,i} + \sum_{i<j}^{N} g_{ij}, \tag{2}$$

with the electron-nucleus (in $h_{D,i}$) and electron–electron interaction operators,

$$V_{\text{nuc},i} = \sum_{J=1}^{M} -Z_J / |\boldsymbol{r}_i - \boldsymbol{R}_J|, \tag{3}$$

$$g_{ij} = 1 / |\boldsymbol{r}_i - \boldsymbol{r}_j|, \tag{4}$$

for the instantaneous point-charge Coulomb interactions of nuclei and electrons (the nucleus–nucleus interaction is left aside for brevity). One may introduce the so-called relativistic correction terms for these instantaneous interactions (usually the Breit operator is used as a correction for the electron–electron interaction), but this does not add any new insight into the conceptual problems to be revisited here.

Before we proceed it appears to be appropriate to state the obvious and to remove common misconceptions. The four-component theoretical foundations indicated above are formulated for a common absolute time valid for all particles in contrast with the principles of special relativity. In addition, one applies the Born–Oppenheimer approximation for the separation of nuclear and electronic degrees of freedom. Moreover, the formulation remains first-quantized even if the language of second-quantization is applied to elegantly rephrase the correlation problem for many-electron wave functions. Needless to say that the four-component formulation covers the whole periodic table of elements though notable numerical effects are visible only for large nuclear charge numbers (i.e., for heavy elements) or deep attractive electron–nucleus potentials, respectively. A typical misperception of this relativistic formulation of quantum chemistry is that the small components of the spinor ψ describe a positron (or a positronic state), while the two upper large components an electron: electrons as well as the so-called positronic

states are each described by sets of all *four* functions. This is also reflected in the calculation of expectation values: the kinetic energy of an electron and properties associated with the application of an external magnetic field require large and small component functions at the same time.

2 Four-component and two-component approaches

Quantum chemical methods based on the four-component many-electron Hamiltonian (Eq. 2) have been successfully devised by many groups (see e.g., Ref. [19] for a review of methods, programs, and applications). Currently, the most successful general purpose program package for highly sophisticated four-component calculations is undoubtedly DIRAC [20]. Unfortunately, the four-component methods are neither truly relativistically invariant — though often termed 'fully' relativistic — nor computationally competitive compared to standard Schrödinger-based quantum chemical methods. However, the latter fact cannot be taken as an objection against the four-component methods, it is merely a statement to indicate that the study of molecules with, say, a hundred atoms is hardly possible, at least for the moment. It should be emphasized that the accuracy of results obtained with four-component methods for atoms and molecules is very remarkable (the reader is referred to the excellent and comprehensive bibliography collected by Pyykkö [21–24] for examples). It is the intention of this work to leave computational aspects aside and to concentrate on conceptual issues. We should thus elaborate on why the four-component methods are not truly relativistic as already mentioned. The four-component electronic Hamiltonian in Eq. 2 simply does not possess the necessary invariance properties. All we can aim for with the four-component methods is the approximation of the relativistically correct numerical values of some physical quantities. But if we do accept that actual 'fully' relativistic, i.e., four-component calculations are based on equations which do not fulfill the basic requirement of special relativity to be covariant under Lorentz transformations, we may devise other methods with merits not present in four-component theory. Consequently, we require any *other* 'fully' relativistic approach for the quantum mechanical description of atoms and molecules to be as reliable as the four-component theory. However, this other approach may improve on issues that can be raised against the four-component methods for conceptual reasons.

The major drawback of the four-component methods emerges from a conceptual problem. This is the occurrence of the positronic states, which are states at negative energies separated by about $-2c^2$ from the electronic bound states. The connotation of 'positronic' suggests that there is a physical meaning inherent in these states. The opposite, however, is the case: the first-quantized Dirac Hamiltonians in Eq. 2 restrict the quantum mechanical treatment to typical chemical systems in a certain energy range; particle creation processes need not to be considered and the number of particles
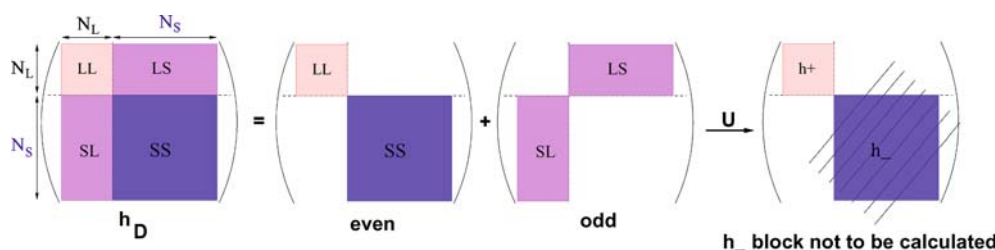
Fig. 1 Schematic representation of the Dirac Hamiltonian in a given basis with N_L basis functions for ψ_L and N_S basis functions for the expansion of ψ_S. Due to the so-called kinetic balance condition [33, 70, 71], a Gaussian-type basis set for the small component ψ_S is significantly larger, which represents a computational bottleneck for four-component methods in self-consistent field calculations as well as integral transformations. The scheme also demonstrates how the Dirac Hamiltonian may be decomposed into a sum of even, i.e., block-diagonal, and odd, i.e., off-diagonal operators

does not change.[1] Moreover, positrons would feature positive energies, but the positronic states have negative energies. Last but not least, in actual calculations the negative-energy states are usually represented by local Gaussian-type basis functions, which do not adequately describe such continuum states (remember that the external electron–nucleus potential in the electronic Hamiltonian of Eq. 2 is only attractive for electronic (bound) states because of the negative sign in Eq. 3 stemming from the negative charge of the electrons). Thus, the positronic states possess *no* physical meaning and are troublesome artifacts in four-component relativistic quantum chemistry.

Now, we have introduced an accurate (four-component) many-electron theory based on Dirac's Hamiltonian plagued with a conceptual inconsistency. It would be desirable to replace the first-quantized four-component theory by an approach of the same accuracy, but without making any reference to negative-energy states. Two different kinds of such 'reduction' approaches have been developed: elimination and transformation techniques (see [25–27] for recent reviews). Since the elimination schemes are not flawless as the operators turn out to be energy dependent (otherwise the lowest order approximation is used), they appear to be less elegant from a formal point of view than the transformation schemes although they are numerically very efficient in practice (see e.g., the popular ZORA method [28–30]).

The transformation techniques aim at a block-diagonalization of the Dirac Hamiltonian (see also Fig. 1 for a graphical representation),

$$h_{\mathrm{bd}} = U h_{\mathrm{D}} U^\dagger = \begin{pmatrix} h_+ & 0 \\ 0 & h_- \end{pmatrix}, \qquad (5)$$

in order to derive an upper block Hamiltonian h_+ of reduced dimension valid for the relativistic description of electrons

only (the lower block h_-, which is responsible for the negative energy states, is then simply neglected). The operator h_+ is sometimes misleadingly also called a *no-pair* operator (strictly speaking, the first-quantized four-component Dirac Hamiltonian is also a no-pair operator).

For the sake of simplicity, the unitary transformation in Eq. 5 has been applied to the *one-electron* Dirac Hamiltonian in an external field. The generalization to arbitrary particle numbers may be derived in terms of the usual tensor product formulation for the construction of the many-particle Hamiltonian (see [26] for details).

The linear transformation U is chosen to be unitary in order to preserve the normalization of states as well as the spectrum of the Hamiltonian. The transformed four-component wave function is then obtained as,

$$\phi = U\psi$$
$$= U \begin{pmatrix} \psi_L \\ \psi_S \end{pmatrix} = \begin{pmatrix} U_{LL}\psi_L + U_{LS}\psi_S \\ U_{SL}\psi_L + U_{SS}\psi_S \end{pmatrix} = \begin{pmatrix} \phi_L \\ \phi_S \end{pmatrix}, \qquad (6)$$

with $\phi_S = \mathbf{0}$ for electronic states. The basic quantum mechanical equation of motion for an relativistic electron in this no-pair formulation is thus obtained as,

$$h_+\phi_L = i\frac{\partial \phi_L}{\partial t}, \qquad (7)$$

or in stationary form for h_+ independent of t,

$$h_+\phi_L = E_+\phi_L. \qquad (8)$$

Until the work of Barysz and Sadlej in 2002 [31], these techniques have been considered approximate and have thus always been termed 'quasi'-relativistic in order to make a distinction from the 'fully' relativistic four-component reference. These authors achieved for the very first time the *direct* numerical calculation of the upper block Hamiltonian h_+ and the transformed upper component ϕ_L without any reference to a small component. *Numerical* means that the resulting one-electron upper block Hamiltonian h_+ was obtained in a matrix form defined with respect to a pre-defined basis set, and its entries can hardly be traced back to well-defined analytic expressions for matrix elements as these numerical values resulted from an iterative numerical scheme. Barysz and Sadlej have thus achieved exact decoupling of upper and lower components of the spinor in a purely numerical scheme.

---

[1] It should be emphasized that the language of second-quantization is also used in the four-component theory as it is used in nonrelativistic quantum chemistry. This use in quantum chemistry is different from physics as it merely serves to elegantly rephrase the calculation of energy expectation values with correlated wave functions. The rules for creation and annihilation operators, which create or annihilate orbitals in electronic configurations, implement the Slater–Condon rules for the evaluation of matrix elements over N-particle functions. In physics, however, second quantization is used for the proper quantization of classical fields.

In order to fully appreciate this achievement,[2] one has to recall that a true two-component method must never make any reference to the small component of the spinor, i.e., the lower block of the Hamiltonian h_-, which would account for the positronic states, is never calculated, neither are off-diagonal coupling matrices calculated. Thus, also the off-diagonal parts U_{LS} and U_{SL} of the unitary matrix are not available in a truely two-component scheme and may only be obtained by evoking a four-component framework, which we explicitly try to circumvent here.

We should not forget to mention that (approximate) no-pair operators have been studied by many groups for quite some time (see e.g., the extensive account by Kutzelnigg in Ref. [33]). The new developments aimed at *exact* decoupling without reference to a small component. The numerical achievement of Barysz and Sadlej [31], which has been called the infinite-order two-component (IOTC) method related to earlier work by Barysz et al. [34], raised the question whether an analytic expansion to infinite-order is possible at all.[3] This issue will be discussed after Sect. 3 where the *generalized* Douglas–Kroll–Hess (DKH) method [35] is introduced.

The most important transformation technique is the so-called the DKH method [36, 37], to which the work of Barysz and Sadlej has a close relation [38]. In the next Section, we introduce the DKH transformation technique in the most general way [35]. The unique character [38] of this scheme is explained and it is then argued that this particular method provides a 'fully' relativistic electrons-only theory free of the positronic-state inconsistencies.

3 Essential DKH theory

Although the unitary matrix U in Eq. 5 can formally be given in closed-form [32], the elements in this matrix U are obtained from a rather complicated equation, which needs to be solved iteratively. It is this equation which has been solved by Barysz and Sadlej [31] to directly obtain a numerical representation of h_+. (We do not want to go into details here but only mention that this direct approach involves an important trick invented by Hess [39], namely the use of a p^2-basis, which we need to discuss later.)

Since the closed-form solution for U allows us to derive an expression for h_+, which cannot readily be evaluated as the elements of U occurring in this expression are not readily calculated, it turned out to be useful to think of a consecutive decoupling in terms of a sequence of unitary transforma-

tions $U = \cdots U_4 U_3 U_2 U_1 U_0$, which block-diagonalizes h_D stepwise,

$$h_{\mathrm{bd}} = \cdots U_4 U_3 U_2 U_1 U_0 \ h_D \ U_0^\dagger U_1^\dagger U_2^\dagger U_3^\dagger U_4^\dagger \cdots . \tag{9}$$

Each unitary matrix U_i shall then be chosen such as to diminish the off-diagonal terms, which are also called *odd terms*, order by order. This order-by-order block diagonalization assumes the existence of an expansion of the block-diagonal Hamiltonian in terms of a suitable expansion parameter, which allows to identify block-diagonal (so-called even) operators, $\mathcal{E}_k$, of given order k. The formal expansion parameter of the DKH decoupling procedure is the external potential V. (Note that we have skipped the subscript '*nuc*' since other scalar potential contributions may be included as well.) Once all these even operators are added we obtain the block-diagonal Hamiltonian,

$$h_{\mathrm{bd}} = \sum_{k=0}^{\infty} \mathcal{E}_k. \tag{10}$$

A partially transformed Hamiltonian h_{pt} obtained after the m-th unitary transformation may be written as

$$h_{\mathrm{pt}} = \sum_{k=0}^{2m+1} \mathcal{E}_k + \sum_{k=2m+2}^{\infty} \mathcal{E}_k^{(m+1)} + \sum_{k=m+1}^{\infty} \mathcal{O}_k^{(m+1)}, \tag{11}$$

where we have already anticipated the so-called $(2m + 1)$-rule, which states that each unitary transformation yields two orders k of the expansion parameter V in Eq. 10. Equation 11 still contains off-diagonal (odd) operators $\mathcal{O}_k$, which need to be removed by subsequent unitary transformations. However, this can only be successful if the magnitude of the $\mathcal{O}_k$ is diminished as k gets larger. Note that we have also anticipated in the superscripts that the final form of the $\mathcal{E}_k$ with $k \geq 2m + 2$ is not reached after the m-th unitary transformation. We have to stipulate that all $\mathcal{E}_k$ with $k \leq 2m + 1$ remain untouched by the following unitary transformations. Otherwise, the low order even terms $\mathcal{E}_0$, $\mathcal{E}_1$, $\mathcal{E}_2$, ... would be changed after every new transformation step. This is not desirable as a useful method should only add higher-order even terms to the low-order even terms, which are to remain untouched and contribute most to h_{bd}.

In order to meet this requirement, it is necessary to expand each unitary matrix U_m into a power series,

$$U_m = \sum_{j=0}^{\infty} a_{m,j} W_m^j = \mathbf{1} + \sum_{j=1}^{\infty} a_{m,j} W_m^j, \tag{12}$$

where $a_{m,0} = 1$ was chosen and each term of the series is classified uniquely according to the order $k = m \times j$ with respect to the *same* expansion parameter V as used in Eq. 10. The analytic expansion of U_m in Eq. 12 is the most general form of what may be called 'a parametrization of a unitary matrix' with W_m being the parameter. This parameter can be chosen freely as long as it is anti-hermitian,

$$W_m^\dagger = W_m^{T\star} = -W_m, \tag{13}$$

---

[2] It should be mentioned that the basic equations were known for some time [32] but assumed to be of purely formal character not useful in actual calculations.

[3] Here, we again meet the implementability and feasibility requirement mentioned earlier: it is, of course, possible to write some sort of series expansion on a piece of paper, in which terms are added up to infinity (compare also the elimination techniques) but — even if such an expansion converges — this does not imply that one can carry out actual calculations to an order sufficiently high to call the result exact (within machine precision).

so that $U^\dagger U = \mathbf{1}$ can be fulfilled and also because we need to be able to readily write the adjoint transformation,

$$U_m^\dagger = \sum_{j=0}^{\infty} (-1)^j a_{m,j} W_m^j = \mathbf{1} + \sum_{j=1}^{\infty} (-1)^j a_{m,j} W_m^j, \qquad (14)$$

for the evaluation of Eq. 10. The use of these most general parametrizations of the unitary matrices has led to the term *generalized* DKH transformation [35] as the resulting Hamiltonians make no reference to a specific parametrization of the unitary matrix (Douglas and Kroll [36] introduced the so-called square-root parametrization, which turned out to be no essential part of the theory [35]; see Sect. 4.3).

To fulfill the unitary condition for each U_i in Eq. 12, we have to transfer the unitarity requirement onto the expansion coefficients (see [35] for details). The identity operator $\mathbf{1}$ in Eq. 12 takes care to leave the low-order even terms $\mathcal{E}_k$ with $k \leq 2m+1$ untouched. Since the order of the term W_m^j in the parametrization expansion for U_m is $m \times j$, each product with even terms $\mathcal{E}_k$ resulting from Eq. 9 is of order $k \times (m \times j)$, which is always larger than $2m + 1$ if the product is even — otherwise the term may be of lower order but odd and will thus be eliminated in this or in a later unitary transformation step. It is important to understand that we obtained odd operators of comparatively low order in V while we derived already the final form of higher-order even operators due to the $(2m + 1)$-rule. Since these higher-order even operators do not change in subsequent transformation steps, we have already obtained the higher-order DKH Hamiltonian and can simply neglect the lower-order odd terms *without* making an approximation.

The step-by-step elimination of odd terms $\mathcal{O}_k$ is essential to the method but how are the odd terms eliminated order by order? Here, we make use of the freedom in the choice of the anti-hermitian parameter W_m. Once the expansions (Eq. 12) are inserted in Eq. 10, we simply need to require,

$$[W_m, \mathcal{E}_0] - \frac{a_{m,0}}{a_{m,1}} \mathcal{O}_m^{(m)} = 0, \qquad (15)$$

in order to eliminate the odd term of order m (the derivation is not complicated but a bit tedious). This equation thus determines W_m — but it also requires W_m to be an *odd* operator. Note that odd operators cannot have a representation in a truly two-component framework like DKH as odd means off-diagonal, i.e., coupling between the upper and lower components of the spinor — but there must not be a small component.

For the derivation of the DKH Hamiltonian, it is *mandatory* [38] to carry out U_0 explicitly as a free-particle Foldy–Wouthuysen transformation [36,40],

$$U_0 h_D U_0^\dagger = \underbrace{\beta E_p - c^2}_{\mathcal{E}_0} + \underbrace{A_p(V + R_p V R_p)A_p}_{\mathcal{E}_1}$$
$$+ \underbrace{\beta A_p[R_p, V]A_p}_{\mathcal{O}_1} \equiv H_1, \qquad (16)$$

using the auxiliary quantities

$$E_p = \sqrt{\mathbf{p}^2 c^2 + c^4}, \qquad A_p = \sqrt{\frac{E_p + c^2}{2E_p}},$$
$$R_p = \frac{c\boldsymbol{\alpha}\cdot\mathbf{p}}{E_p + c^2} \qquad (17)$$

in order to obtain the sequence of two even and one odd term on the right hand side of Eq. 16. All subsequent unitary transformations $U_{i \geq 1}$ can be carried out in an automated manner, which makes the symbolic derivation of DKH Hamiltonians of arbitrary order feasible [41].

The parametrization given by Eq. 12 also takes care of the fact that *no* low-order odd terms may arise: once an odd term of a given order has been eliminated, an odd term of this order will never occur again. The reason for this is that inserting Eq. 12 in Eq. 10 yields operator products, e.g., for fifth-order,

$$U_2 U_1 H_1 U_1^\dagger U_2^\dagger$$
$$\approx \underbrace{\left[\mathbf{1} + a_{2,1} W_2 + a_{2,2} W_2^2\right]}_{\approx U_2}$$
$$\times \underbrace{\left[\mathbf{1} + a_{1,1} W_1 + a_{1,2} W_1^2 + a_{1,3} W_1^3 + a_{1,4} W_1^4\right]}_{\approx U_1}$$
$$\times H_1 \underbrace{\left[\mathbf{1} - a_{1,1} W_1 + a_{1,2} W_1^2 - a_{1,3} W_1^3 + a_{1,4} W_1^4\right]}_{\approx U_1^\dagger}$$
$$\times \underbrace{\left[\mathbf{1} - a_{2,1} W_2 + a_{2,2} W_2^2\right]}_{\approx U_2^\dagger} \qquad (18)$$

which are either of low order in k and even or, if odd, of higher order (note that the approximation signs '$\approx$' in Eq. 18 do *not* imply that the Hamiltonian of fifth-order is approximated).

We finally obtain the DKH operator of given order n, $h_{\text{DKH}n}$, as a partial sum of even terms in expansion (Eq. 10),

$$h_{\text{bd}} = \sum_{k=0}^{n} \mathcal{E}_k + \sum_{k=n+1}^{\infty} \mathcal{E}_k = h_{\text{DKH}n} + O(\tilde{V}^n)$$
$$= h_{\text{DKH}n} + O(V^{n+1}), \qquad (19)$$

with the energy-damped external potential $\tilde{V}$ as expansion parameter — often abbreviated as 'the external potential V'.

It is important to understand that the Taylor expansion for the unitary matrices in Eq. 12 can be truncated at rather low order (compare Eq. 18). The zeroth-order even term $\mathcal{E}_0$ would be multiplied with some W_m^j operator if the expanded unitary transformation is inserted in Eq. 10. For $m = j = 1$ the smallest power in V results for the product term. Thus, the second unitary transformation U_1, which contains the lowest-order W_1 parameter has to be expanded at most up to the W_1^m term, if m is the desired order of the DKH Hamiltonian. Thus, the truncation of Eq. 12 does not introduce any approximation! However, a somewhat unpleasant property of the DKH expansion is that the inner unitary transformations have to be expanded to high powers of the W_m parameter because the W_m parameters of the inner transformations contain the

potential V only to some small power m; U_1 is parametrized by W_1, which contains V^1, U_2 is parametrized by W_2, which is already of order V^2, U_3 is parametrized by W_3, which contains V^3, and so forth.

The original formulation of the DKH method has been given up to second order, DKH2, by Douglas and Kroll [36] and implemented by Hess in 1986 [37, 42] (we owe Barysz and Sadlej the first instructive review of Hess' ingenious computational ingredients for the DKH method [43]; see also [35] for explicit details on the implementation into quantum chemistry program packages). The derivation and implementation of the third-order variant, DKH3, by Nakajima and Hirao [44] initiated a lot of new work on the DKH method. Their work also stimulated the generalization of the DKH method [35] so that the theory can now be presented in the most general way. First, this generalized DKH method has been explicitly developed and implemented up to fifth-order [35] where a dependence of $\mathcal{E}_5$ on some expansion coefficients $a_{m,j}$ occurs for the very first time (up to fourth-order, the DKH Hamiltonians are independent of the chosen parametrization). Of course, due to the embracing contiguous unitary transformations, such an expansion coefficient dependence is transduced to all $\mathcal{E}_{\geq 5}$ operators. This aspect has been further investigated by van Wüllen [45], who also gave explicit sixth order $\mathcal{E}_6$ expressions using the next unitary matrix U_3; we will come back to this issue in Sect. 4.3. We should emphasize that DKHn denotes the sum $h_{\mathrm{DKH},n} = \sum_{k=0}^{n} \mathcal{E}_k$ and not the single highest-order term $\mathcal{E}_n$.

The most important question is why the DKH scheme is unique. We have addressed this question in detail in Ref. [38]. The results of this formal analysis, which was necessary in order to understand whether exact decoupling with DKH can be achieved in actual calculations at all, may be briefly summarized in the following three paragraphs:

The initial transformation U_0 of the DKH procedure has *necessarily* and *uniquely* to be chosen as the well-known free-particle Foldy–Wouthuysen (fpFW) transformation [38]. This result was derived employing the framework of the generalized DKH transformation, i.e., the most general parametrization of unitary matrices defined by Eq. 12. Then, the uniqueness of the fpFW transformation can be understood in terms of this generalized expansion of U_0: all other parametrizations, i.e., all other choices of unitary expansion conditions $a_{0,j}$ yield ill-defined and singular expressions. Only the fpFW transformation produces the first single closed-form odd operator $\mathcal{O}_1$ linear in the potential V and can thus serve as the starting point for creating and annihilating odd terms of higher order in V. Note that the original odd term in the Dirac Hamiltonian in Eq. 1 is the term $c\,\boldsymbol{\alpha}\cdot\boldsymbol{p}$ which does not depend on V and is thus a zeroth-order term to be abbreviated as $\mathcal{O}_0$.

In principle, we may expand the block-diagonal Hamiltonian analogously to Eq. 10 with respect to any other expansion parameter. Reasonable expansion parameters are the inverse speed of light $1/c$, which is in Hartree atomic units used here identically to the fine structure constant α, or the external potential V. However, these formal expansion parameters are accompanied by additional quantities so

that the true expansion parameters are p/mc or the energy-damped potential $\tilde{V}$ (each V is accompanied by huge energy denominators; see below). Note that if we would have chosen an expansion parameter $1/m$ based on the rest mass m of the electron, we would have used the very same true expansion parameter p/mc as in the $1/c$ expansion. The failure of an expansion in terms of p/mc, which would represent a higher-order FW transformation, has been shown to be related to the kinetic term $\mathcal{E}_0$ and stems simply from a nonconvergent Taylor series expansion [38]. Due to the $1/r$-singularity of the Coulomb potential, the FW transformation may be affected by additional problems for $r \rightarrow 0$, which has been known for some time (see e.g., the work of Kutzelnigg [46,47]).

The decisive aspect with respect to the fpFW is the generation of $\mathcal{O}_1$ containing exactly only linear-potential terms rather than the consecutive generation of the higher-order odd terms, which are all of a single well-defined high-order in V. With $\tilde{V}$ being the only remaining expansion parameter, the DKH protocol is thus the *unique* transformation scheme for decoupling the Dirac Hamiltonian [38]. Consequently, the IOTC approach can be discussed in this framework and represents a purely numerical 'one-step' variant of the infinite-order DKH scheme [38].

Here, we shall now proceed with the discussion of exact analytical decoupling within the DKH approach. The results obtained in Ref. [38] have shown that the exact decoupling of the Dirac Hamiltonian in the original DKH framework can only be achieved analytically. The first step towards exact decoupling is thus the implementation of a symbolic derivation of DKH Hamiltonians of arbitrary-order, which can then be evaluated by matrix multiplications. This has been realized within the DKH method for the first time only recently [41] (see Table 1 for representative results). For exact decoupling, an equally important second step has to be taken: since DKH Hamiltonians can only be derived order by order, which has been automated, it is necessary to know in advance the order which is sufficient for decoupling to machine precision. This step has also been taken in Ref. [41]. It was possible to define a truncation error operator which allows to determine the maximum order needed for exact decoupling (within machine precision) for a given potential V owing to the well-defined convergence behaviour of the DKH series.

The presentation so far has been restricted only to the one-electron Dirac operator in an external field, e.g., to the DKH treatment of the hydrogen atom. The reason for this was that all essential steps can be dicussed for this system. We should emphasize again that the extension to many-electron systems is straightforward if one uses the tensor calculus for the construction of many-electron Hamiltonians and states (see [26]).

4 Conceptual issues of DKH theory

Douglas–Kroll–Hess theory appears to be rather complicated and involved. Therefore, we proceed to the investigation of

Table 1 $1/Z^2$-scaled one-electron $1s$ energies ($E_{\text{scaled}} = -E/Z^2$) for the DKH scheme employing an exponential expansion for the parametrization of the unitary transformations in Eq. 12 (in Hartree atomic units). DEQ denotes the exact Drac eigen value known analytically

Z	20	40	60	80	100
n.rel.	0.500000000	0.500000000	0.50000000	0.50000000	0.5000000
DKH1	0.503353740	0.514933891	0.53727855	0.57600742	0.6472394
DKH2	0.502681349	0.511009862	0.52608269	0.55051952	0.5906192
DKH3	0.502691655	0.511137573	0.52662335	0.55204993	0.5942370
DKH4	0.502691274	0.511128036	0.52656307	0.55182946	0.5936474
DKH5	0.502691312	0.511130065	0.52658415	0.55194727	0.5941541
DKH6	<u>0.502691308</u>	0.511129616	0.52657690	0.55189081	0.5938314
DKH7	0.502691308	0.511129699	0.52657886	0.55191021	0.5939588
DKH8	0.502691308	0.511129683	0.52657829	0.55190266	0.5938970
DKH9	0.502691308	0.511129687	0.52657849	0.55190627	0.5939362
DKH10	0.502691308	<u>0.511129686</u>	0.52657841	0.55190444	0.5939107
DKH11	0.502691308	0.511129686	0.52657844	0.55190526	0.5939244
DKH12	0.502691308	0.511129686	<u>0.52657843</u>	0.55190488	0.5939165
DKH13	0.502691308	0.511129686	0.52657843	0.55190508	0.5939217
DKH14	0.502691308	0.511129686	0.52657843	0.55190498	0.5939182
DEQ	0.502691308	0.511129686	0.52657843	0.55190502	0.5939195

The horizontal lines denote the order at which exact decoupling is obtained (within the given accuracy). This order is naturally dependent on the external potential, i.e., on Z. The table has been produced from the results given in Ref. [41]. The DKH energies are given up to the 14th order in the external potential V

additional conceptual difficulties, some of which have not yet been discussed elsewhere.

4.1 The momentum space 'myth'

When Hess presented his work in 1986, he followed the momentum-space representation of Douglas and Kroll. Our derivation so far has not made any reference to a particular choice of representation. And in fact it is not required to represent all operators in an explicit momentum-dependent basis set. The origin of the momentum-space presentation of the DKH method may be traced back to the square-root operator in $\mathcal{E}_0$, i.e., $\mathcal{E}_0 = \left[\beta\sqrt{\boldsymbol{p}^2 c^2 + c^4} - c^2\right]$ [compare Eqs. 16 and 17]. This term requires the evaluation of the square root of the momentum operator. Such a square-root expression can hardly be evaluated in a position-space formulation with linear momentum operators as differential operators (compare for this line of reasoning Ref. [18]). In a momentum-space formulation, however, the momentum operator takes a simple multiplicative form.

Formally, we then have to Fourier-transform all operators which possess in *position* space a simple multiplicative form like the scalar potential V, and obtain integral operators defined by their operator kernels. Hence, also Eq. 15 has to be given in momentum space and is usually written for the corresponding operator kernels as

$$W_m(p_0, p_1, \ldots, p_m) - \frac{a_{m,0}}{a_{m,1}}\beta\frac{\mathcal{O}_m^{(m)}(p_0, p_1, \ldots, p_m)}{E_{p_0} + E_{p_m}} = 0.$$
(20)

However, neither Eq. 15 nor Eq. 20 can be evaluated in the DKH framework because all operators in these equations are odd and cannot therefore be calculated (there is no small component in a two-component theory!). Nonetheless, the

W_m operators are, of course, present in the even terms of the DKH Hamiltonian and need to be evaluated. Since they are multiplied with other odd operators, all odd operators can be grouped and the odd×odd products are even and thus computable.

A quantum chemical calculation of a molecule is usually carried out in a given one-particle position-space basis set, typically consisting of a set of Gaussian-type basis functions. As mentioned above, already the lowest-order even term, $\mathcal{E}_0$, requires the evaluation of the square root of a sum containing the square of the momentum operator. Hess realized [39] that this only requires a basis, which yields a diagonal representation for p^2. Since the nonrelativistic kinetic energy matrix of a single electron, $(t) = (p^2/2)$, is calculated in all quantum chemistry program packages, diagonalization of (t) can easily be carried out and the eigenvectors yield the transformation matrix from the original position-space basis set to a linearly transformed position-space basis, which yields a diagonal representation for p^2. After calculation and addition of all even terms in this p^2-basis by standard matrix operations, the resulting matrix representation of the DKH Hamiltonian can be transformed back to the original basis with the inverse transformation.

4.2 Is the DKH expansion in powers of V unique?

The dependence of the DKH Hamiltonians of fifth and higher order on the expansion coefficients $a_{m,j}$ in the parametrization of the unitary transformation in Eq. 12 clearly shows that the DKH expansion in Eq. 10 cannot be unique in the sense that it is possible to have different expansions of h_{bd} in terms of the external potential [this question must not be confused with the uniqueness of the external potential V as the only possible expansion parameter in Eq. 10]. At first sight, this

is odd because one would expect that an expansion of h_{bd} in powers of V according to Eq. 10 should be unique. At infinite-order, all infinitely many different expansions with respect to the external potential V, formally written in Eq. 10 as a single unique one, do necessarily yield the same spectrum of the block-diagonal Hamiltonian h_{bd}.

Obviously, the different expansions are related to one another. Let us assume two different sequences of unitary transformations,

$$U = \cdots U_4 U_3 U_2 U_1 U_0 \qquad (21)$$

and

$$U' = \cdots U'_4 U'_3 U'_2 U'_1 U'_0, \qquad (22)$$

which differ in the choice of expansion coefficients $a_{m,j}$ in Eq. 12. Then, we can formally write the relation between both the resulting expansions of h_{bd},

$$h_{bd}^{(U)} = \cdots U_4 U_3 U_2 U_1 U_0 h_D U_0^\dagger U_1^\dagger U_2^\dagger U_3^\dagger U_4^\dagger \cdots = U h_D U^\dagger \qquad (23)$$

and

$$h_{bd}^{(U')} = \cdots U'_4 U'_3 U'_2 U'_1 U'_0 h_D U_0'^\dagger U_1'^\dagger U_2'^\dagger U_3'^\dagger U_4'^\dagger \cdots$$
$$= U' h_D U'^\dagger, \qquad (24)$$

as a linear map S

$$S = \cdots U'_4 U'_3 U'_2 U'_1 U'_0 U_0^\dagger U_1^\dagger U_2^\dagger U_3^\dagger U_4^\dagger \cdots = U' U^\dagger, \qquad (25)$$

which transforms the expansion of $h_{bd}^{(U)} = \sum_i \mathcal{E}_i^{(U)}$ to the one for $h_{bd}^{(U')} = \sum_i \mathcal{E}_i^{(U')}$.

However, this formal trick to describe the relation between both expansions does not explain explicitly how the difference between the Hamiltonians (Eq. 23) and Eq. 24 at a given power $n > 4$ of the expansion parameter V is compensated at some higher power $m > n$. This aspect has been studied by van Wüllen who gave explicit formulations for the fifth- and sixth-order Hamiltonians, which shall highlight the compensation mechanism in a systematic manner [45].

4.3 Relevance of the dependence on the expansion parameters $a_{m,j}$

From the last subsection we understand that one should use DKH Hamiltonians either only up to fourth-order or up to infinite-order (within machine precision) in order to avoid ambiguities in the method. However, even for other truncated DKH Hamiltonians of fifth- or higher-order, these ambiguities are very small and can hardly be observed if the basis set used is not sufficiently large (it needs to contain huge coefficients in the exponents of the basis functions, which might then cause numerical instabilities). Thus, a discussion of the coefficient dependence of the results obtained with fifth or higher order DKH Hamiltonians is not relevant as the numerical effect is tiny and one can always restrict a calculation to the fourth or to some high order, say DKH10.

Interestingly, for a long time it was believed that the original square-root parametrization of Douglas and Kroll

is particularly useful since the mandatory Taylor expansion of the square root according to Eq. 12 contains only even terms (all odd terms possess zero coefficients). Nakajima and Hirao [44] then used a different parametrization, which is much more well known in the quantum chemistry community, namely the exponential parametrization (other closed-form parametrizations for Eq. 12 are also known [25–27]). And indeed, the square-root parametrization leads to the smallest number of terms in the DKH Hamiltonian [41], but the exponential parametrization can be expected to converge faster. However, the purpose of the parametrization defined by Eq. 12 is to yield unitary Taylor expansions. The only formal choice for an optimum parametrization is thus one, which requires the expansion parameters $a_{m,j}$ to fulfill the unitarity condition as closely by as possible. On the basis of this requirement it is possible to define the so-called *optimum* parametrization [35].

4.4 Convergence properties of the DKH energies

It has already been mentioned in Sect. 3 that the true expansion parameter in Eq. 10 is the energy-damped potential $\tilde{V}$. The origin of this damping is the energy of the freely moving relativistic electron in E_p by Eq. 15 [38]. In the matrix operations needed for the evaluation of the DKH Hamiltonian, the energy damping occurs automatically via

$$R_i V R_j \longrightarrow \tilde{V}_{ij} = \frac{V_{ij}}{E_i + E_j}$$
$$= \frac{V_{ij}}{\sqrt{\boldsymbol{p}_i^2 c^2 + c^4} + \sqrt{\boldsymbol{p}_j^2 c^2 + c^4}}, \qquad (26)$$

where V_{ij} is a matrix element of the external potential in the p^2-basis. Note that the $\boldsymbol{p}_i^2 = p_i^2$ values are simply the eigenvalues of the nonrelativistic kinetic energy matrix multiplied by two. Since we are working in a p^2-basis, we have associated with each basis vector a value for the squared momentum.

The variational stability of the DKH Hamiltonians — a boundary condition for the convergence of the DKH series — has been proven for DKH2 [48] and new results for the higher-order expansion have been obtained as well [72]. Despite these analytical findings, which may be regarded as a first step towards an even deeper understanding of DKH theory, first numerical results obtained for the eigenvalues of higher-order DKH Hamiltonians created doubts on the method. Nakajima and Hirao [44] obtained a smooth convergence even for the third-order DKH energies. A closer inspection up to fifth-order [35] and later up to 14th order [38] revealed an oscillating but convergent behaviour, which is paralleled by the magnitude and sign of the truncation operator defined in Ref. [38]; the odd orders DKH1, DKH3, DKH5, DKH7, DKH9, ... always yield energies below the four-component reference energy. The first result by Nakajima and Hirao appears to be an artifact of a too small basis set. Later, additional doubts on the DKH methods arose from results

obtained to approximate the degenerate Dirac energies $s_{1/2}$ and $p_{1/2}$, which turned out to be nondegenerate at low-order DKH theory [49–51]. However, all these discrepancies vanish at higher order provided the basis set is properly chosen. The different accuracy of the low-order $s_{1/2}$ and $p_{1/2}$ energies has been rationalized by van Wüllen [51].

4.5 At which stage do we get rid of the negative-energy states?

The exact block diagonalization of the Dirac Hamiltonian ensures that no coupling between electronic and positronic states occurs, but the block-diagonal Hamiltonian is still a 4×4 (super)matrix. The 'true' decoupling is only achieved if we manage to *selectively* calculate h_+ so that the lower block Hamiltonian h_- is *never* obtained. In fact, this is actually done but the central point is usually not explicated.

In order to understand this point, we start from the explicit expansion of Eq. 12 inserted in Eq. 10 as carried out in Eq. 18. For the evaluation of the DKH Hamiltonian we have to carry out all multiplications in Eq. 18. Usually the even terms are then collected and represented as commutators. However, for their evaluation the commutators are expanded again. The resulting products contain only a small number of matrices [41] which contain β operators that may be contracted within the expression for each term of the operator by using proper commutation and anti-commutation rules in order to exploit the relation $\beta^2 = 1$ as often as possible. If there exists an even number of β operators in a term of the DKH Hamiltonian, then all β matrices vanish and the sign of this term might have changed due to anti-commutation relations. If the number is odd, one β operator survives. These surviving β operators lead to the two types of solutions, i.e. the positive- and negative-energy states, because of the diagonal structure $(1,1,-1,-1)$ of the β matrices. Since we are interested only in the upper block Hamiltonian h_+, we keep the upper 2×2 block of the β operator, which is the 2×2 unit matrix 1_2. Thus, the β can be skipped for electronic states, while for negative-energy states (i.e., for the so-called positronic states) we would have to change the sign of all operators linear in β for the calculation of h_- (the other terms are independent of β and remain unchanged).

4.6 Is there an approximate RI step in DKH theory?

An interesting aspect of DKH theory is that (standard) operator matrices for the (nonrelativistic) kinetic energy and external potential can be used for the evaluation of the DKH Hamiltonian. Only one nonstandard matrix is needed, which can be calculated with little additional effort. This is the matrix representation of pVp,

$$\boldsymbol{p} \cdot V \boldsymbol{p} \longrightarrow (p_x V p_x) + (p_y V p_y) + (p_z V p_z), \tag{27}$$

which is evaluated by operation with the momentum operator on the basis functions and integration of the resulting matrix elements. From diagonalization of the kinetic energy

we obtain the eigenvalues to calculate the relativistic energy of a freely moving particle and the eigenvector matrix for the transformation of the V and pVp matrices in the p^2-basis. Thus, the DKH Hamiltonian is calculated as a matrix representation for the relativistic energy of a freely moving particle to which correction terms are added that include order-by-order the action of the external potential.

However, there is a peculiarity involved because some terms in the DKH Hamiltonian are of the form $\boldsymbol{p} V \ldots V \boldsymbol{p}$, i.e., no momentum operator occurs between the potential energy operators and thus no matrix representation is at hand. Hess' solution to this problem was the introduction of a resolution of the identity (RI),

$$\mathbf{1} = \frac{\boldsymbol{p} \cdot \boldsymbol{p}}{p^2}, \tag{28}$$

which generates $\boldsymbol{p} V \boldsymbol{p}$ matrix operators after insertion between the two external potential operators and rearranging the $\boldsymbol{p}$ operators:[4]

$$\boldsymbol{p} V \ldots V \boldsymbol{p} = \boldsymbol{p} V \ldots \frac{\boldsymbol{p} \cdot \boldsymbol{p}}{p^2} \ldots V \boldsymbol{p} = \boldsymbol{p} V \boldsymbol{p} \ldots \frac{1}{p^2} \ldots \boldsymbol{p} V \boldsymbol{p}. \tag{29}$$

So far, we have *not* introduced any approximation. But the evaluation of the operator sequence on the right hand side of Eq. 29 requires a translation into products of operator matrices,

$$\begin{aligned} &\boldsymbol{p} V \boldsymbol{p} \ldots \frac{1}{p^2} \ldots \boldsymbol{p} V \boldsymbol{p} \\ &\longrightarrow (\boldsymbol{p} V \boldsymbol{p})(\ldots)(1/p^2)(\ldots)(\boldsymbol{p} V \boldsymbol{p}), \end{aligned} \tag{30}$$

where $(\ldots)$ symbolizes the matrix form. It is this last translation step which makes the introduced RI approximate and which requires a large basis set, since Eq. 30 is only exact in a complete basis set.

4.7 The density paradoxon

Another interesting aspect of DKH theory is what one might call 'the density paradoxon'. The density distribution ρ of a single electron in an external potential calculated from a four-component wave function is defined as

$$\rho(\boldsymbol{r}) = \psi^\dagger(\boldsymbol{r})\psi(\boldsymbol{r}) = \psi_L^\dagger(\boldsymbol{r})\psi_L(\boldsymbol{r}) + \psi_S^\dagger(\boldsymbol{r})\psi_S(\boldsymbol{r}), \tag{31}$$

and ρ fulfills a continuity equation,

$$\frac{\partial}{\partial t}\rho + \mathrm{div}\,\boldsymbol{j} = 0, \tag{32}$$

with the components of the current density given by

$$j^k = c\psi^\dagger \alpha^k \psi. \tag{33}$$

---

[4] Note the implication that we never encounter a (single) linear momentum between two V operators and that there is always a (single) linear momentum on the right and left hand side of the right and left V operators, respectively, in Eq. 29 if the RI is to be inserted between two V operators. This implication is always fulfilled as can be understood from the construction of the DKH Hamiltonian according to Eq. 18.

The density ρ in Eq. 31 possesses *no* nodes if only one product, i.e. $\psi_L^\dagger \psi_L$ or $\psi_S^\dagger \psi_S$, is zero for a given $\boldsymbol{r}$. This is different in nonrelativistic Schrödinger quantum mechanics. For instance, the radial electron density distribution of a *nonrelativistic* electron in a hydrogen-like atom is zero if the radial function $R_{nl}(r)$ is zero (there exist $(n-l-1)$ radial nodes with n being the principal quantum number). However, the four-component radial electron density is prevented from becoming zero owing to the fact that the large and small radial terms of Eq. 31 possess their nodes at different positions.

Inspection of the density resulting from the states of the block-diagonal Dirac Hamiltonian h_{bd} for a single electron in an external field calculated as $\phi_L^\dagger(\boldsymbol{r})\phi_L(\boldsymbol{r})$,

$$\rho = \psi^\dagger \psi = \psi^\dagger \mathbf{1} \psi = \psi^\dagger U^\dagger U \psi$$
$$= (U\psi)^\dagger U\psi = \phi_L^\dagger \phi_L = \rho_{\mathrm{bd}}, \qquad (34)$$

shows that the four-component density is obtained from the scalar product $\phi_L^\dagger \phi_L$ of the two-component orbital. This is in accordance with the fact that the continuity equation Eq. 32 then reads

$$\frac{\partial \phi_L^\dagger \phi_L}{\partial t} + \mathrm{div}\,\boldsymbol{j'} = 0, \qquad (35)$$

with the transformed current density components

$$j'^{,k} = c(U\psi)^\dagger \alpha^k (U\psi). \qquad (36)$$

One may now ask whether this two-component density $\rho_{\mathrm{bd}} = \phi_L^\dagger \phi_L$ is also able to reproduce the mutual annihilation of nodes described for the four-component density in Eq. 31 above. At first sight, one is tempted to claim that this is not possible as there is no small component whose squared value may add a finite number to compensate for a vanishing large component according to Eq. 31. However, the two-component transformed orbitals in the product $\phi_L^\dagger \phi_L$ are in general complex-valued so that $\phi_L^\dagger \phi_L = |\phi_L|^2 = \phi_{L,\mathrm{Re}}^2 + \phi_{L,\mathrm{Im}}^2$, which is also non-zero if real and imaginary parts, $\phi_{L,\mathrm{Re}}$ and $\phi_{L,\mathrm{Im}}$, possess their nodes at different values for the variable $\boldsymbol{r}$. If ϕ_L were a real-valued function, we would have to have two different radial parts for the two components of the 2-spinor ϕ_L in order to obtain a nodeless radial density distribution as in the four-component case. In a one-component approximation, which neglects all spin-dependent terms and which therefore does not resemble the four-component density, additional nodes may occur as in the Schrödinger case.

5 DKH theory quo vadis?

In this work, we have revisited DKH theory from a conceptual point of view in order to highlight conceptual difficulties which are usually not discussed in standard technical papers on this theory. One aim of this work was also to emphasize the essential parts of the theory to be separated from insignificant details on which time and effort need not be wasted in future. The nice features of DKH theory are that no negative-energy artifacts occur, that accuracy can be systematically improved, and that it is a unique transformation theory.

The nicest computational feature of the method is that its scalar relativistic variant, i.e., where all terms in the Hamiltonian carrying Pauli spin matrices have been neglected after application of Dirac's relation, can be easily implemented in any nonrelativistic quantum chemistry code. Even the scalar-relativistic higher-order DKH operators up to, say, tenth order, DKH10, can be implemented to take little computer time compared to the subsequent SCF iterations. However, for valence-shell properties like bond distances, vibrational frequencies, or electric field gradients, the higher-order contributions turned out to be negligible and the scalar-relativistic effect is quantitatively recovered already at the level of Hess' original DKH2 method [52, 53]. However, for core-dominated properties like NMR chemical shifts and coupling constants, higher orders become relevant. In general, one may recommend the use of the fourth order DKH4 Hamiltonian, which is of higher order and still independent of the choice of parametrization of the unitary matrix.

The epilogue is that the DKH theory offers a way to circumvent the zoo of quasi-relativistic Hamiltonians as such a plethora of operators would open a Pandora's box of quantum chemical studies on numerical effects of the various corrections, which would not add any insight to relativistic quantum chemistry. However, there are still several issues connected with the DKH theory which need to be clarified. Computational challenges are (1) the inclusion of spin–orbit, and many-electron terms in the unitary transformation of the full many-electron Hamiltonian cannot be neglected, especially if higher-order DKH Hamiltonians are employed as these effects are of comparable size. In pioneering work [54–62] only the lowest-order terms have been included so far. (2) the calculation of geometry gradients, which are desperately needed for structure optimizations, has only been tackled for the second-order standard method, DKH2, [63]. (3) the calculation of molecular properties, where picture change effects must be avoided [64], is another challenging topic for future research. Most treatments deal with the standard DKH2 method (see e.g., [65–68]) but first steps have been taken for DKH Hamiltonians up to fifth order [52, 53, 69] and a general arbitrary-order framework is currently being developed (Wolf and Reiher, submitted). All these computational issues tempt us to introduce powerful new approximations but care has to be taken in order to avoid an undesirably large number of DKH variants, which would be a step in the wrong direction.

Also, theoretical challenges are left to be tackled, for instance: (1) The DKH hydrogen atom awaits its solution. A first step would be the investigation of the analytic properties of the DKH 2-spinor for the limiting cases for $r \to 0$ and $r \to \infty$. (2) The treatment of large molecules with many light atoms would be most convenient in a relativistic embedding approach, which allows one to treat light atoms with nonrelativistic Schrödinger quantum mechanics but heavy atoms

with a two-component DKH approach. However, a rigorous interfacing of theories is by no means trivial — if possible at all. (3) Although the number of terms in the DKH Hamiltonian at a given (high) order is large and increases dramatically with the order, the correction obtained for the energy is small compared to the effort, i.e., the smaller the contribution to the energy, the greater the effort to calculate it, which is somewhat unintuitive. It will be interesting to see whether future work allows reducing the number of terms by assessing their order of magnitude on the basis of analytic arguments.

To conclude, the formalism of DKH theory, which experienced significant improvements and offered new insights during the last couple of years, is now well-developed but a lot of new exciting work still needs to be done in order to make this fully relativistic electrons-only theory a widely used approach in *all* areas of theoretical chemistry.

References

1. Einstein A (1905) Ann Phys (Leipzig) 17:891–921
2. Bagus PS, Lee YS, Pitzer KS (1975) Chem Phys Lett 33:408–411
3. Pyykkö P, Desclaux J-P (1979) Acc Chem Res 12:276–281
4. Ziegler T, Snijders JG, Baerends EJ (1981) J Chem Phys 74:1271–1284
5. Pitzer KS (1979) Acc Chem Res 12:271–276
6. Pyykkö P (1988) Chem Rev 88:563–594
7. Schwerdtfeger P (ed) (2002) Relativistic electronic structure theory. Part 1. Fundamentals. Elsevier, Amsterdam
8. Schwerdtfeger P (ed) (2004) Relativistic electronic structure theory. Part 2. Applications. Elsevier, Berlin
9. Hess BA (ed) (2003) Relativistic effects in heavy-element chemistry and physics. Wiley, New York
10. Hirao K, Ishikawa Y (eds) (2004) Recent advances in relativistic effects in chemistry. World Scientific Publishing, Singapore
11. Feynman RP (1961) Quantum electrodynamics. Benjamin New York
12. Gupta SN (1977) Quantum electrodynamics. Gordon and Breach Science Publishers, New York
13. Mohr PJ, Plunien G, Soff G (1998) Phys Rep 293:227–369
14. Dirac PAM (1928) Proc Roy Soc London A 117:610–624
15. Dirac PAM (1928) Proc Roy Soc London A 118:351–361
16. Thaller B (1992) The Dirac equation. Springer, Berlin Heidelberg New York
17. Goidenko I, Labzowsky L, Eliav E, Kaldor U, Pyykkö P (2003) Phys Rev A 67:020102, 1–3
18. Schwabl F (2004) Quantenmechanik für Fortgeschrittene, 3rd edn. Springer, Berlin Heidelberg New York
19. Reiher M, Hinze J (2002) Four-component ab initio methods for electronic structure calculations of atoms, molecules, and solids. In: Hess BA (ed) Relativistic effects in heavy-element chemistry and physics; Wiley, Chichester
20. Jensen HJA, Saue T, Visscher L et al (2004) DIRAC, a relativistic ab initio electronic structure program, Release DIRAC04.0, http://dirac.chem.sdu.dk
21. Pyykkö P (1986) Relativistic theory of atoms and molecules — a bibliography 1916–1985. Volume 41 of Lecture Notes in Chemistry. Springer. Berlin Heidelberg New York
22. Pyykkö P (1993) Relativistic theory of atoms and molecules II – a bibliography 1986–1992. Volume 60 of Lecture Notes in Chemistry. Springer, Berlin Heidelberg New York
23. Pyykkö P (2000) Relativistic theory of atoms and molecules, vol III – A Bibliography 1993–2000. Springer, Berlin Heidelberg New York
24. Pyykkö P (2004) Database 'RTAM' — relativistic quantum chemistry database 1915–2004; http://www.csc.fi/rtam/
25. Wolf A, Reiher M, Hess BA (2002) Two-component methods and the generalised Douglas–Kroll transformation. In: Schwerdtfeger P (ed) Relativistic quantum chemistry, vol I. Theory; Theoretical and computational chemistry. Elsevier, Amsterdam, pp 622–663
26. Wolf A, Reiher M, Hess BA (2004) Transgressing theory boundaries: the generalized Douglas–Kroll transformation. In: Hirao K, Ishikawa Y (eds) Recent advances in relativistic effects in chemistry. World Scientific Publishing, Singapore, pp 137–190
27. Reiher M, Wolf A, Hess BA (2005) Relativistic quantum chemistry: from quantum electrodynamics to quasi-relativistic methods. In: Rieth M, Schommers W (eds) Handbook of theoretical and computational nanotechnology. American Scientific Publishers, Stevenson Ranch (in press)
28. van Lenthe E, Baerends E-J, Snijders JG (1993) J Chem Phys 99:4597–4610
29. van Lenthe E, Baerends E-J, Snijders JG (1994) J Chem Phys 101:9783–9792
30. Chang C, Pélissier M, Durand P (1986) Phys Scr 34:394–404
31. Barysz M, Sadlej AJ (2002) J Chem Phys 116:2696–2704
32. Heully JL, Lindgren I, Lindroth E, Lundquist S, Mårtensson-Pendrill AM (1986) J Phys B 19:2799–2815
33. Kutzelnigg W (1997) Chem Phys 225:203–222
34. Barysz M, Sadlej AJ, Snijders JG (1997) Int J Quant Chem 65:225–239
35. Wolf A, Reiher M, Hess BA (2002) J Chem Phys 117:9215–9226
36. Douglas M, Kroll NM (1974) Ann Phys 82:89–155
37. Hess BA (1986) Phys Rev A 33:3742–3748
38. Reiher M, Wolf A (2004) J Chem Phys 121:2037–2047
39. Hess BA (1985) Phys Rev A 32:756–763
40. Foldy LL, Wouthuysen SA (1950) Phys Rev 78:29–36
41. Reiher M, Wolf A (2004) J Chem Phys 121:10945–10956
42. Jansen G, Hess BA (1989) Phys Rev A 39:6016–6017
43. Barysz M, Sadlej AJ (2001) J Mol Struct (Theochem) 573:181–200
44. Nakajima T, Hirao K (2000) J Chem Phys 113:7786–7789
45. van Wüllen C (2004) J Chem Phys 120:7307–7313
46. Kutzelnigg W (1989) Z Phys D 11:15–28
47. Kutzelnigg W (1990) Z Phys D 15:27–50
48. Brummelhuis R, Siedentop H, Stockmeyer E (2002) Doc Math 7:167–182
49. Kędziera D, Barysz M (2004) Chem Phys Lett 393:521–527
50. Kędziera D, Barysz M (2004) J Chem Phys 121:6719–6727
51. van Wüllen C (2005) Chem Phys 311:105–112
52. Wolf A, Reiher M, Hess BA (2004) J Chem Phys 120:8624–8631
53. Neese F, Wolf A, Fleig T, Reiher M, Hess BA (2005) J Chem Phys 122:204107
54. Samzow R, Hess BA (1991) Chem Phys Lett 184:491–495
55. Samzow R, Hess BA, Jansen G (1992) J Chem Phys 96:1227–1231
56. Hess BA, Marian CM, Wahlgren U, Gropen O (1996) Chem Phys Lett 251:365–371
57. Boettger JC (2000) Phys Rev B 62:7809–7815
58. Mayer M, Krüger S, Rösch N (2001) J Chem Phys 115:4411–4423
59. Matveev A, Rösch N (2003) J Chem Phys 118:3997–4012
60. Majumder S, Matveev AV, Rösch N (2003) Chem Phys Lett 382:186–193
61. Nakajima T, Hirao K (2003) J Chem Phys 119:4105–4111
62. Peralta JE, Scuseria GE (2004) J Chem Phys 120:5875–5881
63. Nasluzov VA, Rösch N (1996) Chem Phys 210:413–425
64. Kellö V, Sadlej AJ (1998) Int J Quant Chem 68:159–174

65. Fukuda R, Hada M, Nakatsuji H (2002) J Chem Phys 118:1015–1026
66. Fukuda R, Hada M, Nakatsuji H (2002) J Chem Phys 118:1027–1035
67. Malkin I, Malkina OL, Malkin VG (2002) Chem Phys Lett 361:231–236
68. Malkin I, Malkina OL, Malkin VG, Kaupp M (2004) Chem Phys Lett 396:268–276
69. Nakajima T, Hirao K (2000) Chem Phys Lett 329:511–516
70. Lee YS, McLean AD (1982) J Chem Phys 76:735–736
71. Stanton RE, Havriliak S (1984) J Chem Phys 81:1910–1918
72. Siedentop H, Stockmeyer E (2005) Phys Lett A 341: 473–478