Perspective: Essentials of Relativistic Quantum Chemistry

Wenjian Liu

1Qingdao Institute for Theoretical and Computational Sciences, Shandong University, Qingdao, Shandong 266237, P. R. China

Abstract

Relativistic quantum chemistry has evolved into a fertile and large field and is now becoming an integrated part of mainstream chemistry. Yet, given the much-involved physics and mathematics (as compared with nonrelativistic quantum chemistry), it is still necessary to clean up the essentials underlying the relativistic electronic structure theories and methodologies (such that uninitiated readers can pick up quickly the right ideas and tools for further development or application) and meanwhile pinpoint future directions of the field. To this end, the three aspects of electronic structure calculations, i.e., relativity, correlation, and QED, will be highlighted.

liuwj@sdu.edu.cn
I. INTRODUCTION: ABC OF RELATIVISTIC QUANTUM MECHANICS

As one of the two pillars of modern physics, the theory of special relativity was founded by Einstein in 1905 [1]. Among others, the best known energy-mass relation

\[ E = \gamma mc^2 \]  

(1)
is most relevant for our purpose. Here, \( m \) is the rest mass of a particle moving with velocity \( v \), while \( \gamma = (1 - \frac{v^2}{c^2})^{-1/2} \) is the Lorentz factor, with \( c \) being the constant velocity of light. This relation can be converted to

\[ E^2 = c^2 p^2 + m^2 c^4 \]  

(2)
via the very definition of the momentum \( p = \gamma mv \). In fact, the energy-momentum relation (2) is more fundamental than the energy-mass relation (1), since the former applies to both massive and massless particles whereas the latter applies only to massive particles. Moreover, relation (1) is merely the positive-energy part of the square root of the right-hand side of relation (2), i.e., \( E = \pm \sqrt{c^2 p^2 + m^2 c^4} \). In view of the correspondence principle, i.e.,

\[ E \rightarrow i\hbar \frac{\partial}{\partial t}, \quad p \rightarrow -i\hbar \nabla, \]  

(3)
relation (2) can be mapped, as done by Klein and Gordon in 1926 [2, 3], to a first-quantized wave equation,

\[ \left( \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 + k^2 \right) \psi_{KG}(x) = 0, \quad k = \frac{mc}{\hbar}, \quad x = rt, \]  

(4)
which treats space and time on the same footing and is manifestly Lorentz covariant. However, it is second order in time, which is fundamentally different from the Schrödinger equation that is first order in time. As a result, the norm-conserving density [4], \( \rho_{KG} = \frac{i}{2mc^2} (\psi_{KG}^* \frac{\partial \psi_{KG}}{\partial t} - \psi_{KG} \frac{\partial \psi_{KG}^*}{\partial t}) \), is not positive definite (since \( \psi_{KG} \) and \( \frac{\partial \psi_{KG}}{\partial t} \) are independent of each other and can have arbitrary values at a given time \( t \)) and hence cannot be interpreted as a probability density. Because of this, the Klein-Gordon equation (4) was not regarded to be physically meaningful until Pauli and Weisskopf [5] recognized, through a theoretical exercise [6], that it is a relativistic wave equation for spin-0, charged and massive particles, which were discovered to be \( \pi^+ \) and \( \pi^- \) mesons in the late 1940s (NB: \( \rho_{KG} \) multiplied by charge \( q \) can be reinterpreted as a charge density which can be either positive or negative). A relativistic first-quantized wave equation that is first order in time and in space was first proposed by Dirac in 1928 [7, 8], by noticing that
the energy-momentum relation (2) can be written as

\[ E^2 = D_0^2, \quad D_0 = c \alpha \cdot p + \beta mc^2 = D_0^\dagger \]

\[ = c^2 \sum_{i=x,y,z} \alpha_i^2 p_i^2 + \beta^2 m^2 c^4 + \sum_{i=x,y,z} [\beta, \alpha_i] + p_i mc^3 + \frac{c^2}{2} \sum_{i \neq j}^3 [\alpha_i, \alpha_j] + p_i p_j \]

\[ = c^2 p^2 + m^2 c^4, \]  

provided that the following conditions hold

\[ \beta^\dagger = \beta, \quad \alpha_i^\dagger = \alpha_i, \quad \beta^2 = \alpha_i^2 = 1, \quad [\beta, \alpha_i]_+ = [\alpha_i, \alpha_j]_+ = 0 \ (i \neq j). \]  

It is then not a difficult math to figure out the explicit, lowest-dimensional matrix expressions for \( \beta \) and \( \alpha \):

\[ \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \alpha = \begin{pmatrix} 0 & \sigma \\ \sigma & 0 \end{pmatrix}, \]

where \( \sigma \) is the vector of the 2 \( \times \) 2 Pauli spin matrices,

\[ \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \]

Recognizing the function \( D_0 \) (5) as the Hamiltonian, the correspondence principle (3) then leads immediately to the free-particle Dirac equation

\[ i\hbar \frac{\partial}{\partial t} \psi(x) = D_0 \psi(x). \]

At variance with the scalar form of \( \psi_{KG} \), the wave function \( \psi \) of the Dirac equation (11) is a bispinor with four components, i.e.,

\[ \psi(x) = \begin{pmatrix} \psi^L(x) \\ \psi^S(x) \end{pmatrix} = \begin{pmatrix} \psi^{L\alpha}(x) \\ \psi^{L\beta}(x) \\ \psi^{S\alpha}(x) \\ \psi^{S\beta}(x) \end{pmatrix}. \]

It can readily be checked that each of the four components satisfies the Klein-Gordon equation (4). It can also be shown that the density \( \rho = \psi^\dagger \psi \) is positive definite. Moreover, the appearance of the Pauli spin matrices (10) in the Dirac \( \alpha \) matrix (9) reveals that the Dirac equation (11) is a relativistic wave equation for spin-\( \frac{1}{2} \) particles, such that the components of the wave function (12) can be labeled by the \( \alpha \) and \( \beta \) spins.
As a matter of fact, the Dirac equation (11) can directly be obtained from the Klein-Gordon equation (4) by decomposing the latter into two coupled first-order equations. Following Kramers[9, 10], this can proceed by rewriting Eq. (4) as

\[
(i\hbar \partial_t - c\sigma \cdot p)(i\hbar \partial_t + c\sigma \cdot p)\psi_{KG}(x) = m^2 c^4 \psi_{KG}(x),
\]

where use of the identity \( p^2 = (\sigma \cdot p)^2 \) has been made. Further replacing \( \psi_{KG} \) with spinor \( \psi^l \) and \((i\hbar \partial_t + c\sigma \cdot p)\psi^l \) with \( mc^2 \psi^r \), we obtain

\[
 i\hbar \frac{\partial}{\partial t} \psi_W(x) = D_0^W \psi_W(x),
\]

where

\[
 D_0^W = \begin{pmatrix} c\sigma \cdot p & mc^2 \\ mc^2 & -c\sigma \cdot p \end{pmatrix}, \quad \psi_W(x) = \begin{pmatrix} \psi^r(x) \\ \psi^l(x) \end{pmatrix}.
\]

Eq. (14) is known as the Dirac equation in the Weyl representation. By further carrying out the following unitary transformation,

\[
 U_W = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = U_W^{-1},
\]

the Dirac equation in the standard representation (11) can be recovered. At this moment it is worthy mentioning that, although electron spin appears naturally in the Dirac equation, it is not a relativistic quantity, since it appears also in the Lévy-Leblond equation[11]

\[
 i\hbar \frac{\partial}{\partial t} \begin{pmatrix} \psi^L(x) \\ \phi^L(x) \end{pmatrix} = \begin{pmatrix} 0 & \sigma \cdot p \\ \sigma \cdot p & -2m \end{pmatrix} \begin{pmatrix} \psi^L(x) \\ \phi^L(x) \end{pmatrix},
\]

which is just the nonrelativistic limit (nrl)[12] of the Dirac equation (11) and can be introduced \textit{a priori} by means of a spinor representation of the (nonrelativistic) Galilei group.

For an electron \((q = -e)\) moving in an external electromagnetic field characterized by the vector potential \( A_{ext} \) and scalar potential \( \phi_{ext} \), the following minimal coupling relations

\[
p \rightarrow \pi = p - qA_{ext}, \quad i\hbar \frac{\partial}{\partial t} \rightarrow i\hbar \frac{\partial}{\partial t} - q\phi_{ext}
\]

for electromagnetic interaction can be invoked, so as to obtain

\[
i\hbar \frac{\partial}{\partial t} \psi(x) = D\psi(x),
\]

\[
 D = c\alpha \cdot (p - qA_{ext}) + \beta mc^2 + q\phi_{ext}.
\]
If the external field is static, suffice it to consider the following eigenvalue problem

\[ D \psi_p(r) = \varepsilon_p \psi_p(r), \tag{21} \]

which has three branches of solutions if the external field arises from a net positive charge distribution: positive-energy continuum, discrete positive-energy bound states and negative-energy continuum, as illustrated by the left panel of Fig. 1. The gap \( \Delta E \) between the lowest positive-energy level \( \epsilon_{1s} \) and the top edge \( -mc^2 \) of the negative-energy continuum can be calculated as

\[
\Delta E = \epsilon_{1s} - (-mc^2) = mc^2\sqrt{1 - (Z/c)^2} + mc^2 = 2mc^2 f_Z,
\]

\[
\frac{1}{2} < f_Z = \frac{1}{2} (1 + \sqrt{1 - (Z/c)^2}) < 1, \tag{22}
\]

where \( \epsilon_{1s} \) is the ground state energy of the Dirac equation for a one-electron atom of nuclear charge \( Z \). It is seen that the gap is indeed very large (e.g., \( f_Z \approx 0.91 \) for Hg\(^{79+} \)). It is not much changed for many-electron systems (\( f_Z \approx 0.92 \) for Hg).

The existence of an empty negative-energy continuum was extremely troublesome in the early days of relativistic quantum mechanics, for it implied that no atom would be stable! For instance, it can be estimated[4] that, in the presence of a radiative field (which always exists in reality), the electron in the ground state of the hydrogen atom can fall down to the top of the negative-energy continuum in less than one nanosecond, and it can even trigger a radiation catastrophe via continuous radiative transitions. To resolve this apparent untruth, Dirac[13] proposed in 1930 that all states of negative energy be filled, such that transitions of electrons to negative-energy states are forbidden by virtue of the Pauli exclusion principle. Since a hole left by exciting an electron from the filled negative-energy sea has a positive energy and the same mass but opposite charge as the leaving electron, it was interpreted by Dirac in 1931[14] as an anti-electron (positive electron/positron). Although highly controversial, this bold prediction...
was confirmed by experiment just one year later\textsuperscript{[15]}. Notwithstanding such a big triumph, Dirac’s hole theory has a number of defects\textsuperscript{[16]}:

(1) It is asymmetric with respect to electrons and positrons.

(2) It characterizes a positron as a virtual hole rather than a real particle.

(3) It involves an infinite negative electric charge filling the whole space even if only one electron is under consideration.

(4) It has to assume that the negative-energy electrons do not generate any potential acting on the positive-energy electrons. Otherwise, no nuclear charge could generate an enough attraction to compensate the infinitely repulsive potential.

(5) It does not, in a strict sense, explain the stability of a positive-energy electron: being infinitely large, the sea can always accept infinitely many electrons. In other words, the Pauli exclusion principle does not really hold for a system of an infinite number of fermions.

(6) It does not apply to spin-0 particles (described by the Klein-Gordon equation\textsuperscript{(4)}) which do not satisfy an exclusion principle.

The above problems associated with the negative-energy continuum drove the pioneers of quantum mechanics to formulate a quantum field theory for electrodynamics (QED) through a particular second quantization of the Dirac matter field and electromagnetic field, where all particles are of positive energy (see Ref.\textsuperscript{[6]} for a historical review of the early days of QED). To see how this can be achieved, we first take a look at the charge-conjugation transformation

\[
\hat{C} = C\beta \hat{K}_0, \quad C = -i\alpha_y
\]  

(23)
of the Dirac equation\textsuperscript{(21)}: taking the complex conjugate ($\hat{K}_0$) followed by multiplying $C\beta$ from the left leads to

\[
[c\alpha \cdot (p + qA_{\text{ext}}) + \beta mc^2 - q\phi_{\text{ext}}]\hat{\psi}_p(r) = \hat{\psi}_p(r)\epsilon_p^C, \quad \epsilon_p^C = -\epsilon_p.
\]  

(24)
The manipulation is facilitated by making use of the following identities

\[
C^\dagger = C^T = C^-1 = -C, \quad C\beta = -\beta C, \quad C\alpha^* = -\alpha C, \quad C^\dagger C = -\alpha^T.
\]  

(25)

To unify the notion, we define

\[
\psi^C_p(x) = \psi^C_p(r)e^{-i\epsilon_p t} = \hat{\psi}_p(x), \quad \psi^C_p(r) = C\beta\psi^*(p), \quad \epsilon_p = -\epsilon_p < 0.
\]  

(26a)
\[\psi_p^C(x) = \psi_p^C(r)e^{-ie_p t} = \hat{\psi}_p(x), \quad \psi_p^C(r) = C\beta\psi_p^*(r), \quad \epsilon_p = -\epsilon_\beta > 0. \quad (26b)\]

That is, apart from its apparent action, charge conjugation will also interchange the indices of the argument as \( p \leftrightarrow \bar{p} \), such that \( \epsilon_p > 0 \) and \( \epsilon_\beta < 0 \) always hold. Another example is \( \hat{C}[a_p \psi_p] = C\beta(a_p \psi_p)^{\dagger T} = C\beta(a^\dagger p \psi_p^*) = a^\dagger p \psi_p^C. \)

It is clear that, for the same time-independent external field \((\phi_{ext}(r), A_{ext}(r))\), if \( \psi_p(x) = \psi_p^C(r)e^{-ie_p t} \) is a stationary state of the Dirac equation for an electron \((q = -1)\) of positive energy \( \epsilon_p \), \( \psi_p^C(x) = \hat{C}\psi_p(x) = \psi_p^C(r)e^{-ie_p t} \) will then be a stationary state of the Dirac equation for a positron \((q = +1)\) of negative energy \( \epsilon_\beta = -\epsilon_p \). Likewise, if \( \psi_\beta(x) = \psi_\beta^C(r)e^{-i\epsilon_\beta t} \) is an electronic negative-energy state \((\text{NES})\), \( \psi_p^C(x) = \hat{C}\psi_\beta(x) = \psi_p^C(r)e^{-i\epsilon_\beta t} \) will be a positronic positive-energy state \((\text{PES}; \epsilon_p = -\epsilon_\beta; \text{cf. the right panel of Fig. 1})\). Note in particular that the probability density of a negative-energy electron \( \psi_p^C(r)e^{-i|\epsilon_\beta|t} \) is indistinguishable from that of a positive-energy positron \( \psi_p^C(r)e^{-i|\epsilon_\beta|t} \), i.e., \( |\psi_p^C(r)|^2 = |\psi_\beta(r)|^2 \). As such, a negative-energy electron propagating backward in time can be regarded as the mirror image of a positive-energy positron propagating forward in time. Therefore, the Dirac matter field should be quantized as

\[\hat{\phi}(x) = b_p \psi_p(x) + b^\dagger \psi_\beta(x), \quad (27a)\]

\[b_p|0\rangle = b_\beta|0\rangle = 0, \quad b^\dagger = b^\dagger_p, \quad b^\dagger = b^\dagger_\beta, \quad \epsilon_p > 0, \quad \epsilon_\beta < 0, \quad (27b)\]

in the interaction picture and the Einstein summation convention over repeated indices, in order for the field to comprise only of positive-energy particles: \( b_p \) annihilates an electron of positive energy \( \epsilon_p \), whereas \( b^\dagger \) creates a positron of positive energy \(|\epsilon_\beta| = -\epsilon_p\). Since any operator must be expanded in a complete (orthonormal) basis spanned by the PES and NES of the same Dirac equation [21], the amplitude accompanying \( b^\dagger \) can only be the electronic NES \( \psi_\beta^C(r)e^{-i|\epsilon_\beta|t} \) instead of the corresponding positronic PES \( \psi_p^C(r)e^{-i|\epsilon_\beta|t} \) (NB: in the presence of an external field, \( \{\psi_p^C(r)\} \) are even not orthogonal to the electronic PES \( \{\psi_\beta(r)\} \), i.e., the inner products \( \langle \psi_p^C|\psi_\beta \rangle \) are generally nonzero). On the other hand, charge conservation dictates that the operator \( b^\dagger \) (instead of \( b_\beta \)) must accompany \( b_p \). Both \( b_p \) and \( b^\dagger \) increase the charge of a state by one unit; \( b_\beta \) does this by destroying an electron whereas \( b^\dagger \) does this by creating a positron. Thus the field operator \( \hat{\phi}(x) \) always increases one unit of charge. Similarly, the field operator \( \hat{\phi}^\dagger(x) \) always decreases one unit of charge. Therefore, the operator \( \hat{\phi}^\dagger(x)\hat{\phi}(x) \) conserves the charge. Had \( b_\beta \) been chosen to accompany \( b_p \), the operator \( \hat{\phi}^\dagger(x)\hat{\phi}(x) \) would not conserve the charge: it would include terms like \( b^\dagger b_p \) and \( b^\dagger b_\beta \) which decrease and increase two units of charge, respectively.
The particular form (27) for the quantized Dirac matter field is the very first cornerstone of QED. It can actually be rewritten as

\[
\hat{\phi}(x) = a_p \psi_p(x) + a_{\bar{p}} \psi_{\bar{p}}(x),
\]

(28a)

\[
a_p |0_{e^-}; N_{e^-}\rangle = a_{\bar{p}} |0_{e^-}; N_{e^-}\rangle = 0, \quad a^+ = a_{p}^+, \quad a^0 = a_{\bar{p}}^+, \quad \epsilon_p > 0, \quad \epsilon_{\bar{p}} < 0,
\]

(28b)

by replacing the genuine vacuum \(|0\rangle\) with the physical vacuum \(|0_{e^-}; N_{e^-}\rangle\) consisting of zero positive-energy electrons and \(N_{e^-} \to \infty\) negative-energy electrons. That is, the particle-hole picture

\[
b_p = a_p, \quad b^p = a^0, \quad b_{\bar{p}} = a_{\bar{p}}, \quad b^{\bar{p}} = a_p, \quad \epsilon_p > 0, \quad \epsilon_{\bar{p}} < 0
\]

(29)

implied in Eq. (27a) is merely a mathematical operation and is only convenient for pictorial interpretation in terms of diagrams, where the expression (28a) is more convenient for algebraic manipulations[16]. At first glance, we have just gone back to the filled Dirac picture, such that the aforementioned problems associated with Dirac’s hole theory would arise again. However, the picture can be reversed: it is perfectly legitimate to quantize the Dirac matter field in terms of the solutions of the positron Dirac equation (24):

\[
\hat{\phi}^C(x) = d_p \psi_p^C(x) + d_{\bar{p}} \psi_{\bar{p}}^C(x),
\]

(30a)

\[
d_p |0\rangle = d_{\bar{p}} |0\rangle = 0, \quad d^p = d_{p}^+, \quad d^{\bar{p}} = d_{\bar{p}}^+, \quad \epsilon_p > 0, \quad \epsilon_{\bar{p}} < 0,
\]

(30b)

where \(d_p\) annihilates a positron of positive energy \(\epsilon_p\), whereas \(d^{\bar{p}}\) creates an electron of positive energy \(|\epsilon_{\bar{p}}|\). The charge-conjugation transformation of \(\hat{\phi}(x)\) (27) leads to[17] [18]

\[
\hat{\phi}^C(x) = C\beta \hat{\phi}^{+T}(x)
\]

(31)

\[
= C\beta [b^{\bar{p}} \psi_{\bar{p}}^*(x)] + C\beta [b_p \psi_p^*(x)]
\]

(32)

\[
= b^p \psi_p^C(x) + b_{\bar{p}} \psi_{\bar{p}}^C(x), \quad \text{s.t. } \epsilon_p > 0, \epsilon_{\bar{p}} < 0.
\]

(33)

By comparing Eq. (33) with Eq. (30a) we obtain

\[
d_p = b_p, \quad d^{\bar{p}} = b_{\bar{p}}, \quad \epsilon_p > 0, \quad \epsilon_{\bar{p}} < 0.
\]

(34)

That is, the \(d\) and \(b\) types of annihilation and creation processes are the same, although their carriers are different (positrons vs electrons). This is more transparent[17] for the case of free particles for which Eqs. (21) and (24) are identical (i.e., \(\psi_p^C(x) = \psi_p(x)\) and \(\psi_{\bar{p}}^C(x) = \psi_{\bar{p}}(x)\)), such that it is immaterial to interpret which set of the PES and NES as electrons or positrons. More over, just like Eq. (28), Eq. (30) can be rewritten as

\[
\hat{\phi}^C(x) = c_p \psi_p^C(x) + c_{\bar{p}} \psi_{\bar{p}}^C(x),
\]

(35a)
Now the vacuum is \( |0_{e^+}; N_{e^+}\rangle\) in lieu of the original \( |0; N_{e^-}\rangle\). Since the two types of (second) quantization of the same Dirac matter field are equivalent, they can simply be averaged with an equal weight. To show how this can be done, let us look at the four-current operators for electrons and positrons:

\[
\begin{align*}
\tilde{j}^\mu_{e^-}(x) &= -e\hat{\phi}^+(x)ca^\mu\hat{\phi}(x), \quad \alpha^\mu = (e^{-1}, \alpha), \quad e = +1, \\
&= -e\{\hat{\phi}^+(x)ca^\mu\hat{\phi}(x)\} - e\langle 0; N_{e^-} | \hat{\phi}^+(x)ca^\mu\hat{\phi}(x) | 0; N_{e^-}\rangle \\
&= -e\hat{\psi}_p^+(r)ca^\mu\hat{\psi}_p(r), \\
\tilde{j}^\mu_{e^+}(x) &= e\hat{\phi}^C(x)ca^\mu\hat{\phi}^C(x) \\
&= e\{\hat{\phi}^C(x)ca^\mu\hat{\phi}^C(x)\} + e\langle 0; N_{e^+} | \hat{\phi}^C(x)ca^\mu\hat{\phi}^C(x) | 0; N_{e^+}\rangle \\
&= e\hat{\psi}_\mu^+(r)ca^\mu\hat{\psi}_\mu^+(r),
\end{align*}
\]

where the first terms of Eqs. (37)/(38) and (40)/(41) are normal ordered with respect to \( |0; N_{e^-}\rangle\) and \( |0; N_{e^+}\rangle\), respectively. By means of the relation (31), the first term of Eq. (41) can be written as

\[
\begin{align*}
e\{\hat{\phi}^C(x)ca^\mu\hat{\phi}^C(x)\} &= ec\{\hat{\phi}^T \beta C^T \alpha^\mu C \beta \hat{\phi}^{TT}(x)\} \\
&= ec\{\hat{\phi}_\gamma(x)(\alpha^\mu)^T \gamma^T \hat{\phi}_\gamma^{TT}(x)\} \\
&= e\{ca^\mu\hat{\phi}(x)\hat{\phi}(x)\} \\
&= -e\{\hat{\phi}^+(x)ca^\mu\hat{\phi}(x)\},
\end{align*}
\]

where the normal ordering is now taken with respect to \( |0; N_{e^-}\rangle\). Likewise, the second term of Eq. (41) can be written as

\[
e\hat{\psi}_\mu^+(r)ca^\mu\hat{\psi}_\mu^+(r) = e\hat{\psi}_\mu^+(r)ca^\mu\hat{\psi}_\mu^+(r) = e\langle 0; N_{e^-} | ca^\mu\hat{\phi}(x)\hat{\phi}^+(x) | 0; N_{e^-}\rangle.
\]

Therefore, \(\tilde{j}^\mu_{e^+}(x)\) (39) can be written as

\[
\begin{align*}
\tilde{j}^\mu_{e^+}(x) &= eca^\mu\hat{\phi}(x)\hat{\phi}^+(x) \\
&= -e\hat{\phi}^+(x)ca^\mu\hat{\phi}(x) + e\langle 0; N_{e^-} | ca^\mu\hat{\phi}(x)\hat{\phi}^+(x) | 0; N_{e^-}\rangle.
\end{align*}
\]
The four-current operator averaged over electrons and positrons then reads

\[ j^\mu(x) = \frac{1}{2}(\bar{\gamma}_e^\mu(x) + \bar{\gamma}_e^\mu(x)) \]  
\[ = - \frac{1}{2} e[\hat{\phi}^+(x), ca^\mu \hat{\phi}] = - \frac{1}{2} eca^\mu_{\nu\rho} [\hat{\phi}^\nu(x), \hat{\phi}_\rho(x)] \]  
\[ = - e\{\hat{\phi}^+(x)ca^\mu \hat{\phi} + j^\mu_{fp}(r), \]  
\[ j^\mu_{fp}(r) = - e\langle vac | \frac{1}{2}[\hat{\phi}^+(x), ca^\mu \hat{\phi}(x)]|vac\rangle \]  
\[ = - \frac{1}{2} e[\psi^+_p(r)ca^\mu \psi_p(r) - \psi^+_p(r)ca^\mu \psi_p(r)]. \]  

Note in passing that the vacuum \(|\text{vac}\rangle\) in Eqs. (51) and (52) can either be \(|0\rangle\) along with the definition (27) or \(|0_{e^-}; N_{e^-}\rangle\) along with the definition (28). The zero component of \(j^\mu_{vp}(r)\) (53), i.e., the vacuum density \(\rho_{vp}(r)\), reads

\[ \rho_{vp}(r) = - e\langle vac | \frac{1}{2}[\hat{\phi}^+(x), \hat{\phi}(x)]|vac\rangle \]  
\[ = - \frac{1}{2} e[n_-(r) - n_+(r)], \]  
\[ n_+(r) = \psi^+_p(r)\psi_p(r) = \psi^C_p(r)\psi^C_p(r), \]  
\[ n_-(r) = \psi^+_p(r)\psi_p(r) = \psi^C_p(r)\psi^C_p(r), \]  

where \(n_+(r)\) and \(n_-(r)\) are the number densities of the PES and NES of the electron Dirac equation (21), respectively. By virtue of the identity \(n_+ + n_- = n_+ + n_- = 2n_-\), with \(n_+\) and \(n_-\) (\(= n_+\)) being the free-particle number densities, we have

\[ \rho_{vp}(r) = - e[n_-(r) - n_-(r)], \]  

which is clearly the charge polarization of the vacuum. Moreover, Eq. (55) reveals that the NES of the electron Dirac equation (21) are all occupied by electrons \(e^-\) with charge -1 (i.e., filled Dirac sea of electrons), whereas the PES by positrons \(e^+\) with charge +1. As shown above, the latter arises actually from the filled Dirac sea of positrons (i.e., \(|0_{e^+}; N_{e^+}\rangle\), as a direct consequence of charge conjugation. Therefore, the genuine vacuum \(|0\rangle\) can be viewed\([19]\) as the superposition of \(|0_{e^-}; N_{e^-}\rangle\) and \(|0_{e^+}; N_{e^+}\rangle\): the electrons and positrons annihilate each other spontaneously, so as to leave an empty vacuum. In other words, the original hole theory of Dirac\([13]\) for relativistic electrons should be generalized to “charge-conjugated hole theory” or simply “extended hole theory”. This feature is incorporated automatically into the symmetrized four-current operator (50), introduced first by Schwinger in 1951\([20]\). Given its great importance, the expression (50) should be viewed as another cornerstone of QED. Note in passing that, in the free-particle (fp) representation, \(j^\mu_{fp}(r)\) (52) vanishes pointwise, thereby leading to \(j^\mu_{fp}(x) = - e\{\hat{\phi}^+(x)ca^\mu \hat{\phi}\}\.\)
Last but not least, it is worthy mentioning that the commutator form of contraction \((54)\) is just a special case of the equal-time contraction (ETC)\([20]\) of fermion operators

\[
A(t)B(t) = \langle \text{vac}|T[A(t)B(t)]|\text{vac}\rangle \equiv \frac{1}{2} \langle \text{vac}|T[A(t)B(t')]|\text{vac}\rangle|\nu' - t \rightarrow 0^\pm
\]

\[
= \langle \text{vac}|\frac{1}{2}[A(t),B(t)]|\text{vac}\rangle,
\]

which is symmetric in time. That is, the two expressions \(A(t)B(t)\) and \(-B(t)A(t)\) obtained by letting \(t'\) approach \(t\) from the past and future are both considered and averaged here. Eq. \((61)\) is fundamentally different from the following ETC

\[
\bar{A}(t)\bar{B}(t) = \lim_{\eta \rightarrow 0^+} \langle 0; \bar{0}|T[A(t)B(t + \eta)]|0; \bar{0}\rangle,
\]

which is asymmetric in time and holds only in the nrl. Note that the ETC \((61)\) is only implicit in the Feynman fermion propagator\([19]\), such that its importance is often overlooked in the literature. Instead, we should regard it as an essential ingredient to distinguish relativistic from nonrelativistic quantum mechanics. As a time-independent analog of the ETC \((61)\), the charge-conjugated contraction (CCC) of fermion operators was also introduced\([21]\),

\[
\bar{a}^{\bar{p}}a_{\bar{q}} = \langle 0; N_e^-|\frac{1}{2} [a^{\bar{p}},a_{\bar{q}}]|0; N_e^-\rangle, \quad p, q \in \text{PES, NES}
\]

\[
= \frac{1}{2} \langle 0; N_e^-|a^{\bar{p}}a_{\bar{q}}|0; N_e^-\rangle|\epsilon_{\bar{p}} < 0, \epsilon_{\bar{q}} < 0 - \frac{1}{2} \langle 0; N_e^-|a_{\bar{q}}a^{\bar{p}}|0; N_e^-\rangle|\epsilon_{\bar{q}} < 0, \epsilon_{\bar{p}} > 0 \rangle
\]

\[
= -\frac{1}{2} \delta_{\bar{p}} \delta_{\bar{q}} \text{sgn}(\epsilon_{\bar{q}}), \quad p, q \in \text{PES, NES},
\]

which distinguishes from the standard contraction

\[
\bar{a}^{\bar{p}}a_{\bar{q}} = \langle 0; N_e^-|a^{\bar{p}}a_{\bar{q}}|0; N_e^-\rangle, \quad p, q \in \text{PES, NES}
\]

\[
= \delta_{\bar{q}} \delta_{\bar{p}} n_{\bar{q}},
\]

Although the introduction of CCC \((65)\) looks very trivial, it is a key ingredient in a time-independent Fock space formulation of relativistic quantum mechanics. In particular, it allows to construct\([19, 21]\) an effective QED (eQED) Hamiltonian in a bottom-up fashion (i.e., without ever recourse to QED, a time-dependent perturbation theory). In contrast, the standard contraction \((67)\), the time-independent analog of Eq. \((62)\), will result in wrong, nonrelativistic type of potential energy expressions even for relativistic operators. As an illustration, we look at the
number operator, which reads

\[ \hat{N} = a_p^\dagger + a_p + a_\tilde{p}^\dagger + a_\tilde{p} \quad (68) \]

\[ = \{a_p^\dagger\} + \{a_p\} + \langle 0; N_{e^-} | \frac{1}{2} [a_p^\dagger, a_p] | 0; N_{e^-} \rangle + \langle 0; N_{e^-} | \frac{1}{2} [a_\tilde{p}^\dagger, a_\tilde{p}] | 0; N_{e^-} \rangle \quad (69) \]

\[ = \{a_p^\dagger\} + \{a_p\} - \frac{1}{2} \delta_p + \frac{1}{2} \delta_\tilde{p} \quad (70) \]

\[ = \{a_p^\dagger\} + \{a_p\} \quad (71) \]

according to Eq. (65), but reads

\[ \hat{N} = \{a_p^\dagger\} + \{a_p\} + N_{e^-} \quad (72) \]

according to Eq. (67), with \( N_{e^-} \to \infty \) in line with the filled Dirac sea. It can readily be checked that the correct (71) and incorrect (72) results can also be obtained by using the contractions (65) and (67), respectively, in terms of the \( b \)-operators (29) and the associated vacuum \( |0\rangle \).

Having discussed pedagogically the basics of QED (including first quantization (19) of special relativity (2), second quantization of the Dirac field (27) or equivalently (28), extended hole theory, symmetrized 4-current (50), equal-time contraction (61), as well as charge-conjugated contraction (65)), we just comment briefly on the applications of QED. Undoubtedly, QED is the most accurate theory ever designed in physics. For instance, the anomalous magnetic moment, \( (g - 2)/2 \), of the electron has been determined to the 11th decimal place\(^{22}\), which leads further to improved values for the electron mass\(^{23}\) and the fine structure constant\(^ {24}\). However, the situation is very different for bound states of many-electron systems\(^ {25}\), for which QED is computationally too expensive: the more the electrons, the higher order of perturbation and hence the more Feynman diagrams are required to achieve high precision. Because of this, relativistic QED has thus far been applied successfully only to single ions of at most 5 electrons (see Refs. 26 and 27 for recent reviews). As for molecular systems, only nonrelativistic QED has been applied to the lightest molecules (e.g., \( \text{H}_2 \)\(^ {28}\), \( \text{D}_2 \)\(^ {29}\) and \( \text{HD} \)\(^ {30}\), etc.). So the question is how to account for QED effects in heavy atoms and molecules. To show relevance of this question, we just quote a few results here: (1) according to the rough estimates\(^ {31}\), the leading-order QED (Lamb shift) effects can be as large as 1 kcal/mol in chemical processes involving heavy elements. (2) According to the most recent and to date most accurate relativistic calculations\(^ {32}\) of the first ionization potential (IP) and electron affinity (EA) of the gold atom, QED effects are roughly the same as electron correlation beyond the gold standard CCSD(T) (coupled-cluster with singles and doubles and perturbative triples). (3) As for core properties such as the K-edge electron spectra, QED effects become significant already for the third row of
the periodic table[33]. It is therefore clear that we do need a feasible relativistic QED approach for the electronic structure and spectroscopies of heavy atoms and even molecules. It is also clear that, at variance with the “relativity-QED then correlation” paradigm of QED, we should think of something like “first relativity then correlation and finally QED”[34]. Such effective QED (eQED) approaches[19, 21, 35, 36] do exist, which will be discussed in Sec. III. Before this, we need to know how to solve the time-independent Dirac equation (21) via a finite basis expansion (see Sec. II). After having presented the eQED Hamiltonians in Sec. III, we will discuss in Sec. IV the correlation problem of NES (or virtual positrons) as well as a relativistic theory of real positrons. Sec. V is devoted to a summary of no-pair relativistic Hamiltonians, whereas Sec. VI to the no-pair correlation problem. The account will be closed with perspectives in Sec. VII.

II. THE MATRIX DIRAC EQUATION

For brevity, consider first the Dirac equation for an electron moving in a local potential $V$

$$
\begin{pmatrix}
V & c\sigma \cdot p \\
c\sigma \cdot p & V - 2c^2
\end{pmatrix}
\begin{pmatrix}
\psi^L_p \\
\psi^S_p
\end{pmatrix}
= 
\begin{pmatrix}
\psi^L_p \\
\psi^S_p
\end{pmatrix}
\epsilon_p,
$$

(73)

where the rest-mass energy $mc^2$ has been subtracted to align the energy scale to that of the Schrödinger equation. Early attempts[37] to solve this equation in a basis expansion were plagued by the occurrence of matrix eigenvalues in the forbidden region between the lowest positive-energy and the highest negative-energy operator eigenvalues. This phenomenon is usually called variational collapse and is often traced back to the lack of a lower-bound property of the Dirac operator. Actually, the “variational collapse” is due to the fact that inappropriately chosen basis sets are unable to describe the kinetic energy correctly and to guarantee the correct nrl[38]. It can be removed rigorously via the minimax principle[39, 40], without the need to impose a lower-bound property on the matrix representation of the Dirac equation (73). On the practical side, several prescriptions have been proposed to construct suitable basis sets, including restricted kinetic balance (RKB)[41], unrestricted kinetic balance (UKB)[42] [NB: the acronyms RKB and UKB were first coined by Dyall and Fægri Jr.[43]], dual kinetic balance (DKB)[44] and inverse kinetic balance (IKB)[45]. According to the thorough formal and numerical analyses[45], the following conclusions can be drawn:

(I) RKB is the least adequate condition for constructing the small-component spinor basis
\(\{f_\mu\}_{\mu=1}^{2N_L}\) directly from the large-component set \(\{g_\mu\}_{\mu=1}^{2N_L}\), viz.,

\[
f_\mu = \frac{\alpha}{2} \sigma \cdot p g_\mu, \quad \alpha = c^{-1}, \quad \mu = 1, \ldots, 2N_L.
\]

(74)

Why this is the case can best be understood in terms of the modified Dirac equation\[46, 47\]

\[
D^M \psi_p^M = S^M \psi_p^M \epsilon_p, \quad (75)
\]

\[
T_M = \begin{pmatrix} 1 & 0 \\ 0 & \frac{\alpha}{2} \sigma \cdot p \end{pmatrix}, \quad (76)
\]

\[
D^M = T_M^\dagger D T_M = \begin{pmatrix} V^T & T \\ T & \frac{\alpha^2}{4} \sigma \cdot p V \sigma \cdot p - T \end{pmatrix}, \quad (77)
\]

\[
S^M = T_M^\dagger T_M = \begin{pmatrix} 1 & 0 \\ 0 & \frac{\alpha^2}{2} T \end{pmatrix}, \quad (78)
\]

\[
\psi_p^M = T_M^{-1} \psi_p = \begin{pmatrix} \psi_p^L \\ \phi_p^L \end{pmatrix}, \quad (79)
\]

\[
\psi_p^S = \frac{\alpha}{2} \sigma \cdot p \phi_p^L. \quad (80)
\]

It has been proven\[48\] that the large \((\psi_p^L)\) and pseudo-large \((\phi_p^L)\) components must be expanded in the same spinor basis \(\{g_\mu\}\) in order to guarantee the correct nrl, a prerequisite to ensure that the energies of the PES are correct to \(O(c^{-2})\). Relation (80) then implies immediately the RKB (74). The expansion of \(\psi_p\) in the RKB basis (74) and that of \(\psi_p^M\) in the \(\{g_\mu\}\) basis, i.e.,

\[
\psi_p = \begin{pmatrix} \psi_p^L \\ \psi_p^S \end{pmatrix} = \begin{pmatrix} g_\mu A_{\mu p} \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ f_\mu B_{\mu p} \end{pmatrix}, \quad (81)
\]

\[
\psi_p^M = \begin{pmatrix} \psi_p^L \\ \phi_p^L \end{pmatrix} = \begin{pmatrix} g_\mu A_{\mu p} \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ g_\mu B_{\mu p} \end{pmatrix}, \quad (82)
\]

give rise to the same matrix Dirac equation

\[
\begin{pmatrix} V & T \\ T & \frac{\alpha^2}{4} W - T \end{pmatrix} \begin{pmatrix} A_p \\ B_p \end{pmatrix} = \begin{pmatrix} S & 0 \\ 0 & \frac{\alpha^2}{2} T \end{pmatrix} \begin{pmatrix} A_p \\ B_p \end{pmatrix} \epsilon_p, \quad (83)
\]
where the individual matrices are all of dimension $2N_L$, with the elements being

\begin{align*}
V_{\mu\nu} &= \langle g_\mu | V | g_\nu \rangle, \\
T_{\mu\nu} &= \langle g_\mu | \frac{p^2}{2} | g_\nu \rangle, \\
W_{\mu\nu} &= \langle g_\mu | \sigma \cdot p V | g_\nu \rangle, \\
S_{\mu\nu} &= \langle g_\mu | g_\nu \rangle.
\end{align*}

Eq. (83) is therefore of dimension $4N_L$ with $2N_L$ PES and $2N_L$ NES, which are separated by ca. $2mc^2 \approx 1$ MeV. When solving the equation (83) iteratively, the energetically lowest PES are chosen to be occupied in each iteration cycle, so as to avoid variational collapse. While the rotations between the occupied and unoccupied PES lower the total energy, those between the occupied PES and unoccupied NES raise the total energy, to a much lesser extent though.

The following points concerning RKB still deserve to be highlighted. (a) The RKB condition does not provide full variational safety, because the NES are in error of $O(c^0)$ \[45\]. Depending very much on the construction of the large-component basis, some bounds failures (or prolapse\[49\]) of $O(c^{-4})$ may occur. Nevertheless, such bounds failures will diminish when approaching to the basis set limit, at a rate that is not much different from the nonrelativistic counterpart\[50\]. (b) It turns out that the use of spherical Gaussians with principal quantum number $n$ larger than the angular momentum $l$ plus one leads to terrible variational collapse\[45\], although such functions are valid in the nonrelativistic case. Therefore, the use of spherical Gaussians subject to the restriction $n = l + 1$ (i.e., 1s, 2p, 3d, 4f, 5g, etc.) is not merely a matter of economy but also a must.

(II) IKB is the charge-conjugated version of RKB. It guarantees the correct nrl for the NES instead of the PES. Because of this, it requires basis functions that are very different from the standard ones and is therefore only of conceptual interest rather than of practical usage.

(III) DKB combines the good of both RKB and IKB and even provides full variational safety\[45\]. However, such an advantage is largely offset by its complicated nature and doubled number of integrals compared to RKB. It is therefore recommended only for calculations of tiny quantities (e.g., QED and parity non-conserving effects), where the complexity of DKB is only minor compared to the high precision to be achieved.

(IV) UKB is not uniquely defined. A scalar UKB basis does not transform as the basis of irreducible representations of double point groups or of time-reversal symmetry, a not
serious problem though. More problematic is that a UKB basis often suffers from severe linear dependence. Moreover, UKB does not offer a faster convergence to the basis set limit than RKB.

In short, RKB is the right choice for discretizing the Dirac equation in the absence of external magnetic fields. Since RKB is also a built-in condition for two-component relativistic theories[48], it should be regarded as a cornerstone of relativistic quantum chemistry. In the presence of external magnetic fields, RKB can be generalized to

\[
Z = \begin{pmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{pmatrix}, \tag{85}
\]

which leads to a most general expansion of \( \psi_p \),

\[
\psi_p = Z \tilde{\psi}_p, \quad \tilde{\psi}_p = \begin{pmatrix} g_\mu A_\mu p \\ g_\mu B_\mu p \end{pmatrix} = \begin{pmatrix} Z_{11} g_\mu A_\mu p + Z_{12} g_\mu B_\mu p \\ Z_{21} g_\mu A_\mu p + Z_{22} g_\mu B_\mu p \end{pmatrix}. \tag{86}
\]

Specific examples for the \( Z \) operator (85) can be found from Refs. 35 and 51 and are not repeated here.

### III. THE eQED HAMILTONIAN

Given the one-electron Dirac operator, the question is how to construct a relativistic many-electron Hamiltonian. The common practice is to add in simply the Coulomb interaction. Since the instantaneous Gaunt and Breit interactions can also be derived in a semiclassical manner[52], they can likewise be included, thereby leading to the Dirac-Coulomb-Gaunt/Breit
(DC/DCG/DCB) Hamiltonian

\[ H = \sum_{i=1}^{N} D(i) + \frac{1}{2} \sum_{i \neq j}^{N} V(r_{ij}), \quad (87) \]

\[ D = c \alpha \cdot p + (\beta - 1) mc^2 - \sum_{A}^{N_A} \frac{Z_{A}}{|R_{A} - r|}, \quad (88) \]

\[ V(r_{12}) = V_{C}(r_{12}) + V_{B}(r_{12}), \quad (89) \]

\[ V_{C}(r_{12}) = \frac{1}{r_{12}}, \quad (90) \]

\[ V_{B}(r_{12}) = V_{G}(r_{12}) + V_{g}(r_{12}), \quad (91) \]

\[ V_{G}(r_{12}) = -\frac{\alpha_{i} \cdot \alpha_{j}}{r_{12}}, \quad (92) \]

\[ V_{g}(r_{12}) = \frac{\alpha_{1} \cdot \alpha_{2}}{2r_{12}} - \frac{(\alpha_{1} \cdot r_{12})(\alpha_{2} \cdot r_{12})}{2r_{12}^{3}}. \quad (93) \]

On the formal side, the DCB Hamiltonian should be adopted as it is correct to \( O(\alpha^2) \), whereas so is neither DC nor DCG. Yet, on the practical side, the DCG Hamiltonian is more appealing for it describes all inter-electronic spin-same-orbit, spin-other-orbit, orbit-orbit, and spin-spin interactions of \( O(\alpha^2) \) and is computationally cheaper than DCB. That is, the difference between DCB and DCG is merely a scalar gauge term \( V_{g} \) that is of minor importance but leads to complicated integrals. As such, the DC and DCG Hamiltonians have been the major basis of relativistic quantum chemistry for molecular chemistry and physics. However, unlike the Schrödinger-Coulomb (SC) Hamiltonian that has well-defined mathematical and spectral properties, such ad hoc relativistic Hamiltonians have serious problems\textsuperscript{[16, 21]}. Without going into details, suffice it to say here that such first-quantized Hamiltonians violate a fundamental law of relativistic quantum mechanics, viz., it is the charge instead of the number of particles that is conserved. Therefore, it is pointless to solve the DC/DCG/DCB equation \( H\Psi = E\Psi \) exactly, unless one is interested in its mathematical solutions. Instead, to conserve the number of electrons, it is only consistent to adopt the no-pair approximation (NPA) from the outset, regardless of the existence of bound states or not. An immediate consequence is that the resulting energy \( E_{np} \) is not unique but is always dependent on how the projection operator is defined. Since the projector can only be defined in terms of the PES of some effective potential, it can be said that \( E_{np} \) is always potential dependent, a situation that is very different from the FCI (full configuration interaction) solution of the Schrödinger equation. Even though such ambiguity can largely be removed by optimizing the potential/projector at a correlated level (e.g., no-pair full multiconfiguration self-consistent field including orbital rotations to the unoccupied NES\textsuperscript{[53]}), how to account for the (dynamic) correlation of NES still remains to be resolved. This
requires a “with-pair relativistic Hamiltonian” in the first place.

As emphasized in the Introduction, the correct description of relativistic electrons must be done via second quantization. More specifically, it is the “extended hole theory”, the field Dirac picture coupled with charge conjugation, that is the proper tool for constructing many-electron relativistic Hamiltonians. To begin with, a primitive second quantization of the Dirac matter field can be introduced, viz.,

$$\hat{\phi}(r) = a_p \psi_p(r), \quad a_p |_{\text{vac}} = 0, \quad p \in \text{PES, NES},$$  (94)

where the spinors are eigenfunctions of the following effective Dirac equation

$$(D + U)\psi_p = e_p \psi_p,$$  (95)

with $U$ being some local or nonlocal screening potential. The term ‘primitive’ here means that this form of second quantization does not distinguish the empty from the filled Dirac picture. This gives rise to the following normal-ordered, second-quantized DC/DCG/DCB Hamiltonian

$$\mathcal{H} = D_p^q a^p_q + \frac{1}{2} g^{rs}_{pq} a_r^s, \quad p, q, r, s \in \text{PES, NES},$$  (96)

$$D_p^q = \langle \psi_p | D | \psi_q \rangle, \quad g^{rs}_{pq} = \langle \psi_p \psi_q | V(r_{12}) | \psi_r \psi_s \rangle,$$  (97)

$$a^p_q = a^p a^q, \quad a^p_{rs} = a^p a^q a^s a^r.$$  (98)

The filled Dirac picture can be realized in a finite basis representation by setting the Fermi level below the energetically lowest of the $\tilde{N}$ ($= N_e$) occupied NES. The physical energy of an $N$-electron state can be calculated as the difference between those of states $\Psi(N; \tilde{N})$ and $\Psi(0; \tilde{N})$,

$$E = \langle \Psi(N; \tilde{N}) | \mathcal{H} | \Psi(N; \tilde{N}) \rangle - \langle \Psi(0; \tilde{N}) | \mathcal{H} | \Psi(0; \tilde{N}) \rangle,$$  (99)

provided that the charge-conjugation symmetry is incorporated properly. To do so, we first shift the Fermi level just above the top of the NES. This amounts to normal ordering the Hamiltonian $\mathcal{H}$ (96) with respect to the non-interacting vacuum $|0; \tilde{N} \rangle$ ($= |0_e; N_e \rangle$) of zero positive energy electrons and $\tilde{N}$ negative-energy electrons. Here, the CCC (65) of fermion operators must be invoked, so as to obtain

$$a^p_q = \{a^p a^q\}_n + \langle 0; \tilde{N} | \frac{1}{2} [a^p, a^q] | 0; \tilde{N} \rangle, \quad p, q \in \text{PES, NES},$$  (100)

$$= \{a^p a^q\}_n - \frac{1}{2} \delta^p_q \text{sgn}(\epsilon_q), \quad p, q \in \text{PES, NES},$$  (101)

$$D_p^q a^p_q = D_p^q \{a^p a^q\}_n + C_{1n}, \quad C_{1n} = -\frac{1}{2} D^p_q \text{sgn}(\epsilon_p),$$  (102)
where the subscript \( n \) of the curly brackets emphasizes that the normal ordering is taken with respect to the reference \(|0; \tilde{N}\rangle\). More specifically,

\[
\{a^p a_q\}_n = \begin{cases} 
  a^p a_q, & \epsilon_p > 0, \epsilon_q > 0, \\
  a^p a_q, & \epsilon_p > 0, \epsilon_q < 0, \\
  a^p a_q, & \epsilon_p < 0, \epsilon_q > 0, \\
  -a_q a^p, & \epsilon_p < 0, \epsilon_q < 0.
\end{cases} \tag{103}
\]

By applying the relation (65) repeatedly we obtain

\[
a_{r s}^{p q} = \{a_{r s}\}_n - \frac{1}{2} \{\delta_r^p \delta_s^q \text{sgn}(\epsilon_r) + \delta_s^p \delta_r^q \text{sgn}(\epsilon_s) - \delta_r^p \delta_s^q \text{sgn}(\epsilon_r) - \delta_s^p \delta_r^q \text{sgn}(\epsilon_s)\}_n
\]

\[
+ \frac{1}{4} (\delta_r^p \delta_s^q - \delta_s^p \delta_r^q) \text{sgn}(\epsilon_r) \text{sgn}(\epsilon_s), \tag{104}
\]

and hence

\[
\frac{1}{2} g_{r s}^{p q} a_{r s}^{p q} = \frac{1}{2} g_{r s}^{p q} \{a_{r s}\}_n + Q_{p}^q \{a_{q}^p\}_n + C_{2n}, \tag{105}
\]

\[
Q_{p}^q = \tilde{Q}_{p}^q + \bar{Q}_{p}^q = -\frac{1}{2} g_{p s}^q \text{sgn}(\epsilon_s), \tag{106}
\]

\[
\bar{Q}_{p}^q = -\frac{1}{2} g_{p s}^q \text{sgn}(\epsilon_s), \tag{107}
\]

\[
\tilde{Q}_{p}^q = \frac{1}{2} g_{s q}^p \text{sgn}(\epsilon_s), \tag{108}
\]

\[
C_{2n} = \frac{1}{8} g_{p q}^s \text{sgn}(\epsilon_p) \text{sgn}(\epsilon_q) = -\frac{1}{4} Q_{p}^p \text{sgn}(\epsilon_p). \tag{109}
\]

Note that the implicit summations in \( C_{1n} \) (102), \( \tilde{Q} \) (107), \( \bar{Q} \) (108), and \( C_{2n} \) (109) include all the PES and NES, whether occupied or not. The Hamiltonian \( H \) (96) in the filled Dirac picture can then be written as

\[
H = H_a^{\text{QED}} + C_n, \tag{110}
\]

\[
H_a^{\text{QED}} = H_a^{\text{FS}} + Q_{p}^q \{a_{q}^p\}_n, \tag{111}
\]

\[
H_a^{\text{FS}} = D_{r}^p \{a_{r}^p\}_n + \frac{1}{2} g_{r s}^{p q} \{a_{r s}\}_n, \tag{112}
\]

\[
C_n = C_{1n} + C_{2n} = \langle 0; \tilde{N}|H|0; \tilde{N}\rangle = -\frac{1}{2} D_{p}^p \text{sgn}(\epsilon_p) - \frac{1}{4} Q_{p}^p \text{sgn}(\epsilon_p). \tag{113}
\]

\( H_a^{\text{QED}} \) (111) is just the desired “with-pair relativistic Hamiltonian” or simply effective QED (eQED) Hamiltonian, whereas \( H_a^{\text{FS}} \) is the so-called Fock space Hamiltonian advocated by Kutzelnigg [54] (see also Ref. [52], which missed, by construction, the vacuum polarization (VP) \( \tilde{Q} \) (107) (see Fig. 2(b)) and electron self-energy (ESE) \( \bar{Q} \) (108) (see Fig. 2(c)). If wanted,
FIG. 2. Diagrammatical representation of the one-photon exchange, vacuum polarization and electron self-energy

$H_{QED}^{(111)}$ can also be expressed in terms of the $b$-operators in view of the relations (29), viz.,

$$
H_{b}^{QED} = (D + Q)_{b}^{q \{ b^{\dagger} b^{\dagger} \}} + (D + Q)_{b}^{q \{ b_{b} b_{q} \}} + (D + Q)_{p}^{q \{ b_{b} b^{\dagger} \}} + (D + Q)_{p}^{q \{ b_{b} b^{\dagger} \}}
$$

$$
+ \frac{1}{4} \delta_{pq}^{rs} \{ b^{\dagger} b_{b} b_{b} b_{r} \} + \frac{1}{2} \delta_{pq}^{rs} \{ b^{\dagger} b_{b} b_{b} b_{r} \} + \frac{1}{2} \delta_{pq}^{rs} \{ b_{b} b_{b} b_{b} b_{r} \}
$$

$$
+ \frac{1}{2} \delta_{pq}^{rs} \{ b^{\dagger} b_{b} b^{\dagger} b^{\dagger} \} + \frac{1}{4} \delta_{pq}^{rs} \{ b_{b} b_{b} b^{\dagger} b^{\dagger} \},
$$

where the normal ordering is taken with respect to $|0\rangle$. Note that the eQED Hamiltonian (111)/(114) can also be obtained by a diagrammatical procedure[17, 19]. Had the standard contraction (67) of fermion operators been taken, we would obtain the following “configuration space” (CS) Hamiltonian

$$
H_{a}^{CS} = H_{a}^{FS} + \delta_{pq}^{ij} \{ a^{p}_{q} \} n.
$$

At variance with the $Q$ potential (106), the potential $\delta_{pq}^{ij}$ here arises from the occupied NES $\{ \psi_{j} \}$ alone (which is also a conventional interpretation of Fig. 2(b)). It is infinitely repulsive, leading to that no atom would be stable. This is of course plainly wrong.

Finally, the proper evaluation of the $Q$ potential (106) should be discussed. The Coulomb-only $\tilde{Q}$ term (107) is the full vacuum polarization[55] due to the polarization density $\rho_{vp}$ (55). In practice, it can be split into the Uehling [56] and Wichmann-Kroll [57] terms, which can then be evaluated with the analytic formulae[58, 59]. The ESE term $\tilde{Q}$ (108) is more difficult to handle. In addition to the Coulomb interaction, the transverse-photon contribution should
also be included. In the Coulomb gauge adopted here, the transverse part of the ESE reads
\[
(\bar{Q}_T)^q_p = \frac{1}{2} \langle p | \Sigma^C_T(\epsilon_p) + \Sigma^C_T(\epsilon_q) | q \rangle,
\]
\[
\langle p | \Sigma^C_T(\epsilon_p) | q \rangle = \langle p_s | \int_0^\infty \frac{dk f^C_T(k, r_1, r_2)}{\epsilon_p - \epsilon_s - (ck - i\gamma) \text{sgn}(\epsilon_s)} | sq \rangle,
\]
\[
f^C_T(k, r_1, r_2) = \frac{\sin(kr_{12})}{\pi r_{12}} [\alpha_1 \cdot \alpha_2 - \frac{(\alpha_1 \cdot \nabla_1)(\alpha_2 \cdot \nabla_2)}{k^2}] .
\]
Therefore, the total ESE (still denoted as \(\bar{Q}\)) can be written in a symmetric form
\[
Q^q_p = \frac{1}{2} \langle p | \Sigma^C + \Sigma^C_T(\epsilon_p) + \Sigma^C_T(\epsilon_q) | q \rangle, \quad \langle p | \Sigma^C | q \rangle = g^q_{ps} \text{sgn}(\epsilon_s).
\]
It has recently been shown that the ESE \((119)\) can be fitted into a simple and accurate semilocal model operator for each atom\[61, 62\]. Therefore, the VP-ESE (Lamb shift) can readily be included in the mean-field treatment, so as to account for screening effects on the VP-ESE automatically.

IV. APPLICATION OF THE eQED HAMILTONIAN

In this section, the occupied PES and NES are to be denoted respectively by \(\{i, j, \cdots\}\) and \(\{\tilde{i}, \tilde{j}, \cdots\}\), whereas the unoccupied PES by \(\{a, b, \cdots\}\). Unspecified orbitals are denoted as \(\{p, q, r, s\}\). When necessary, the NES will explicitly be designated by \(\{\tilde{p}, \tilde{q}, \tilde{r}, \tilde{s}\}\).

A. The second-order QED energy of an \(N\)-electron system

The eQED Hamiltonian \(H^\text{QED}_a (111)\) or \(H^\text{QED}_b (114)\) can be employed in the Bloch equation to determine the wave operators order by order. The resulting energy expressions are in full agreement with those obtained by the S-matrix formulation of QED\[21\]. However, the procedure treating all the PES as particles is rather involved. It is more expedite\[16\] to calculate the physical energy according to Eq. \((99)\) by treating the occupied PES also as holes. That is, to calculate the first term on the right-hand side of Eq. \((99)\), the eQED Hamiltonian \(H^\text{QED}_a (111)\) can further be normal-ordered with respect to the non-interacting reference \(|N; \tilde{N}\rangle\), the zero order of \(\Psi(N; \tilde{N})\). Since the normal ordering is now taken with respect to the occupied PES alone, the standard contraction of Fermion operators, e.g.,
\[
ap^p a^\dagger_q = \langle N; \tilde{N} | \{a^p a^\dagger_q\}_n | N; \tilde{N}\rangle = \langle N; 0 | a^p a^\dagger_q | N; 0 \rangle = \delta^p_n q, \quad \epsilon_q > 0,
\]
should be invoked. More specifically,

\[ D^q_p \{ a^p_q \}_n = D^q_p \{ a^p_q \}_F + D^i_p, \quad (121) \]

\[ Q^p_q \{ a^p_q \}_n = Q^p_q \{ a^p_q \}_F + Q^i_p, \quad (122) \]

\[ \frac{1}{2} g_{pq}^{rs} a^p_q a^p_q = \frac{1}{2} g_{pq}^{rs} a^p_q a^p_q F + (V_{HF})^q_p a^p_q F + \frac{1}{2} (V_{HF})^i_p. \quad (123) \]

Consequently, we have

\[ \mathcal{H} = H_F^{QED} + C_F, \quad (124) \]

\[ H_F^{QED} = f^q_p \{ a^p_q \}_F + \frac{1}{2} g_{pq}^{rs} a^p_q a^p_q F, \quad (125) \]

\[ f^q_p = (f^q_e)_p + Q^q_p, \quad (126) \]

\[ (f^q_e)_p = D^q_p + (V_{HF})^q_p, \quad (V_{HF})^q_p = g^{dij}_{pq}, \quad (127) \]

\[ C_F = C_n + E^{[1]} \quad (128) \]

\[ E^{[1]} = E_{np}^{[1]} + Q^i_p, \quad (129) \]

\[ E_{np}^{[1]} = (D + \frac{1}{2} V_{HF})^i_p. \quad (130) \]

To facilitate the use of many-body perturbation theory (MBPT) for electron correlation, the Hamiltonian (124) can further be partitioned as

\[ \mathcal{H} = H_{0A} + V_{0A} + V_{1A} + V_{2A}, \quad (131) \]

\[ H_{0A} = \epsilon_p \{ a^p_p \}_F + \sum_i \epsilon_i - \frac{1}{2} \epsilon_p \text{sgn}(\epsilon_p), \quad (132) \]

\[ V_{0A} = (Q - U + \frac{1}{2} V_{HF})^i_p + (\frac{1}{2} U - \frac{1}{4} Q)^p_p \text{sgn}(\epsilon_p), \quad (133) \]

\[ V_{1A} = (V_{1A})^q_p \{ a^p_q \}_F, \quad (V_{1A})^q_p = (Q + V_{HF} - U)^q_p, \quad (134) \]

\[ V_{2A} = \frac{1}{2} g_{pq}^{rs} a^p_q a^p_q F, \quad (135) \]

where the appearance of the counter potential

\[ -U^q_p \{ a^p_q \} = -U^q_p \{ a^p_q \}_n + \frac{1}{2} U^p_p \text{sgn}(\epsilon_p) \]

\[ = -U^q_p \{ a^p_q \}_F - U^i_p + \frac{1}{2} U^p_p \text{sgn}(\epsilon_p) \quad (136) \]

is due to the fact that the general mean-field equation (95) has been employed to determine the spinors and energy levels. As for the second term of Eq. (99), the Hamiltonian (110) can be
partitioned as
\[
    \mathcal{H} = H_{0B} + V_{0B} + V_{1B} + V_{2B},
\]
\[
    H_{0B} = \epsilon_p \{a^p_a\}_n - \frac{1}{2} \epsilon_p \text{sgn}(\epsilon_p),
\]
\[
    V_{0B} = (\frac{1}{2} U - \frac{1}{4} Q)^p_p \epsilon_p \text{sgn}(\epsilon_p),
\]
\[
    V_{1B} = (V_{1B})^q_p \{a^q_p\}_n, \quad V_{1B} = Q - U,
\]
\[
    V_{2B} = \frac{1}{2} s^r_s \{a^r_p\}_n.
\]

Following the standard MBPT, we obtain immediately
\[
    E^{(0)} = -\frac{1}{2} \epsilon_p \text{sgn}(\epsilon_p - \epsilon_F) + \frac{1}{2} \epsilon_p \text{sgn}(\epsilon_p) = \sum_i^N \epsilon_i,
\]
\[
    E^{(1)} = V_{0A} - V_{0B} = (\frac{1}{2} V_{HF} - U + Q)_i,
\]
\[
    E^{(2)} = E^{(2)}_1 + E^{(2)}_2,
\]
\[
    E^{(2)}_1 = \left( (V_{1A})^a_i (V_{1A})^a_i \right) \frac{\epsilon_i - \epsilon_a}{\epsilon_i - \epsilon_a} + \left( (V_{1A})^a_i (V_{1A})^a_i \right) \frac{\epsilon_i - \epsilon_a}{\epsilon_i - \epsilon_a}
\]
\[
    - \left( (V_{1B})^a_i (V_{1B})^a_i \right) \frac{\epsilon_i - \epsilon_i}{\epsilon_i - \epsilon_i},
\]
\[
    E^{(2)}_2 = (V_{HF} - U)_i Q^i_a + Q^a_i (V_{HF} - U)_i + Q^a_i Q^i_a
\]
\[
    + \frac{(V_{HF})^i_a Q^i_a + Q^a_i (V_{HF})^i_a}{\epsilon_i - \epsilon_a} - \frac{Q^i_a Q^i_a - U^i_a Q^i_a - Q^i_a U^i_a}{\epsilon_i - \epsilon_i},
\]
\[
    E^{(2)}_2 = \frac{1}{4} \frac{\delta^a_b}{\delta^m_n \delta^a_b} \frac{\delta^m_n \delta^a_b}{m,n=i,j,a,b} - \frac{1}{4} \frac{\delta^p_q \delta^i_j}{\delta^p_q \delta^i_j} \frac{\delta^p_q \delta^i_j}{p,q=i,j,a,b}
\]
\[
    = \begin{pmatrix}
        \frac{1}{4} \frac{\delta^a_b \delta^i_j}{\delta^a_b \delta^i_j} + \frac{1}{2} \frac{\delta^a_b \delta^i_j}{\delta^a_b \delta^i_j} \\
        \frac{1}{4} \frac{\delta^i_j \delta^i_j}{\delta^i_j \delta^i_j} + \frac{1}{2} \frac{\delta^i_j \delta^i_j}{\delta^i_j \delta^i_j}
    \end{pmatrix}.
\]
FIG. 3. Diagrammatical representation of the second-order QED energy. (a) antisymmetrized two-body; (b) one-body. For the $\Psi(N; \tilde{N})$ state, the particles (up-going lines) and holes (down-going lines) are $\{a, b\}$ and $\{i, j, \tilde{i}, \tilde{j}\}$, respectively, and the one-body potential represented by the square is $V_{1A}$. For the $\Psi(0; \tilde{N})$ state, the particles and holes are $\{a, b, i, j\}$ and $\{\tilde{i}, \tilde{j}\}$, respectively, and the one-body potential is $V_{1B}$. A global negative sign should be inserted to the terms of $\Psi(0; \tilde{N})$.

The first and second terms of $E_1^{(2)}$ (145) and $E_2^{(2)}$ (149) arise from the $\Psi(N; \tilde{N})$ and $\Psi(0; \tilde{N})$ states, respectively. The one-body $E_1^{(2)}$ (145) can further be decomposed into two terms, $E_{FS, 1}^{(2)}$ (147) and $E_{Q, 1}^{(2)}$ (148). Both $E_{FS, 1}^{(2)}$ (147) and $E_2^{(2)}$ (150) arise from the Fock space Hamiltonian $H_{aFS}^{\text{FS}}$ (112), while $E_{Q, 1}^{(2)}$ (148) is due exclusively to the VP and ESE [NB: $E_{Q, 1}^{(2)}$ and the Q term in $E^{(1)}$ will not show up if the Q potential is included in the mean-field equation (73)]. The two terms of $E^{(2)}$ (144) can be represented by the same Goldstone-like diagrams shown in Fig. 3. It is just that the particles and holes, as well as the one-body potential, are interpreted differently.

Note that the frequency-dependent Breit interaction

$$V_T(\omega, r_{12}) = -\alpha_1 \cdot \alpha_2 \frac{\cos(|q|r_{12})}{r_{12}}, \quad \omega = qc$$

$$+ [(\alpha_1 \cdot \nabla_1), [(\alpha_2 \cdot \nabla_2), \frac{\cos(|q|r_{12}) - 1}{q^2 r_{12}}]]$$  \hspace{1cm} (151)

must be employed to account for the contribution of NES to correlation, which would otherwise be severely overestimated if the frequency dependence is neglected [63]. This amounts to replacing the integrals $g_{pq}^{rs}$ in $E_2^{(2)}$ (150) with

$$g_{pq}^{rs} = \langle pq | V_C(r_{12}) + \frac{1}{2} V_T(\epsilon_r - \epsilon_p, r_{12}) + \frac{1}{2} V_T(\epsilon_s - \epsilon_q, r_{12}) | rs \rangle.$$  \hspace{1cm} (152)

Here we should recall again the most recent and to date most accurate relativistic calculations [32] of the IP and EA of Au: the calculated IP (9.2288 eV) deviates from the experimental one (9.2256 eV) somewhat larger than the corresponding EA (calculated 2.3072 vs. experimental 2.3086 eV).
This is counterintuitive, since IPs can usually be calculated more accurately than EAs. There could be two possible reasons for this: (a) the no-pair correlation still need to be improved and (b) the missing contribution of NES to correlation has to be taken into account. Even if the contribution of NES to correlation is not the reason for such discrepancy, it is certainly important for core properties.

It also deserves to be mentioned that the same $E_2^{(2)}$ would correspond to the seven Feynman diagrams shown in Fig. 4 if the occupied PES are to be treated as particles instead of holes as in diagram Fig. 3(a). The first three and the next four of these Feynman diagrams are usually called non-radiative and radiative contributions in QED but all of which are contributions to electron correlation in the present context. It is now clear that only the $Q$ potential entering the eQED Hamiltonian arises from the unconventional contraction of fermion operators, whereas the treatment of electron correlation follows standard many-body theories in conjunction with the filled Dirac picture. This is because the $n$-th order correlation energy $E^{(n)} = \langle \text{vac}|V\Omega^{(n-1)}|\text{vac}\rangle$ arises from the (full) contraction between the first-order fluctuation potential $V$ and the $(n-1)$th-order wave operator $\Omega^{(n-1)}$ that are already normal ordered separately and hence originate from “different times”. In the parlance of diagrams, all the terms herein refer to reducible Feynman diagrams (see Fig. 4). In other words, only those irreducible Feynman diagrams that go beyond the eQED Hamiltonian (which is, by definition, linear in the two-particle interaction involving only one-photon exchange) must be treated via full QED, a time-dependent perturbation theory.

Finally, it is instructive to compared the second-order QED energy $E^{(2)}$ with that of the configuration space approach

$$E^{(2)}_{\text{CS}} = E^{(2)}_{\text{CS,1}} + E^{(2)}_{\text{CS,2}},$$

$$E^{(2)}_{\text{CS,1}} = \frac{(V_{HF} - U)^i_\alpha(V_{HF} - U)^i_\beta}{\epsilon_i - \epsilon_\alpha} + \frac{(V_{HF} - U)^i_\beta(V_{HF} - U)^i_\alpha}{\epsilon_i - \epsilon_\beta},$$

$$E^{(2)}_{\text{CS,2}} = \frac{1}{4} \epsilon_i + \frac{\delta^{ab}}{\delta_{ij}} \frac{\delta^{ij}}{\delta_{ab}} + \frac{1}{4} \epsilon_j + \frac{\delta^{ij}}{\delta_{ab}} \frac{\delta^{ij}}{\delta_{ab}} + \frac{1}{2} \epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b.$$
It is seen that \( E_{CS,1}^{(2)} \) and \( E_{CS,2}^{(2)} \) agree respectively with \( E_{FS,1}^{(2)} \) and \( E_{2}^{(2)} \) only in the first terms involving solely the PES, but are very different from the latter in the terms involving the NES. In particular, the denominator of the last term of Eq. (155) can be zero (e.g., \( \epsilon_a = \epsilon_i + |X| \) and \( \epsilon_f = \epsilon_j - |X| \)). Since there exists an infinite number of such “+−” intermediates, this problem has been termed continuum dissolution\([64]\). As shown here, it is purely an artefact due to the underlying empty Dirac picture. As such, only no-pair projected wave functions are acceptable in the configuration space formulation. Efforts\([65]\) to solve exactly the DC/DCG/DCB equation \( H \Psi = E \Psi \) are then purely mathematical exercises. Since there is no analytic Hamiltonian in Fock space as well, the term “exact (analytic) relativistic wave function” is simply meaningless\([66]\).

**B. Mean-field theory of real positrons**

As an another application of the eQED Hamiltonian (111), we present here a mean-field theory for a system of \( N \) electrons and \( \tilde{M} \) positrons. The energy up to first order reads

\[
E_{ep}^{[1]} = \langle N; \tilde{N}| A^\dagger H_{a}^{\text{QED}} A |N; \tilde{N}\rangle, \quad A = \Pi_{i}^{\tilde{M}} \tilde{a}_{i} \\
= \left[ \sum_{i=1}^{N} (D + Q)_{i} + \frac{1}{2} \sum_{i,j=1}^{N} g^{ij}_{ij} \right] + \left[ - \sum_{i=1}^{\tilde{M}} (D + Q)_{i} + \frac{1}{2} \sum_{i,j=1}^{\tilde{M}} \tilde{g}^{ij}_{ij} \right] \\
- \sum_{i=1}^{\tilde{M}} \sum_{j=1}^{N} \tilde{g}_{ij}^{ij}, \tag{157}
\]

Use of Wick’s theorem for expressing products of normal-ordered operators as a linear combination of contracted ones has been made when going from Eq. (156) to (157). The first and second terms are the average energies of the \( N \) electrons and of the \( \tilde{M} \) positrons, respectively, whereas the third, cross term represents their mutual interaction. The negative sign in the second and third terms results from the normal ordering (103) implicit in \( H_{a}^{\text{QED}} \) (111), and can be understood as a negative occupation number \( (n_{i} = -1) \) of the hole arising from the ionization \( a_{i} |N; \tilde{N}\rangle \). The other occupied NES not involved in the ionization \( \Pi_{i}^{\tilde{M}} a_{i} |N; \tilde{N}\rangle \) have been normal-ordered away, and can therefore be viewed as unoccupied, just like the unoccupied PES. As such, the expression (157) can be written as

\[
E_{ep}^{[1]} = \sum_{k} n_{k} (D + Q)^{k}_{k} + \frac{1}{2} \sum_{k,l} n_{k} n_{l} \tilde{g}^{kl}_{kl}, \quad k, l \in \text{PES, NES}, \tag{158}
\]

by assigning an occupation number \( n_{k} \) to each orbital \( \psi_{k} \): \( n_{k} \) is zero for the unoccupied PES and NES, +1 for the \( N \) occupied PES, and −1 for the \( \tilde{M} \) occupied NES. Formally, this agrees
with the empty Dirac picture. However, such agreement between the empty and filled Dirac pictures holds only at the mean-field level but not at the correlated level (see Sec. IV A). More generally, such agreement holds for all one-body but not for any two-body operators[16].

To minimize the energy $E_{e_p}^{[1]}$ subject to the orthonormal conditions, we can introduce the following canonical Lagrangian

$$L = E_{e_p}^{[1]} - \sum_k n_k \langle \psi_k | \psi_k \rangle - 1 | \epsilon_k, \quad k \in \text{PES, NES}. \quad (159)$$

The condition $\frac{\delta L}{\delta \psi_i^\dagger} = 0$ then gives rise to

$$fn_i | \psi_i \rangle = \epsilon_i n_i | \psi_i \rangle, \quad i \in \text{PES, NES}, \quad (160)$$

where

$$f = D + Q + \sum_k n_k g^{kj}, \quad k \in \text{PES, NES}, \quad (161)$$

$$f_p = (D + Q)^q_p + \sum_k n_k g^{kp}, \quad k \in \text{PES, NES}. \quad (162)$$

As it stands, Eq. (160) determines only the occupied PES and NES but which can be extended to the unoccupied ones (which are arbitrary anyway), viz.,

$$f | \psi_p \rangle = \epsilon_p | \psi_p \rangle, \quad p \in \text{PES, NES}. \quad (163)$$

The energetically lowest PES and highest NES are to be occupied in each iteration.

Some remarks are in order: (a) the cross, exchange term $-\sum_{ij} g_{ij}$ vanishes in the nrl, meaning that electrons and positrons are distinguishable particles in the nonrelativistic world, such that their mutual anti-symmetrization is no longer required. In other words, only QED treats electrons and positrons on an equal footing. (b) If the VP-ESE term $Q$ is neglected, the present mean-field theory of electrons and positrons will reduce to that formulated by Dyall in a different way[67].

V. NO-PAIR RELATIVISTIC HAMILTONIANS

There have been a number of comprehensive reviews[35, 48, 68–71] on the no-pair relativistic Hamiltonians. Therefore, only a brief summary of the essentials is necessary here. The no-pair relativistic Hamiltonians can be classified into four-component (4C), quasi-four-component (Q4C) and two-component (2C) ones, the last of which can further be classified into approximate (A2C) and exact (X2C) two-component ones.
A. Four-component

First of all, confining the orbital indices of $H_+^{\text{QED}}$ only to PES leads to the following no-pair QED Hamiltonian

$$H_+^{\text{QED}} = D_p^q a_p^q + Q_p^q a_p^q + \frac{1}{2} g_{pq}^{rs} a_p^q a_s^r, \quad p, q, r, s \in \text{PES}$$

(164)

$$= (f_e + Q) p^q (a_p^q)_{\text{N}} + \frac{1}{2} g_{pq}^{rs} (a_p^q)_{\text{N}} + E^{[1]}, \quad p, q, r, s \in \text{PES}$$

(165)

with $(f_e)_p^q$ and $E^{[1]}$ defined in Eqs. (127) and (129), respectively. Here, the subscript N indicates that the normal ordering is taken with respect to $|N; \tilde{0}\rangle$. The $H_+^{\text{QED}}$ Hamiltonian (164), along with $\tilde{Q}$ (107), $\bar{Q}$ (119) and $g_{pq}^{rs}$ (152), was already obtained by Shabaev [36] but in a top-down fashion. The aforementioned potential dependence in the calculated energies can be removed by introducing the following correction [72]

$$E^{(2)}_{PC} = \frac{(V_{HF})_i^q U_i^q + U_i^q (V_{HF})_i^q - U_i^q U_i^q}{\epsilon_i - \epsilon_i},$$

(166)

where $U$ is the potential in Eq. (73). One then has a potential-independent no-pair QED (PI-QED) Hamiltonian [21]

$$H_+^{\text{PI-QED}} = (f_e + Q - U) p^q (a_p^q)_{\text{N}} + \frac{1}{2} g_{pq}^{rs} (a_p^q)_{\text{N}}$$

$$+ E^{[1]} + E^{(2)}_{PC}, \quad p, q, r, s \in \text{PES}.$$  

(167)

Neglecting the $Q$ term in $H_+^{\text{PI-QED}}$ leads to the potential-independent no-pair DCB (PI-DCB) Hamiltonian

$$H_+^{\text{PI-DCB}} = (f_e - U) p^q (a_p^q)_{\text{N}} + \frac{1}{2} g_{pq}^{rs} (a_p^q)_{\text{N}} + E^{[1]}_{\text{NP}} + E^{(2)}_{PC}, \quad p, q, r, s \in \text{PES}.$$  

(168)

Further neglecting the $Q$ term in $H_+^{\text{PI}}$ leads to the standard no-pair DCB Hamiltonian

$$H_+^{\text{DCB}} = (f_e) p^q (a_p^q)_{\text{N}} + \frac{1}{2} g_{pq}^{rs} (a_p^q)_{\text{N}} + E^{[1]}_{\text{NP}}, \quad p, q, r, s \in \text{PES},$$

(169)

which has been the basis of “no-pair relativistic quantum chemistry”.

B. Quasi-four-component

The previous no-pair four-component approaches first generate both PES and NES at the mean-field level but then discard the NES at a correlated level. The question is how to avoid the NES from the outset. Actually, this can be done in two different ways. One is to retain the aesthetically simple four-component structure but freeze the NES, while the other is to remove the
FIG. 5. The $R_i(\mathbf{r})$ operator \[ (171) \] with $V = V_N + V_H + V_{LDA}$ as a function of the distance from the position of Rn. The radial expectation values of $1s_{1/2}$, $2s_{1/2}$, $2p_{1/2}$, $2p_{3/2}$, and $3s_{1/2}$ are 0.015, 0.063, 0.051, 0.060, and 0.163 a.u., respectively.

NES so as to obtain a two-component approach. While the former employs the untransformed Hamiltonian and introduces approximations from the very beginning, the latter invokes an effective Hamiltonian and has to introduce suitable approximations at a later stage. Note that in each case the approximations introduced to the Hamiltonians are orders of magnitude smaller than other sources of errors (e.g., incompleteness in the one- and many-particle bases) and are therefore hardly “approximate”. Moreover, since the two paradigms stem from precisely the same physics, they should be made fully equivalent.

To realize the first paradigm, we first take a look at the $S/L$ ratio between the small and large components of a PES $\psi_i$, which can be obtained from the second row of Eq. (73)

\[ \psi_i^S = \frac{\alpha}{2} R_i \sigma \cdot p \psi_i^L, \quad (170) \]

\[ R_i(\mathbf{r}) = \left[ 1 + \frac{\alpha^2}{2} (\epsilon_i - V(\mathbf{r})) \right]^{-1} \rightarrow 1. \quad (171) \]

The major effect of $\sigma \cdot p$ is to change the parity of the large component to that of the small component. So the $S/L$ ratio is determined mainly by the $R_i(\mathbf{r})$ operator \[ (171) \]. As can be seen from Fig. 5 the effect of $R_i(\mathbf{r})$ is extremely short ranged: each $R_i(\mathbf{r})$ becomes just a constant factor beyond a small radius $r_c$ (ca. 0.05 a.u., roughly the radii of $2s$ and $2p$). Imagine we have first solved the (radial) Dirac equation for each isolated (spherical and unpolarized) atom and
thus obtained the corresponding atomic 4-spinors (A4S) \{\varphi_\mu\}. Then, the atoms are brought together to synthesize the molecule. While both the large and small components of \varphi_\mu will change, the \(S/L\) ratio will not\[73]\[74]. The mathematical realization\[75] of such a physical picture is to expand the molecular 4-spinors (M4S) \(\psi_i\) in the basis only of positive-energy A4S \{\varphi_{+,\mu}\}, viz.,

\[
|\psi_i\rangle = \sum_\mu |\varphi_{+,\mu}\rangle C_{\mu i} = \sum_\mu \left( \begin{array}{c} |\varphi_{+,\mu}^L\rangle \\ |\varphi_{+,\mu}^S\rangle \end{array} \right) C_{\mu i}, \quad (172)
\]

which gives rise to the following projected four-component (P4C) approach\[75]\n
\[
h_{+}^{P4C} C = S_{+}^{P4C} C e,
\]

\[
(h_{+}^{P4C})_{\mu\nu} = \langle \varphi_{+\mu}^L | V | \varphi_{+\nu}^{L}\rangle + \langle \varphi_{+\mu}^S | c \sigma \cdot p | \varphi_{+\nu}^{L}\rangle 
+ \langle \varphi_{+\mu}^L | c \sigma \cdot p | \varphi_{+\nu}^{S}\rangle + \langle \varphi_{+\mu}^S | V - 2mc^2 | \varphi_{+\nu}^{S}\rangle, \quad (174)
\]

\[
(S_{+}^{P4C})_{\mu\nu} = \langle \varphi_{+\mu}^L | \varphi_{+\nu}^{L}\rangle + \langle \varphi_{+\mu}^S | \varphi_{+\nu}^{S}\rangle. \quad (175)
\]

The dimension of \(h_{+}^{P4C}\) is \(2NL\) instead of \(4NL\). That is, the molecular NES are excluded completely. Physically, this amounts to neglecting rotations between the PES and NES of the isolated atoms, a kind of polarization on the atomic vacuum induced by the molecular field. As molecular formation is a very-low energy process, its \(O(c^{-4})\) perturbation on the vacuum introduces no discernible errors at all\[73]\[75]\[76]. By further introducing a “model small component approximation” (MSCA), a quasi-four-component (Q4C) approach\[76] can be obtained, which is four-component in structure but is computationally very much like a two-component approach. Without going into further details (see Refs. 48 and 71 for the matrix elements \(f_{pq}^{Q4C}\) of \(f_{+}^{Q4C}\)), we now have the following second-quantized, normal-ordered many-electron Hamiltonian

\[
H_{+}^{Q4C} = E_{ref}^{Q4C} + f_{pq}^{Q4C} \{a_p^p\} + \frac{1}{2} g_{pq}^{rs} \{a_p^p a_r^r\}. \quad (176)
\]

Q4C shares precisely the same integral transformation and correlation treatment as two-component approaches\[48]\[71], but does not suffer from picture-change errors (PCE)\[77] which otherwise plague two-component approaches. Moreover, the model spectral form \[61]\[62] of the Q potential \(\langle 106\rangle\) can readily be incorporated into \(f_{pq}^{Q4C}\), thereby leading to an QED@Q4C approach.
C. Two-component

By definition, a two-component relativistic theory is to transform away the positronic degrees of freedom of the Dirac operator, so as to obtain a Hamiltonian that describes only electrons. This can be done with either unitary transformation or elimination of the small component (ESC). However, neither route can be done in closed form, except for the trivial free-particle case. As such, only approximate two-component (A2C) operator (analytic) Hamiltonians such as the Breit-Pauli Hamiltonian and the zeroth-order regular approximation (ZORA)\cite{78, 79} can be obtained in this way. The situation is changed dramatically when going to the matrix formulation, where the exact decoupling is readily achieved. In essence, the matrix formulation amounts to block-diagonalizing the matrix Dirac equation (83), which can be done in one step\cite{80–83}, two steps\cite{84–86} and multiple steps\cite{87–89}. The three types of formulations share the same decoupling condition and differ only in the renormalization\cite{48}. There exist even closed mapping relations among three formulations\cite{48}. Since the initio free-particle transformation invoked in the two-step and multiple-step formulations is only necessary for finite orders\cite{90, 91} but not for infinite order, it is clear that it is the one-step formulation that should be advocated. This approach has been coined “exact two-component” (X2C)\cite{92}. For generality, we extend Eq. (83) to a generic eigenvalue problem

\[
\mathbf{h} \mathbf{C} = \mathbf{M} \mathbf{E}, \quad (177)
\]

and

\[
\mathbf{h} = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} = \mathbf{h}^\dagger, \quad \mathbf{M} = \begin{pmatrix} S_{11} & 0 \\ 0 & S_{22} \end{pmatrix} = \mathbf{M}^\dagger, \quad \mathbf{C} = \begin{pmatrix} A_+ & A_- \\ B_+ & B_- \end{pmatrix}. \quad (178)
\]

To decouple the PES and NES, we first introduce the following formal relations

\[
B_+ = \mathbf{X} A_+, \quad A_- = \tilde{\mathbf{X}} B_-, \quad (179)
\]

between the small- and large-component coefficients for the PES and NES, respectively. The following unitary transformation matrix \(\mathbf{U}_X\) can then be introduced\cite{48}

\[
\mathbf{U}_X = \Omega_N \Omega_D, \quad \Omega_N = \begin{pmatrix} R_+^\dagger & 0 \\ 0 & R_-^\dagger \end{pmatrix}, \quad \Omega_D = \begin{pmatrix} \mathbf{I} & \mathbf{X}^\dagger \\ \tilde{\mathbf{X}}^\dagger & \mathbf{I} \end{pmatrix}, \quad (180)
\]

where\cite{81}

\[
R_+ = (S_{11}^J S_+)^{-\frac{1}{2}} = S_{11}^{-\frac{1}{2}} (S_{11}^{-\frac{1}{2}} S_+ S_{11}^{-\frac{1}{2}})^{-\frac{1}{2}} S_{11}^J, \quad (181)
\]

\[
R_- = (S_{22}^J S_-)^{-\frac{1}{2}} = S_{22}^{-\frac{1}{2}} (S_{22}^{-\frac{1}{2}} S_- S_{22}^{-\frac{1}{2}})^{-\frac{1}{2}} S_{22}^J, \quad (182)
\]

\[
\hat{S}_+ = S_{11} + \mathbf{X}^\dagger S_{22} \mathbf{X}, \quad (183)
\]

\[
\hat{S}_- = S_{22} + \tilde{\mathbf{X}}^\dagger S_{11} \tilde{\mathbf{X}}. \quad (184)
\]
The requirement that \( U_X \mu_{X} \dagger = M \) leads to

\[
\tilde{X} = -S_{11}^{-1}X^\dagger S_{22},
\]

(185)

meaning that \( \tilde{X} \) is determined directly by \( X \), which is further determined by \( (U_X hU_X^\dagger)_{21} = 0 \), viz.,

\[
h_{21} + h_{22}X = S_{22}X S_{11}^{-1}L_{UESC}^+, \quad L_{UESC}^+ = h_{11} + h_{12}X,
\]

(186)

\[
h_{22}X = S_{22}X S_{11}^{-1}L_{NESC}^+.
\]

(187)

The \( U_X \)-transformation of Eq. (177) then yields

\[
(U_X hU_X^\dagger)C_X = \begin{pmatrix} f_{X2C}^+ & 0 \\ 0 & f_{X2C}^- \end{pmatrix} C_X = MC_X E,
\]

(188)

\[
C_X = (U_X^\dagger)^{-1}C = M^{-1}U_X MC = \begin{pmatrix} C_+ & 0 \\ 0 & C_- \end{pmatrix}.
\]

(189)

The upper-left block of Eq. (188) defines the equation for the PES,

\[
\begin{align*}
f_{X2C}^+ C_+ &= S_{11} C_+ E_+, \\
f_{X2C}^- &= R_{+}^\dagger L_{+}^X R_+, \quad X = \text{NESC, SESC}, \\
L_{+}^NESC &= h_{11} + h_{12}X + X^\dagger h_{21} + X^\dagger h_{22}X,
\end{align*}
\]

(190)

(191)

(192)

Here the acronyms UESC, NESC and SESC refer to the unnormalized, normalized\cite{footnote_unnormalized} and symmetrized\cite{footnote_symmetrized} eliminations of the small component, respectively. Eq. (187) arises from Eq. (186) via the relation \( S_{11}^{-1}L_{UESC}^+ = S_{+}^{-1}L_{NESC}^+ \) (because \( L_{UESC}^+ A_+ = S_{11} A_+ E_+ \) and \( L_{NESC}^+ A_+ = S_{+} A_+ E_+ \)), whereas Eq. (193) arises from \( L_{+}^{\text{ESC}} = \frac{1}{2}(L_{UESC}^+ + L_{NESC}^+) \) and the decoupling condition (186). Likewise, the lower-right block of Eq. (188) defines the equation for the NES,

\[
\begin{align*}
f_{X2C}^- C_- &= S_{22} C_- E_-,
\end{align*}
\]

(193)

(194)

\[
\begin{align*}
f_{X2C}^+ &= R_{-}^\dagger L_{-}^X R_-, \quad X = \text{NESC, SESC}, \\
L_{-}^NESC &= h_{22} + h_{21}X + X^\dagger h_{12} + X^\dagger h_{11}X,
\end{align*}
\]

(195)

(196)

(197)

(198)

\[
\begin{align*}
L_{-}^{\text{ESC}} &= \frac{1}{2}(S_{-} S_{22}^{-1} L_{UESC}^+ + c.c.), \quad L_{UESC}^+ = h_{22} + h_{21}X, \\
C_- &= R_{-}^{-1} B_-.
\end{align*}
\]

(199)

It can be proven\cite{footnote_proven} that \( C_+ (C_-) \) is closest to \( A_+ (B_-) \) in the least-squares sense.
The above manipulation can further be extended to include magnetic fields as well\[35, 71\]. Moreover, at variance with the explicit expression (191), $\hat{\mathcal{F}}_{X^2C}$ can also be constructed on the fly, by an orthonormalization and back-transformation procedure\[83\]. The following remarks are still in order.

(1) The one-step matrix formulation of two-component relativistic theories was initiated by Dyall\[93\] in 1997. However, the proper formulation of the (energy-independent) decoupling condition (186)/(187)\[80\] as well as the correct renormalization (181)\[81\] were found only later on. It was also found\[48\] that the same results can be obtained by converting the Foldy-Wouthuysen (FW) Hamiltonian \[95\] (which has no closed form though) directly into matrix form in terms of the RKB basis. That is, the matrix and operator (more precisely operator-like) formulations of $X^2C$ are identical, as should be. The situation is different for finite-order $A^2C$ approaches. To see this, we look at the ZORA equation\[78, 79\],

$$ (V + T^{ZORA})\psi^ZORA_p = \psi^ZORA_p \epsilon^ZORA_p, \quad (200) $$

$$ T^{ZORA} = \sigma \cdot p - \frac{1}{2 - \alpha^2 V} \sigma \cdot p. \quad (201) $$

In view of the identity $1/(2 - \alpha^2 V) \times (2 - \alpha^2 V) = 1$, the matrix elements of $T^{ZORA}$ can be calculated as

$$ \langle \sigma \cdot p g_\mu | \frac{1}{2 - \alpha^2 V} | \sigma \cdot p g_\rho \rangle [(2T)^{-1}]_{\rho\nu} \langle \sigma \cdot p g_\nu | 2 - \alpha^2 V | \sigma \cdot p g_\sigma \rangle = 2T_{\mu\nu}, \quad (202) $$

which leads to

$$ T^{ZORA} = T^{XZORA}, \quad X^{ZORA} = (T - \frac{\alpha^2}{4} W)^{-1} T. \quad (203) $$

Therefore, the matrix representation of the ZORA equation (200) reads

$$ (V + T^{XZORA})A^{ZORA} = S A^{ZORA} e^{ZORA}. \quad (204) $$

Thanks to the use of the resolution of the identity (RI) in terms of the $\{\sigma \cdot p g_\mu\}$ basis in Eq. (202), the matrix ZORA equation (204) agrees with the operator ZORA equation (200) only when the basis $\{g_\mu\}$ is complete. This is totally different from the matrix counterpart (i.e., $X^2C$) of the (non-expanded) FW Hamiltonian \[95\], where the use of the same RI is not an approximation but only a formal step\[48\]. Since Eq. (204) is never used in practice, the commonly called ZORA (and the infinite-order regular approximation\[96\]) is a genuine analytic relativistic theory. In contrast, other relativistic theories, whether finite-order\[90, 91\]
or infinite-order\cite{80,89}, are all algebraic. However, the analyticity of a relativistic Hamiltonian should not be celebrated, simply because only Fock space is the right framework for relativistic quantum mechanics, which gives rise to only algebraic relativistic many-electron Hamiltonians (see Sec. III).

(2) The eigenvalue equation (190) and the decoupling condition (186)/(187) are coupled and have to be solved iteratively\cite{80,82}. The so-obtained results agree with those by the parent matrix Dirac equation (83) up to machine accuracy, thereby justifying the name “exact two-component”\cite{92}. However, the computation is much more expensive than solving Eq. (83) directly, even for a one-electron system. Therefore, a suitable approximation to $X$ must be found in order to make X2C practical. To this end, we take a look at the matrix presentation of the key relation (170) in a RKB basis (74) (without caring for the inherent singularities\cite{97}), viz.,

\[
B_{+,i} = \frac{1}{2} T^{-1} R_{+,i} \quad R_{\mu\nu}^{(i)} = \langle g_{\mu} | \sigma \cdot p R_{i} | \sigma \cdot p | g_{\nu} \rangle,
\]

(205)

\[
= U^{(i)} A_{+,i},
\]

(206)

where $U^{(i)}$ is the energy/state-dependent equivalent\cite{93} of the state-universal $X$. Since the $R_{i}(r)$ operator is extremely shorted ranged (see Fig. (5)), it can be envisaged that the molecular $U^{(i)}$ (and hence $X$) should be strongly block-diagonal in atoms. As can be seen from Fig. 6, this is indeed the case. Note in particular that, to enhance the interatomic interaction, we have set the interatomic distance of Au$_2$ to 1.5 Å, which is much shorter than the equilibrium distance of 2.47 Å. Therefore, a general deduction is that the molecular $X$ can well be approximated as the superposition of the atomic ones\cite{73,74,76,98}

\[
X = \sum_{F} X_{F},
\]

(207)

which stays in the same spirit as P4C (see Point 3 below). Here, each atomic $X_{F}$ can, in view of the very definition (179), be obtained by solving the (radial) Dirac equation for a neutral or ionic spherical and unpolarized configuration. The atomic approximation to $X$ works very well not only for ground state energies of molecular systems\cite{73,76}, but also for electric\cite{99,104} and magnetic\cite{105,106} response properties, analytic energy gradients and Hessian\cite{107}, as well as periodic systems\cite{108}. In contrast, the widely used approximation $X_{1e}$ obtained by diagonalizing the one-electron Dirac matrix is not accurate enough for nuclear magnetic shielding, and cannot be applied to periodic systems. There have been
FIG. 6. Distribution of the matrix elements of $X = B^+A_+^1(A_+A_+^+)^{-1}$ for Au$_2$ at a distance of 1.5 Å. Dirac-Hartree-Fock (DHF) result with the uncontracted ANO-RCC basis set (594 functions for each atom).

attempts$^{[109–111]}$ to approximate the renormalization matrix $R_+$ also as the superposition of the atomic ones. Since $R_+$ is much less local than $X$, such approximation does introduce discernible errors$^{[112]}$. Nevertheless, such errors are tolerable for large systems in view of the dramatic gain in computational efficiency (especially in gradient and Hessian calculations$^{[107]}$). The atomic approximation to $X$ (and $R_+$) can obviously be generalized to a diatomic (fragmental) approximation$^{[73, 76]}$, which is of course only necessary if one is interested in highly distorted molecular systems in which two heavy atoms are very close to each other in distance. It is of particular interest to note that the atomic/fragmental approximation to both $X$ and $R_+$ (i.e., the X2C/AU Ansatz defined in Ref. $^{[107]}$) allows one to interpret$^{[48]}$ X2C as a seamless bridge between the Dirac and Schrödinger equations, because it can treat the heavy and light atoms in the system relativistically and nonrelativistically, respectively, unlike that the Dirac (Schrödinger) equation treats the whole system relativistically (nonrelativistically).

(3) $L_+^\text{NESCR}^{[192]}$ is closely related to $h_+^{P4C}^{[174]}$. To see this, we assume the A4S $\{|\varphi_{+,\mu}\rangle\}$ in the latter are further expanded in a RKB basis, viz.,

$$|\varphi_{+,\mu}\rangle = \sum_{\Lambda \in K} \left( \frac{g_{\Lambda} a_{K,\Lambda\mu}}{2} \sigma \cdot p g_{\Lambda} b_{K,\Lambda\mu} \right), \quad \forall \mu \in K; \quad b_K = X_K a_K$$

for each atom $K$. We then have

$$h_+^{P4C} = a^\dagger L_+^\text{NESCR} a, \quad a = \sum_K a_K, \quad (209)$$

$$S_+^{P4C} = a^\dagger S_+ a, \quad (210)$$
where \( L^{\text{NESC}}_+ \) in Eq. (209) and \( \tilde{S}^+ \) in Eq. (210) have adopted the atomic approximation (207) to \( X \). It is hence clear that P4C is just NESC, provided that the atomic-natural-spinor-type generally contracted RKB basis and the atomic approximation (207) to \( X \) are used in both cases. However, P4C\[75\] and NESC\[93\] were introduced in completely different ways, in the same year though. Unlike NESC, P4C is not limited to the RKB condition. Rather, it can also adopt, e.g., numerical A4S.

(4) All physical operators are subject to the same transformation going from the Dirac equation to a two-component theory. Neglecting this will result in PCE\[77\]. This can readily be done in the case of X2C, thanks to the simple relations \( A_+ = R_+ C_+ \) and \( B_+ = XA_+ \). Moreover, the MSCA (which takes care of both scalar and spin-orbit one-centered two-electron picture-change effects) underlying Q4C \[76\] can also be employed in X2C. A more dramatic simplification of X2C is to assemble the two-electron spin-orbit part of \( f^\text{X2C} \) from DHF calculations of spherically averaged atomic configurations and then neglect all molecular relativistic two-electron integrals\[113\]. All in all, the second-quantized, normal-ordered, PCE-corrected many-electron X2C Hamiltonian can be written as\[73\]

\[
H^\text{X2C}\_+ = E^\text{X2C}_\text{ref} + f^\text{X2C} \{ a^p_q \} + \frac{1}{2} g^{rs}_{pq} \{ a^p_{rs} \}.
\] (211)

Note in passing that if the model spectral form \[61, 62\] of the \( Q \) potential \( (106) \) is included in \( h \)\[178\], we would obtain automatically a QED@X2C approach\[35\].

At this stage it should have been clear that the Q4C and X2C formalisms render no-pair four- and two-component relativistic calculations completely identical in all aspects of simplicity, accuracy and efficiency, at both the mean-field and correlated levels (a point that was observed more than a decade ago\[73\]).

D. Spin-separated two-component

There are various situations where one would like to treat spin-free (sf) and spin-dependent (sd) relativistic effects separately. For instance, the terms “intersystem crossing” and “multistate reaction” are both rooted in the perturbative treatment of spin-orbit coupling (SOC). In addition to SOC, spin-dependent interactions include also spin-spin coupling (SSC). While SOC contains both one- and two-body terms, SSC is purely a two-body operator arising from the spin separation of the Gaunt interaction\[114\] and should be taken into account in calculations of magnetic properties\[115, 116\]. Here we outline briefly how to extract SOC from the
X2C Hamiltonian. The very first issue lies in that \( f_{+}^{X2C} \) is defined only in matrix form, such that the Dirac identity

\[
(\sigma \cdot A)B(\sigma \cdot C) = A \cdot (BC) + i\sigma \cdot [A \times (BC)]
\]  

(212)
cannot be used. However, we can start with the partitioning of the Dirac matrix (83) into a scalar and a spin-orbit term

\[
\begin{pmatrix}
V & T \\
T & \frac{\alpha^2}{4}W - T
\end{pmatrix} = \begin{pmatrix}
V & T \\
T & \frac{\alpha^2}{4}W_{sf} - T
\end{pmatrix} + \begin{pmatrix}
0 & 0 \\
0 & \frac{\alpha^2}{4}W_{sd}
\end{pmatrix},
\]

(213)

where

\[
(W_{sf})_{\mu\nu} = \langle g_\mu | p \cdot V p | g_\nu \rangle, \quad (W_{sd})_{\mu\nu} = \langle g_\mu | i\sigma \cdot (pV \times p) | g_\nu \rangle.
\]

(214)
The first, spin-free term can be block-diagonalized in the same way as before, so as to obtain

\[
h_{+sf}^{X2C} = R_{+,0}^+ (V + TX_0 + X_0^+T + X_0^+[\frac{\alpha^2}{4}W_{sf} - T]X_0)R_{+,0},
\]

(215)

where \( p, q, r, s \) refer to real-valued spin orbitals. Allowing the spin-free \( U_0 \) transformation [cf. Eq. (180)] to the second term of Eq. (213) leads to

\[
\begin{pmatrix}
\frac{\alpha^2}{4}R_{+,0}^+X_{0}^+W_{sd}X_{0}R_{+,0} & \frac{\alpha^2}{4}R_{+,0}^+X_{0}^+W_{sd}R_{-,0} \\
\frac{\alpha^2}{4}R_{-,0}^+W_{sd}X_{0}R_{+,0} & \frac{\alpha^2}{4}R_{-,0}^+W_{sd}R_{-,0}
\end{pmatrix},
\]

(216)

where the upper-left block is just the first-order SOC (to be denoted as so-DKH1)

\[
h_{SO,1e}^{(1)} = \frac{\alpha^2}{4}R_{+,0}^+X_{0}^+W_{SO}X_{0}R_{+,0}.
\]

(217)

Higher-order SOC can readily be obtained by carrying out further DKH-type unitary transformations that eliminate at each step the lowest-order odd terms in \( W_{sd} \). In particular, the so-DKH2 and so-DKH3 operators \( h_{SO,1e}^{(n)} \) can be constructed essentially for free (see Ref. [94] for the explicit expressions), because all necessary quantities are already available after constructing \( h_{+sf}^{X2C} \). As for the two-electron SOC, a mean-field approximation to the first-order terms is sufficient

\[
f_{SO,2e}^{(1)} = \frac{\alpha^2}{4}R_{+,0}^+[G_{LL}^{SO} + G_{LS}^{SL}X_0 + X_0^+G_{SL}^{SL} + X_0^+G_{SS}^{SS}X_0]R_{+,0},
\]

(218)
where $G_{SO}^{XY} (X, Y \in \{L, S\})$ are the matrices of the effective one-electron operators $G_{SO}^{XY}$,

\[
G_{SO}^{XY} = i\sigma \cdot g^{XY} = i\sum_l g^{XY,l}, \quad X, Y \in \{L, S\}, \quad l \in \{x, y, z\},
\]  

(219)

\[
\delta_{\mu\nu}^{LL,l} = -\sum_{\lambda\kappa} 2K_{\lambda\mu,\nu}^{l}P_{\lambda\kappa}^{SS},
\]  

(220)

\[
\delta_{\mu\nu}^{LS,l} = -\sum_{\lambda\kappa}(K_{\lambda\mu,\nu}^{l} + K_{\lambda\nu,\mu}^{l})P_{\lambda\kappa}^{LS},
\]  

(221)

\[
\delta_{\mu\nu}^{SS,l} = -\sum_{\lambda\kappa}(2K_{\lambda\mu,\nu}^{l} + K_{\lambda\nu,\mu}^{l} - K_{\lambda\mu,\nu}^{l})P_{\lambda\kappa}^{LL},
\]  

(222)

\[
K_{\lambda\mu,\nu}^{l} = \sum_{mn}\varepsilon_{lmn}(\mu_m\nu_{n}^{1}) = -K_{\lambda\mu,\nu}^{l}, \quad \mu_m = \partial_m\mu_l, \quad m, n \in \{x, y, z\},
\]  

(223)

\[
P_{\lambda\kappa}^{LL} = R_{+,0}P_{\lambda\kappa}^{0}P_{++,0}^{+}, \quad P_{\lambda\kappa}^{LS} = P_{\lambda\kappa}^{LL}X_{0}^{0}, \quad P_{\lambda\kappa}^{SS} = X_{0}P_{\lambda\kappa}^{LL}X_{0}^{0},
\]  

(224)

Here, $\kappa, \lambda, \mu, \nu$ refer to atomic (Gaussian) spin orbitals, $\varepsilon_{lmn}$ is the Levi-Civita symbol, while $P = \frac{1}{2}(P^{x} + P^{y})$ is the spin-averaged molecular density matrix, with $P^{x}$ and $P^{y}$ being the converged sf-X2C-ROHF/ROKS (restricted open-shell Hartree-Fock/Kohn-Sham) spin density matrices. The terms in Eq. (221) and the first two terms in Eq. (222) arise from the Coulomb interaction and represent the so-called spin-same-orbit coupling, whereas the term (220) and the third term of Eq. (222) originate from the Gaunt interaction and hence represent the spin-other-orbit coupling. In view of the short-range nature of SOC, a one-centre approximation to the integrals $K_{\lambda\mu,\nu}^{l}$ (223) can further be invoked. In this case, only the atomic blocks of the molecular density matrix $P_{\mu\nu}^{XY}$ contribute to $G_{SO}^{XY}$. Yet, $f_{SO,2e}^{(1)}$ is still a full matrix. If wanted, the SSC (114) can readily be added to $f_{SO,2e}^{(1)}$. The second-quantized, normal-ordered, spin-separated X2C Hamiltonian then reads

\[
H_{X2C}^{SOC} = E_{ref} + H_{sf} + H_{sd}^{[n]}, \quad n = 1 \text{ or } 3,
\]  

(225)

\[
H_{sf} = [h_{X2C}^{(1)}]_{pq}^{\{a_p\}} + \frac{1}{2}s_{pq}^{rs}\{a_r\},
\]  

(226)

\[
H_{sd}^{[n]} = [h_{SO,1e}^{[n]} + f_{SO,2e}^{(1)}]_{pq}^{\{a_p\}},
\]  

(227)

which is the simplest variant in the whole family of spin-separated X2C Hamiltonians (NB: $[n]$ denotes up to $n$-th order). It has been combined with both spin-adapted open-shell time-dependent density functional theory (119-121) and equation-of-motion coupled cluster for calculating fine structures of electronically excited states (102, 103).

Note in passing that, if the decoupling matrix $X_{0}^{1}$ (179) and the renormalization matrix $R_{+,0}^{1}$ (181) are set to identity in both $h_{SO,1e}^{(1)}$ (217) and $f_{SO,2e}^{(1)}$ (218) (i.e., so-DKH1), $H_{sd}^{(1)}$ (227) will reduce to the Breit-Pauli spin-orbit Hamiltonian (so-BP). So $H_{sd}^{(1)}$ can be understood as a bracketed (stabilized) so-BP. While so-BP can only used as a first-order perturbation operator on top of
the nonrelativistic problem, \( H_{sd}^{[n]} \) is bounded from below and can hence be treated variationally. On the other hand, if \( \mathbf{X}_0 \) and \( \mathbf{R}_{+,0} \) in \( h_{SO,1e}^{(1)} \) (217) and \( f_{SO,2e}^{(1)} \) (218) are both set to the free-particle counterparts, \( H_{sd}^{(1)} \) (227) will reduce to the original mean-field so-DKH1[122]. Moreover, \( H_{sd}^{[3]} \) is extremely accurate for both core and valence states[117] and can therefore be regarded as an equivalent of the non-perturbative SOX2CAMF operator[113].

The various Hamiltonians discussed so far, including \( H_a^{QED} \) (111), \( H_+^{QED} \) (167), \( H_+^{QED} \) (164), \( H_+^{DPCB} \) (168), \( H_+^{DCB} \) (169), \( H_+^{Q4C} \) (176), \( H_+^{X2C} \) (211), \( H_+^{X2CSOC} \) (225), and those A2C and nonrelativistic ones, share the same generic form

\[
H = E_{ref} + f_p^q \{ a_p^q \} + \frac{1}{2} g_{pq}^{rs} \{ a_r^q a_s^p \}. \tag{228}
\]

It is just that the Fockian operator \( f \) has to be interpreted differently. A complete and continuous “Hamiltonian Ladder” can then be depicted[34, 35]. The following points deserves to be emphasized again:

(a) Relativistic Hamiltonians can only be formulated in Fock space, whereas all first-quantized relativistic Hamiltonian suffer from contaminations of NES.

(b) \( H_a^{QED} \) (111) is the most accurate relativistic many-electron Hamiltonian and serves as the basis of the emerging field of “molecular QED”.

(c) Under the NPA, four- and two-component approaches are fully equivalent in all aspects of simplicity, accuracy and efficiency. Therefore, one should speak of “four-component and two-component equally good”, instead of “four-component good, two-component bad” or “two-component good, four-component bad”.

(d) X2C is computationally the same as but is much simpler and more accurate than A2C. As such, A2C should be regarded as outdated.

(e) sf-X2C+so-DKH1 is computationally the same as but is more accurate than NR+so-BP. As such, NR+so-BP should be regarded as outdated.

VI. NO-PAIR CORRELATION

Having discussed extensively the QED and relativistic many-electron Hamiltonians, we comment briefly on the correlation problem. Due to the large gap between the NES and PES, a second-order treatment of the NES is sufficient (see Sec. IV A). Therefore, the major challenge
still resides in the no-pair correlation within the manifold of PES. In this regard, like the non-relativistic case, one has to face two general issues, i.e., the slow basis-set convergence and the strong correlation problem. The former can only be improved by the so-called explicitly correlated methods. However, relativistic explicit correlation is plagued by two conceptual points: (a) no-pair projected or second-quantized relativistic Hamiltonians are simply incompatible with explicit correlation due to the lack of analytic operators. (b) The fact\[123\] that the two limits \( c \to \infty \) and \( r_{12} \to 0 \) do not commute makes how to apply the correlation factor \( f_{12} \) (which itself is a complicated quantity\[66\]) an open question. Rather unexpectedly, although the small-component \( \psi_p^S \) of a PES is indeed smaller (albeit in the mean) than the large-component \( \psi_p^L \) by a factor of \( c^{-1} \), the small-small component \( \Psi^{SS}(r_1, r_2) \) of a two-electron wave function \( \Psi(r_1, r_2) \) is of the same order of magnitude as the large-large component \( \Psi^{LL}(r_1, r_2) \) at the coalescence point\[66\]. This means simply that there is no obvious argument to favor the incorporation of the correlation factor \( f_{12} \) in a way that is in line with “first \( c \to \infty \) and then \( r_{12} \to 0 \)” or “first \( r_{12} \to 0 \) and then \( c \to \infty \)”. These issues have recently been scrutinized in depth\[66, 124–126\]. Since there are no new numerical results thereafter, we do not repeat the discussions here. However, it does deserve to be mentioned that the short-range density-functional type of corrections for basis-set incompleteness\[127–129\] is highly promising, not only because of its simplicity but also because it is rooted in second quantization and is hence compatible with relativistic Hamiltonians. The remaining issue is to develop suitable short-range relativistic density functionals for this purpose\[130\].

Compared to the slow basis-set convergence problem, the strong correlation problem is even more intricate in practice. A system is characterized as strongly correlated if a qualitative description already requires a multiconfigurational wave function. The main issue here lies in that the static and dynamic components of electron correlation are often strongly entangled and even interchangeable. The situation is further worsened by SOC. Although a number of relativistic schemes have been developed in the past\[112, 131–148\], approaches that can provide a balanced and self-adaptive description of the static and dynamic components of correlation still remain to be formulated. It is believed that the ultimate way is to introduce some selection procedure that can adapt to the variable static correlation automatically and meanwhile can be terminated at a stage at which the residual dynamic correlation can well be described by a low-order approach. This leads naturally to “selected configuration interaction plus second-order perturbation theory” (sCIPT2), a very old idea that can be traced back to the end of 1960s and has recently been revived in various ways (see Ref.\[149\] for a recent review). Such approaches
are most suited for relativistic calculations because of the following reasons:

1. A compact yet high-quality variational space can readily be determined by an iterative selection procedure, thereby avoiding problems inherent in the scenario of complete active space (CAS). For instance, the size of the CAS would be doubled in the presence of SOC, so as to limit severely the applicability of CAS-based four- or two-component relativistic correlation methods. This problem can only be resolved by selection.

2. The selection of important configurations is particularly effective for SOC, thanks to the short-range nature of SOC. This had better be combined with a local representation from the outset.

3. Unlike nonlinear wave function Ansätze, the symmetry adaptation of CI wave functions can readily be achieved by means of the spin-dependent unitary group approach.

4. SOC is strongly dominated by the one-body terms, such that a second-order perturbative treatment of dynamic correlation, on top of a well-controlled variational space, should be sufficient.

An X2C-based heat bath CI version of sCIPT2 has just been realized, showing great promises although SOC is included therein only at the correlated level but not at the orbital level. The combination of the QED@Q4C or QED@X2C Hamiltonian with the recently proposed iCIPT2 (iterative CI with selection plus second-order perturbation) should be even more promising, because iCIPT2 is spin-symmetry adapted and has the capability of targeting directly high-lying excited states that have little or even no overlap with the low-lying ones.

VII. SUMMARY

Ironically in history, just one year after he proposed the famous relativistic equation of motion, Dirac himself stated that ‘relativistic effects are of no importance in the consideration of atomic and molecular structure and ordinary chemical reactions’. Unfortunately, such a naive point of view lasted for nearly half a century, until the mid-1970s when relativistic effects were found to be indeed very important for electronic structure, sometimes even of light atoms. Since then the field of relativistic quantum chemistry has witnessed fast development, especially in the last 15 years, as evidenced by nearly 20,000 relativistic articles (see
Fig. 7) as well as more than 10 relativistic books\cite{52, 165-174}. With the advent of powerful computational software\cite{75, 175-188}, it can be envisaged that relativistic quantum chemistry will play an increasingly important role in the exploration of molecular science. Apart from further improvement in the computational efficiency, the most important and urgent methodological developments include (1) combination of the QED@Q4C/X2C and sf-X2C+so-DKH1 Hamiltonians with sophisticated, symmetry-adapted wave function-based no-pair correlation methods (e.g., iCIPT2) for high-precision calculations of electronic structure and (2) full implementation of the eQED Hamiltonian to establish the field of “molecular QED” for ultrahigh-precision calculations of spectroscopic parameters. Here, efficient implementation of the frequency-dependent Breit integrals, which scale formally as the 8th power of the basis-set size, must first be accomplished. Works along these directions are being carried out in our laboratory.

ACKNOWLEDGEMENT

This research was financially supported by National Natural Science Foundation of China (Grant Nos. 21833001 and 21973054).

DATA AVAILABILITY STATEMENT

The data that supports the findings of this study are available within the article.
FIG. 7. (a) Number of relativistic articles per year; (b) Distributions of relativistic articles in journals.
JACS: J. Am. Chem. Soc.; IJQC: Int. J. Quantum Chem.; PCCP: Phys. Chem. Chem. Phys.; TCA: Theor. Chem. Acc.; JCTC: J. Chem. Theory Comput.
[1] A. Einstein, Ann. Phys. 17, 891 (1905).
[2] O. Klein, Z. Phys. 37, 895 (1926).
[3] W. Gordon, Z. Phys. 40, 117 (1926).
[4] W. Greiner, Relativistic quantum mechanics, 3rd ed. (Springer-Verlag, Berlin, 1990).
[5] W. Pauli and V. Weisskopf, Helv. Phys. Acta 7, 709 (1934).
[6] S. Esposito, Ann. Phys. 16, 824 (2007).
[7] P. A. M. Dirac, Proc. Roy. Soc. A 117, 610 (1928).
[8] P. A. M. Dirac, Proc. Roy. Soc. A 118, 351 (1928).
[9] H. A. Kramers, Die Grundlagen der Quantentheorie: Quantentheorie des Electrons und der Strahlung (Akad. Verlagsges, Leipzig, 1933).
[10] J. Karworski, “Dirac operator and its properties,” in Handbook of Relativistic Quantum Chemistry, edited by W. Liu (Springer-Verlag, Berlin, 2017) pp. 3–49.
[11] J.-M. Lévy-Leblond, Commun. Math. Phys. 6, 286 (1967).
[12] W. Kutzelnigg, Z. Phys. D 11, 15 (1989).
[13] P. A. M. Dirac, Proc. Roy. Soc. A 126, 360 (1930).
[14] P. A. M. Dirac, Proc. Roy. Soc. A 133, 60 (1931).
[15] C. D. Anderson, Phys. Rev. 41, 405 (1932).
[16] W. Liu, Phys. Chem. Chem. Phys. 14, 35 (2012).
[17] W. Liu, “With-pair relativistic Hamiltonians,” in Handbook of Relativistic Quantum Chemistry, edited by W. Liu (Springer-Verlag, Berlin, 2017) pp. 345–373.
[18] The charge-conjugation transformation (51) of a field operator is bound to the particle-hole picture and hence should not be applied to expressions (28a) and (35a).
[19] W. Liu, Int. J. Quantum Chem. 115, 631 (2015), (E)116, 971 (2016).
[20] J. Schwinger, Phys. Rev. 82, 664 (1951).
[21] W. Liu and I. Lindgren, J. Chem. Phys. 139, 014108 (2013), (E)144, 049901 (2016).
[22] T. Aoyama, M. Hayakawa, T. Kinoshita, and M. Nio, Phys. Rev. Lett. 109, 111807 (2012).
[23] S. Sturm, F. Köhler, J. Zatorski, A. Wagner, Z. Harman, G. Werth, W. Quint, C. H. Keitel, and K. Blaum, Nature 506, 467 (2014).
[24] V. Shabaev, D. Glazov, N. Oreshkina, A. Volotka, G. Plunien, H.-J. Kluge, and W. Quint, Phys. Rev. Lett. 96, 253002 (2006).
[25] P. Indelicato and P. J. Mohr, “Introduction to bound-state quantum electrodynamics,” in Handbook of Relativistic Quantum Chemistry, edited by W. Liu (Springer-Verlag, Berlin, 2017) pp. 131–241.

[26] P. Pyykko, Chem. Rev. 112, 371 (2012).

[27] A. N. Artemyev, “QED effects and challenges,” in Handbook of Relativistic Quantum Chemistry, edited by W. Liu (Springer-Verlag, Berlin, 2017) pp. 243–265.

[28] K. Piszczatowski, G. Łach, M. Przybytek, J. Komasa, K. Pachucki, and B. Jeziorski, J. Chem. Theory Comput. 5, 3039 (2009).

[29] J. Liu, D. Sprecher, C. Jungen, W. Ubachs, and F. Merkt, J. Chem. Phys. 132, 154301 (2010).

[30] K. Pachucki and J. Komasa, Phys. Chem. Chem. Phys. 12, 9188 (2010).

[31] K. G. Dyall, C. W. Bauschlicher Jr, D. W. Schwenke, and P. Pyykkö, Chem. Phys. Lett. 348, 497 (2001).

[32] L. F. Pašteka, E. Eliav, A. Borschevsky, U. Kaldor, and P. Schwerdtfeger, Phys. Rev. Lett. 118, 023002 (2017).

[33] J. Niskanen, K. Jänkälä, M. Huttula, and A. Föhlisch, J. Chem. Phys. 146, 144312 (2017).

[34] W. Liu, Int. J. Quantum Chem. 114, 983 (2014).

[35] W. Liu, Phys. Rep. 537, 59 (2014).

[36] V. Shabaev, J. Phys. B 26, 4703 (1993).

[37] Y.-K. Kim, Phys. Rev. 154, 17 (1967).

[38] W. Schwarz and E. Wechsel-Trakowski, Chem. Phys. Lett. 85, 94 (1982).

[39] J. D. Talman, Phys. Rev. Lett. 57, 1091 (1986).

[40] J. Dolbeault, M. J. Esteban, and E. Séré, J. Funct. Anal. 174, 208 (2000).

[41] R. E. Stanton and S. Havriliak, J. Chem. Phys. 81, 1910 (1984).

[42] Y. Ishikawa, R. C. Binning Jr, and K. M. Sando, Chem. Phys. Lett. 101, 111 (1983).

[43] K. G. Dyall and K. Fægri Jr, Chem. Phys. Lett. 174, 25 (1990).

[44] V. Shabaev, I. Tupitsyn, V. Yerokhin, G. Plunien, and G. Soff, Phys. Rev. Lett. 93, 130405 (2004).

[45] Q. Sun, W. Liu, and W. Kutzelnigg, Theor. Chem. Acc. 129, 423 (2011).

[46] W. Kutzelnigg, Int. J. Quantum Chem. 25, 107 (1984).

[47] K. G. Dyall, J. Chem. Phys. 100, 2118 (1994).

[48] W. Liu, Mol. Phys. 108, 1679 (2010).

[49] K. Fægri Jr, Theor. Chem. Acc. 105, 252 (2001).

[50] W. Kutzelnigg, J. Chem. Phys. 126, 201103 (2007).

[51] Y. Xiao, Q. Sun, and W. Liu, Theor. Chem. Acc. 131, 1080 (2012).
[52] K. G. Dyall and K. Fægri Jr, *Introduction to relativistic quantum chemistry* (Oxford University Press, 2007).

[53] A. Almoukhalalati, S. Knecht, H. J. A. Jensen, K. G. Dyall, and T. Saue, J. Chem. Phys. 145, 074104 (2016).

[54] W. Kutzelnigg, Chem. Phys. 395, 16 (2012).

[55] H. Persson, I. Lindgren, S. Salomonson, and P. Sunnergren, Phys. Rev. A 48, 2772 (1993).

[56] E. A. Uehling, Phys. Rev. 48, 55 (1935).

[57] E. H. Wichmann and N. M. Kroll, Phys. Rev. 101, 843 (1956).

[58] G. Soff and P. J. Mohr, Phys. Rev. A 38, 5066 (1988).

[59] A. G. Fainshtein, N. Manakov, and A. A. Nekipelov, J. Phys. B 24, 559 (1991).

[60] I. Lindgren, *Relativistic Many-Body Theory: A New Field-Theoretical Approach* (Springer-Verlag, New York, 2011).

[61] V. Shabaev, I. Tupitsyn, and V. Yerokhin, Phys. Rev. A 88, 012513 (2013).

[62] V. Shabaev, I. Tupitsyn, and V. Yerokhin, Comput. Phys. Commun. 223, 69 (2018).

[63] E. Lindroth, A.-M. Mårtensson-Pendrill, A. Ynnerman, and P. Öster, J. Phys. B 22, 2447 (1989).

[64] G. Brown and D. Ravenhall, Proc. Roy. Soc. A 208, 552 (1951).

[65] H. Nakatsuji and H. Nakashima, Phys. Rev. Lett. 95, 050407 (2005).

[66] Z. Li, S. Shao, and W. Liu, J. Chem. Phys. 136, 144117 (2012).

[67] K. G. Dyall, Chemi. Phys. 395, 35 (2012).

[68] T. Saue, J. Chem. Phys. 12, 3077 (2011).

[69] D. Peng and M. Reiher, Theor. Chem. Acc. 131, 1081 (2012).

[70] W. Liu, Natl. Sci. Rev. 3, 204 (2016).

[71] W. Liu, “No-pair relativistic Hamiltonians: Q4C and X2C,” in *Handbook of Relativistic Quantum Chemistry*, edited by W. Liu (Springer-Verlag, Berlin, 2017) pp. 375–393.

[72] J. Sapirstein, K. Cheng, and M. Chen, Phys. Rev. A 59, 259 (1999).

[73] D. Peng, W. Liu, Y. Xiao, and L. Cheng, J. Chem. Phys. 127, 104106 (2007).

[74] K. G. Dyall and T. Enevoldsen, J. Chem. Phys. 111, 10000 (1999).

[75] W. Liu, G. Hong, D. Dai, L. Li, and M. Dolg, Theor. Chem. Acc. 96, 75 (1997).

[76] W. Liu and D. Peng, J. Chem. Phys. 125, 044102 (2006), (E)125, 149901 (2006).

[77] V. Kellö and A. J. Sadlej, Int. J. Quantum Chem. 68, 159 (1998).

[78] C. Chang, M. Pelissier, and P. Durand, Phys. Scr. 34, 394 (1986).

[79] E. van Lenthe, E.-J. Baerends, and J. G. Snijders, J. Chem. Phys. 99, 4597 (1993).
[80] W. Kutzelnigg and W. Liu, J. Chem. Phys. 123, 241102 (2005).
[81] W. Liu and D. Peng, J. Chem. Phys. 131, 031104 (2009).
[82] W. Liu and W. Kutzelnigg, J. Chem. Phys. 126, 114107 (2007).
[83] M. Iliaš and T. Saue, J. Chem. Phys. 126, 064102 (2007).
[84] M. Barysz and A. J. Sadlej, J. Chem. Phys. 116, 2696 (2002).
[85] D. Ke, M. Barysz, *et al.*, Chem. Phys. Lett. 446, 176 (2007).
[86] H. J. A. Jensen, Proceedings of the International Conference on Relativistic Effects in Heavy Element Chemistry and Physics, Mülheim/Ruhr, 6-10 April, 2005.
[87] M. Reiher and A. Wolf, J. Chem. Phys. 121, 2037 (2004).
[88] M. Reiher and A. Wolf, J. Chem. Phys. 121, 10945 (2004).
[89] D. Peng and K. Hirao, J. Chem. Phys. 130, 044102 (2009).
[90] B. A. Hess, Phys. Rev. A 33, 3742 (1986).
[91] G. Jansen and B. A. Heß, Phys. Rev. A 39, 6016 (1989).
[92] The acronym ‘X2C’ (pronounced as ‘ecstacy’) for exact two-component Hamiltonians was proposed by W. Liu after intensive discussions with H. J. Aa. Jensen, W. Kutzelnigg, T. Saue and L. Visscher during the Twelfth International Conference on the Applications of Density Functional Theory (DFT-2007), Amsterdam, August 26-30, 2007. Note that the ‘exact’ here emphasizes that all the solutions of the matrix Dirac equation can be reproduced up to machine accuracy. It is particularly meaningful when compared with the approximate two-component Hamiltonians.
[93] K. G. Dyall, J. Chem. Phys. 106, 9618 (1997).
[94] Z. Li, Y. Xiao, and W. Liu, J. Chem. Phys. 141, 054111 (2014).
[95] L. L. Foldy and S. A. Wouthuysen, Phys. Rev. 78, 29 (1950).
[96] K. G. Dyall and E. van Lenthe, J. Chem. Phys. 111, 1366 (1999).
[97] W. Kutzelnigg and W. Liu, Mol. Phys. 104, 2225 (2006).
[98] K. G. Dyall, J. Chem. Phys. 115, 9136 (2001).
[99] W. Xu, J. Ma, D. Peng, W. Zou, W. Liu, and V. Staemmler, Chem. Phys. 356, 219 (2009).
[100] W. Xu, Y. Zhang, and W. Liu, Sci. China Ser. B: Chem. 52, 1945 (2009).
[101] Y. Zhang, W. Xu, Q. Sun, W. Zou, and W. Liu, J. Comput. Chem. 31, 532 (2010).
[102] Z. Li, B. Suo, Y. Zhang, Y. Xiao, and W. Liu, Mol. Phys. 111, 3741 (2013).
[103] Z. Cao, Z. Li, F. Wang, and W. Liu, Phys. Chem. Chem. Phys. 19, 3713 (2017).
[104] W. Liu and Y. Xiao, Chem. Soc. Rev. 47, 4481 (2018).
[105] Q. Sun, W. Liu, Y. Xiao, and L. Cheng, J. Chem. Phys. 131, 081101 (2009).
[106] Q. Sun, Y. Xiao, and W. Liu, J. Chem. Phys. 137, 174105 (2012).
[107] W. Zou, G. Guo, B. Suo, and W. Liu, J. Chem. Theory Comput. 16, 1541 (2020).
[108] R. Zhao, Y. Zhang, Y. Xiao, and W. Liu, J. Chem. Chem. 144, 044105 (2016).
[109] D. Peng and M. Reiher, J. Chem. Phys. 136, 244108 (2012).
[110] J. Seino and H. Nakai, J. Chem. Phys. 136, 244102 (2012).
[111] J. Seino and H. Nakai, J. Chem. Phys. 137, 144101 (2012).
[112] P. K. Tamukong, M. R. Hoffmann, Z. Li, and W. Liu, J. Phys. Chem. A 118, 1489 (2014).
[113] J. Liu and L. Cheng, J. Chem. Phys. 148, 144108 (2018).
[114] W. Kutzelnigg and W. Liu, J. Chem. Phys. 112, 3540 (2000).
[115] O. Vahtras, O. Loboda, B. Minaev, H. Ågren, and K. Ruud, Chem. phys. 279, 133 (2002).
[116] L. Lang and F. Neese, J. Chem. Phys. 150, 104104 (2019).
[117] Z. Li, Y. Xiao, and W. Liu, J. Chem. Phys. 137, 154114 (2012).
[118] Z. Li and W. Liu, “Spin separation of relativistic Hamiltonians,” in Handbook of Relativistic Quantum Chemistry, edited by W. Liu (Springer-Verlag, Berlin, 2017) pp. 411–447.
[119] Z. Li and W. Liu, The Journal of chemical physics 133, 064106 (2010).
[120] Z. Li, W. Liu, Y. Zhang, and B. Suo, The Journal of chemical physics 134, 134101 (2011).
[121] Z. Li and W. Liu, J. Chem. Phys. 135, 194106 (2011), (E)138, 029904 (2013).
[122] B. A. Hess, C. M. Marian, U. Wahlgren, and O. Gropen, Chem. Phys. Lett. 251, 365 (1996).
[123] W. Kutzelnigg, in Aspects of Many-Body Effects in Molecules and Extended Systems (Springer, 1989) pp. 353–366.
[124] S. Shao, Z. Li, and W. Liu, “Basic structures of relativistic wave functions,” in Handbook of Relativistic Quantum Chemistry, edited by W. Liu (Springer-Verlag, Berlin, 2017) pp. 481–496.
[125] S. Shao, Z. Li, and W. Liu, “Coalescence conditions of relativistic wave functions,” in Handbook of Relativistic Quantum Chemistry, edited by W. Liu (Springer-Verlag, Berlin, 2017) pp. 497–530.
[126] W. Liu, S. Shao, and Z. Li, “Relativistic explicit correlation: Problems and solutions,” in Handbook of Relativistic Quantum Chemistry, edited by W. Liu (Springer-Verlag, Berlin, 2017) pp. 531–545.
[127] E. Giner, B. Pradines, A. Ferté, R. Assaraf, A. Savin, and J. Toulouse, J. Chem. Phys. 149, 194301 (2018).
[128] E. Giner, A. Scemama, J. Toulouse, and P.-F. Loos, J. Chem. Phys. 151, 144118 (2019).
[129] P.-F. Loos, B. Pradines, A. Scemama, J. Toulouse, and E. Giner, J. Phys. Chem. Lett. 10, 2931 (2019).
[130] J. Paquier, E. Giner, and J. Toulouse, arXiv preprint arXiv:2002.06935 (2020).
[131] P. Å. Malmqvist, B. O. Roos, and B. Schimmelpfennig, Chem. Phys. Lett. 357, 230 (2002).
[132] P. Å. Malmqvist, Int. J. Quantum Chem. 30, 479 (1986).
[133] P.-Å. Malmqvist and B. O. Roos, Chem. Phys. Lett. 155, 189 (1989).
[134] B. O. Roos and P.-Å. Malmqvist, Phys. Chem. Chem. Phys. 6, 2919 (2004).
[135] H. Jørgen Aa. Jensen, K. G. Dyall, T. Saue, and K. Fægri Jr, J. Chem. Phys. 104, 4083 (1996).
[136] T. Fleig, C. M. Marian, and J. Olsen, Theor. Chem. Acc. 97, 125 (1997).
[137] Y. S. Kim and Y. S. Lee, J. Chem. Phys. 119, 12169 (2003).
[138] I. Kim and Y. S. Lee, J. Chem. Phys. 139, 134115 (2013).
[139] J. Thyssen, T. Fleig, and H. J. A. Jensen, J. Chem. Phys. 129, 034109 (2008).
[140] M. Abe, G. Gopakmar, T. Nakajima, and K. Hirao, in *Radiation Induced Molecular Phenomena in Nucleic Acids* (Springer, 2008) pp. 157–177.
[141] J. E. Bates and T. Shiozaki, J. Chem. Phys. 142, 044112 (2015).
[142] R. D. Reynolds, T. Yanai, and T. Shiozaki, J. Chem. Phys. 149, 014106 (2018).
[143] T. Shiozaki and W. Mizukami, J. Chem. Theory Comput. 11, 4733 (2015).
[144] G. Moritz, A. Wolf, and M. Reiher, J. Chem. Phys. 123, 184105 (2005).
[145] S. Knecht, Ö. Legeza, and M. Reiher, J. Chem. Phys. 140, 041101 (2014).
[146] S. Battaglia, S. Keller, and S. Knecht, J. Chem. Theory Comput. 14, 2353 (2018).
[147] B. Zhang, J. E. Vandezande, R. D. Reynolds, and H. F. Schaefer III, J. Chem. Theory Comput. 14, 1235 (2018).
[148] A. J. Jenkins, H. Liu, J. M. Kasper, M. J. Frisch, and X. Li, J. Chem. Theory Comput. 15, 2974 (2019).
[149] N. Zhang, W. Liu, and M. R. Hoffmann, J. Chem. Theory Comput. (2020), 10.1021/acs.jctc.9b01200
[150] F. Wu, W. Liu, Y. Zhang, and Z. Li, J. Chem. Theory Comput. 7, 3643 (2011).
[151] J. Liu, Y. Zhang, and W. Liu, J. Chem. Theory Comput. 10, 2436 (2014).
[152] Z. Li, H. Li, B. Suo, and W. Liu, Acc. Chem. Res. 47, 2758 (2014).
[153] H. Li, W. Liu, and B. Suo, J. Chem. Phys. 146, 104104 (2017).
[154] M. Gould and G. Chandler, Int. J. Quantum Chem. 25, 1089 (1984).
[155] S. Yabushita, Z. Zhang, and R. M. Pitzer, J. Phys. Chem. A 103, 5791 (1999).
[156] B. Mussard and S. Sharma, J. Chem. Theory Comput. 14, 154 (2017).
[157] W. Liu and M. R. Hoffmann, J. Chem. Theory Comput. 12, 1169 (2016), (E)12, 3000 (2016).
[158] W. Liu and M. R. Hoffmann, Theor. Chem. Acc. 133, 1481 (2014).
[159] Y. Lei, W. Liu, and M. R. Hoffmann, Mol. Phys. 115, 2696 (2017).
[160] C. Huang, W. Liu, Y. Xiao, and M. R. Hoffmann, J. Comput. Chem. 38, 2481 (2017), (E)39, 338 (2018).
[161] C. Huang and W. Liu, J. Comput. Chem. 40, 1023 (2019).
[162] P. A. M. Dirac, Proc. Roy. Soc. A 123, 714 (1929).
[163] P. Pyykö, in Advances in quantum chemistry, Vol. 11 (Elsevier, 1978) pp. 353–409.
[164] P. Pyykö, J. Comp. Chem. 34, 2667 (2013), see http://www.rtam.chem.helsinki.fi updated to Version 22.1 (2020).
[165] P. Schwerdtfeger, Relativistic Electronic Structure Theory: Part 1. Fundamentals (Elsevier, 2002).
[166] P. Schwerdtfeger, Relativistic Electronic Structure Theory: Part 2. Applications (Elsevier, 2004).
[167] B. A. Hess, Relativistic Effects in Heavy-Element Chemistry and Physics (J. Wiley, 2003).
[168] K. Hirao and Y. Ishikawa, Recent Advances in Relativistic Molecular Theory, Vol. 5 (World Scientific, 2004).
[169] I. P. Grant, Relativistic Quantum Theory of Atoms and Molecules, Theory and Computation (Springer, New York, 2007).
[170] M. Barysz and Y. Ishikawa, Relativistic Methods for Chemists, Vol. 10 (Springer Science & Business Media, 2010).
[171] U. Kaldor and S. Wilson, Theoretical Chemistry and Physics of Heavy and Superheavy Elements, Vol. 11 (Springer Science & Business Media, 2013).
[172] M. Reiher and A. Wolf, Relativistic Quantum Chemistry: The Fundamental Theory of Molecular Science, 2nd ed. (Wiley-VCH, Weinheim, 2015).
[173] M. Dolg, Computational Methods in Lanthanide and Actinide Chemistry (John Wiley & Sons, 2015).
[174] W. Liu, Handbook of Relativistic Quantum Chemistry (Springer, Berlin, 2017).
[175] G. T. Te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. van Gisbergen, J. G. Snijders, and T. Ziegler, J. Comput. Chem. 22, 931 (2001).
[176] W. Liu, F. Wang, and L. Li, J. Theor. Comput. Chem. 2, 257 (2003).
[177] W. Liu, F. Wang, and L. Li, “Relativistic density functional theory: The BDF program package,” in Recent Advances in Relativistic Molecular Theory, edited by K. Hirao and Y. Ishikawa (World Scientific, Singapore, 2004) pp. 257–282.
[178] Y. Zhang, B. Suo, Z. Wang, N. Zhang, Z. Li, Y. Lei, W. Zou, J. Gao, D. Peng, Z. Pu, Y. Xiao, Q. Sun, F. Wang, Y. Ma, X. Wang, Y. Guo, and W. Liu, J. Chem. Phys. 152, 064113 (2020).
[179] I. P. Grant, in Supercomputing, Collision Processes, and Applications, Physics of Atoms and Molecules, edited by K. L. Bell, K. A. Berrington, D. S. F. Crothers, A. Hibbert, and K. T. Taylor (Springer,
[180] T. Yanai, H. Nakano, T. Nakajima, T. Tsuneda, S. Hirata, Y. Kawashima, Y. Nakao, M. Kamiya, H. Sekino, and K. Hirao, in *Computational Science – ICCS 2003*, Lecture Notes in Computer Science, Vol. 2660, edited by P. M. A. Sloot, D. Abramson, A. V. Bogdanov, Y. E. Gorbachev, J. J. Dongarra, and A. Y. Zomaya (Springer, Berlin, Heidelberg, 2003) pp. 84–95.

[181] T. Nakajima, M. Katouda, M. Kamiya, and Y. Nakatsuka, *Int. J. Quantum Chem.* **115**, 349 (2015).

[182] COLOGNE, a quantum chemical electronic structure program, Release COLOGNE15, written by E. Kraka, W. Zou, M. Filatov, J. Gräfenstein, D. Izotov, J. Gauss, Y. He, A. Wu, V. Polo, L. Olsson, Z. Konkoli, Z. He, and D. Cremer. Southern Methodist University, Dallas, TX, 2015.

[183] TURBOMOLE V7.4.1 2019, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from [http://www.turbomole.com](http://www.turbomole.com).

[184] DIRAC, a relativistic ab initio electronic structure program written by A. S. P. Gomes, T. Saue, L. Visscher, H. J. Aa. Jensen, and R. Bast, with contributions from I. A. Aucar, V. Bakken, K. G. Dyall, S. Dubillard, U. Ekström, E. Eliav, T. Enevoldsen, E. Faßhauer, T. Fleig, O. Fossgaard, L. Halbert, E. D. Hedegård, B. Heimlich–Paris, T. Helgaker, J. Henriksson, M. Iliaš, Ch. R. Jacob, S. Knecht, S. Komorovský, O. Kullie, J. K. Lærdahl, C. V. Larsen, Y. S. Lee, H. S. Nataraj, M. K. Nayak, P. Norman, G. Olejniczak, J. Olsen, J. M. H. Olsen, Y. C. Park, J. K. Pedersen, M. Pernpointner, R. di Remigio, K. Ruud, P. Salek, B. Schimmelpfennig, B. Senjean, A. Shee, J. Sikkema, A. J. Thorvaldsen, J. Thyssen, J. van Stralen, M. L. Vidal, S. Villaume, O. Visser, T. Winther, and S. Yamamoto (see [http://www.diracprogram.org](http://www.diracprogram.org)).

[185] ReSpect, a relativistic spectroscopy DFT program written by M. Repisky, S. Komorovsky, V. G. Malkin, O. L. Malkina, M. Kaupp, and K. Ruud, with contributions from R. Bast, R. Di Remigio, U. Ekstrom, M. Kadek, S. Knecht, L. Konecny, E. Malkin, and I. Malkin Ondik (see [http://www.respectprogram.org](http://www.respectprogram.org)).

[186] M. Hayami, J. Seino, Y. Nakajima, M. Nakano, Y. Ikabata, T. Yoshikawa, T. Oyama, K. Hiraga, S. Hirata, and H. Nakai, *J. Comput. Chem.* **39**, 2333 (2018).

[187] T. Shiozaki, WIREs Comput. Mol. Sci. **8**, e1331 (2018).

[188] D. B. Williams-Young, A. Petrone, S. Sun, T. F. Stetina, P. Lestrange, C. E. Hoyer, D. R. Nascimento, L. Koulias, A. Wildman, J. Kasper, J. J. Goings, F. Ding, A. E. DePrince III, E. F. Valeev, and X. Li, *WIREs Comput. Mol. Sci.* **10**, e1436 (2020).