Further analysis of the complete Breit interaction

Eva Lindroth*
Clarendon Laboratory, OX1 3PU Oxford, England

Ann-Marie Mårtensson-Pendrill
Department of Physics, University of Göteborg and Chalmers University of Technology, S-412 96 Göteborg, Sweden

(Received 28 September 1988)

In a recent paper [Phys. Rev. A 37, 1087 (1988)] Gorceix and Indelicato study the difference between the multiconfiguration Dirac-Fock (MCDF) expectation value of the Lorentz and Coulomb gauge forms for the Breit interaction. After a slight rearrangement of their expressions, we have found that the difference can be written as a commutator with the full Hamiltonian. This commutator is nonvanishing due to projection operators onto positive energy states implicit in the MCDF approach and gives contributions $O(Z^2\alpha^2)$. It is shown that the diagram involving the excitation of a single virtual electron-positron pair, which removes the discrepancy if the low-frequency limit is used also where it is not valid, does not do so for the proper frequency-dependent form and that a more detailed analysis is required.

I. INTRODUCTION

In a recent paper Gorceix and Indelicato[1] compare numerical results obtained for the retardation effects using the Feynman (Lorentz) and Coulomb gauge expressions between wave functions obtained in a multiconfiguration Dirac-Fock (MCDF) procedure. They found deviations proportional to $\alpha^5 Z^2 m_e c^2$ between the results obtained in the two gauges. In previous extensive studies of two-electron ions[2,3] they found that the effect of retardation is, indeed, essential, to obtain agreement with experiment. These results for the discrepancy between the gauges were based on only two configurations, and it is possible that by using a more complete set the discrepancy may be reduced, as suggested by Grant.[4] However, as we shall see below, this will not be sufficient—a gauge noninvariance to this order remains within the no-pair approximation used implicitly in the MCDF procedure.

In Coulomb gauge the correction to be added to the expression, e^2/r_{12}, for single-photon exchange, can be written in the form[5,6]

$$B_\omega(\mathbf{r}_1,\mathbf{r}_2)=e^2\left[-\boldsymbol{\alpha}_1\cdot\boldsymbol{\alpha}_2\frac{\cos\omega r_{12}}{r_{12}}+\left[(\boldsymbol{\alpha}_1\cdot\boldsymbol{\nabla}_1),\left[(\boldsymbol{\alpha}_2\cdot\boldsymbol{\nabla}_2),\frac{\cos\omega r_{12}-1}{\omega^2 r_{12}}\right]\right]\right]. \tag{1a}$$

The value of ω is the energy of the exchanged photon divided by $\hbar c$. In the Feynman gauge, the correction is instead

$$B_\omega(\mathbf{r}_1,\mathbf{r}_2)=e^2\left[\frac{\cos\omega r_{12}-1}{r_{12}}-\frac{\boldsymbol{\alpha}_1\cdot\boldsymbol{\alpha}_2}{r_{12}}\cos\omega r_{12}\right]. \tag{1b}$$

In the limit $\omega\to 0$, the Feynman gauge reduces to the Gaunt interaction, $-e^2\boldsymbol{\alpha}_1\cdot\boldsymbol{\alpha}_2/r_{12}$ and the Coulomb gauge expression reduces to the original Breit interaction, which is correct to $O(\alpha^2)$ in the low-frequency limit, whereas the leading correction due to the frequency dependence for the Feynman gauge is $-\omega^2 e^2 r_{12}/2$, which is $O(\alpha^2)$. The expressions (1a) and (1b) were derived under assumption of energy conservation between the initial and final state in the matrix elements. Following Brown[7] and Mittlemann[8] a general two-electron matrix element can be written as

$$\langle cd|B_{12}|ab\rangle=\langle cd|[B_{\omega_{ca}}(\mathbf{r}_1,\mathbf{r}_2)+B_{\omega_{db}}(\mathbf{r}_1,\mathbf{r}_2)]/2|ab\rangle\ , \tag{2}$$

where $\omega_{ca}=|\varepsilon_c-\varepsilon_a|/\hbar c$. Evaluating the expectation value of this interaction implies a neglect of any other interaction which may occur while the virtual photon moves between the electrons.

The one-electron energies ε_i are obtained as eigenvalues to the Dirac one-electron Hamiltonian,

$$\begin{aligned} &h|a\rangle=\varepsilon_a|a\rangle\ , \\ &h=\left[\beta mc^2+c\boldsymbol{\alpha}\cdot\mathbf{p}-\frac{Ze^2}{r}+v\right]=h_{\rm nuc}+v\ . \end{aligned} \tag{3}$$

(We normally work in Hartree atomic units, $m=e=4\pi\varepsilon_0=\hbar=1$, $c=1/\alpha$, but here we have chosen to keep also m and e explicit.) The potential v gives an approximate description of the electron-electron interac-

tion. For a two-electron system the Hamiltonian can be written as

$$H = h_{\mathrm{nuc},1} + h_{\mathrm{nuc},2} + V_{12} = h_1 + h_2 + (V_{12} - v_1 - v_2) , \quad (4)$$

where V_{12} describes the interaction between the two electrons. Due to the existence of negative-energy solutions to the one-electron equation (3), the two-electron Hamiltonian in (4) must be used with care to ensure that excitations *into* negative-energy states are not allowed in the wave function—only excitations *from* them, corresponding to the creation of virtual electron-positron pairs, are permitted. The situation is similar to that in a large atom where the valence electrons are not allowed to be excited into the occupied core orbitals. This is accounted for (see, e.g., Ref. 9) by a second-quantized representation of interactions and using Wick's theorem to get a normal ordering of creation and annihilation operators with respect to a suitably chosen "vacuum" level. In the relativistic many-electron case, failure to respect normal ordering for the negative-energy states leads to the well-known problems of "continuum dissolution" first noted by Brown and Ravenhall,[10] and brought back to attention by Sucher.[11] This problem occurs whether V_{12} is the pure Coulomb interaction or includes also the effect of transverse photons, be it in the Coulomb gauge (1a) or Feynman gauge (1b) form, with or without the proper frequency dependence. However, below we consider mainly the case $V_{12} = e^2/r_{12}$, which was used in the MCDF calculations in Ref. 1, with the effect of transverse photons evaluated as a first-order perturbation.

II. THE DIFFERENCE BETWEEN THE GAUGES

Gorceix and Indelicato[1] have obtained convenient expressions for the contributions to the difference between the Feynman and Coulomb gauge expressions for the case of a state-independent local potential in the one-electron Hamiltonian. Combining them we get

$$\begin{aligned}\delta E^{F-C} &= \frac{e^2}{2c^2}\{\langle\Psi|[h_1,[h_1,F_{\omega_1}]]+[h_2,[h_2,F_{\omega_2}]]|\Psi\rangle \\ &\quad + \langle\Psi|[h_1,[h_2,(F_{\omega_1}+F_{\omega_2})]]|\Psi\rangle\} \\ &= \frac{e^2}{2c^2}\langle\Psi|[h_1+h_2,([h_1,F_{\omega_1}]+[h_2,F_{\omega_2}])]|\Psi\rangle ,\end{aligned} \quad (5)$$

where

$$F_\omega(r_{12}) = \frac{\cos\omega r_{12} - 1}{\omega^2 r_{12}} \quad (6)$$

and ω_1 and ω_2 are defined as in Eq. (2) from the energies of one-particle states entering the matrix element for coordinates 1 and 2, respectively. This expression vanishes if Ψ is a single-determinantal wave function with orbitals obtained in a state-independent local potential, but the expression holds also if Ψ is a superposition of such determinants. However, as we shall see below, δE^{F-C} does then, in general, *not* vanish.

In the first expression in (5), the effect of the double commutator with h for either of the coordinates is to remove the ω^2 in the denominator of F, leaving $(\cos\omega r_{12} - 1)/r_{12}$, which is the first term in the Feynman gauge expression (1b). The second term in the Feynman gauge expression, $-\boldsymbol{\alpha}_1\cdot\boldsymbol{\alpha}_2\cos\omega r_{12}$, is common to both gauges and does not contribute to the difference. The second term in (5) involving a commutator with h_1 of a commutator with h_2 gives the last term in the Coulomb gauge expression (1a). Here, only the $c\boldsymbol{\alpha}\cdot\mathbf{p}$ term of h is desired and we may have to correct for the commutator with the one-electron potential, v in (3).[12] We consider first the hydrogenic single-electron Hamiltonian and postpone to Sec. II C a brief discussion of the modification necessary in the general case.

A. Projection operators and gauge invariance

The frequency dependence complicates the commuting properties of F; although a local potential commutes with the expression given in (6) for any fixed ω, it does not in general commute with F, since ω in F is defined by the surrounding one-electron states and the potential mixes in states with different energies. However, any state-independent local potential commutes with the low-frequency limit $F_0 = -r_{12}/2$, which reproduces the $O(Z^3\alpha^2)$ contributions to the matrix elements in both gauges, and we will first investigate the gauge invariance in this case, which was studied numerically by Gorceix and Indelicato.[1] The expression for the difference can then be written as

$$\delta E^{F-C} = \frac{e^2}{2c^2}\langle\Psi|[H,[h_1+h_2,F_0]]|\Psi\rangle$$

if the electron-electron interaction V_{12} commutes with the internal commutators. This holds for the pure Coulomb interaction as well as for the frequency-independent Gaunt and Breit interactions. It is tempting to conclude that the expectation value of a commutator with the full Hamiltonian would vanish, as it would, indeed for an exact nonrelativistic wave function. However, in the relativistic case, the additional complication of projection operators appear: The eigenvalue equation solved, e.g., in MCDF is assumed to be an approximate "no-pair" eigenfunction satisfying

$$\Psi = \Lambda^{++}\Psi, \quad \Lambda^{++}(H-E)\Lambda^{++}\Psi = 0, \quad \Lambda^{++} = \lambda_1^+\lambda_2^+ ,$$

where λ^+ (λ^-) projects onto the positive (negative) -energy subspace for the respective coordinates. Using the relations $\lambda^+ + \lambda^- = 1$ for both coordinates, $H\Lambda^{++}\Psi$ can be rewritten

$$\begin{aligned}H\Lambda^{++}\Psi &= [\Lambda^{++}H + (1-\Lambda^{++})H]\Lambda^{++}\Psi \\ &= E\Lambda^{++}\Psi + (\lambda_1^-\lambda_2^+ + \lambda_1^+\lambda_2^- + \lambda_1^-\lambda_2^-)H\Lambda^{++}\Psi ,\end{aligned}$$

and only the contributions from the first term, $E\Lambda^{++}\Psi$, vanish in the expression for the difference between the gauges. Since the projection operators commute with the

single-electron Hamiltonians, only V_{12} (and v_i when applicable, as discussed in Sec. II C) contributes, and the expression for the difference then becomes

$$\delta E^{F-C}=\frac{e^2}{2c^2}\langle\Psi|\{V_{12}(\lambda_1^-\lambda_2^+ +\lambda_1^+\lambda_2^- +\lambda_1^-\lambda_2^-)[h_1+h_2,F_0]\}|\Psi\rangle +\text{H.c.} , \tag{7}$$

where the outer commutator with the anti-Hermitian inner commutator has been replaced by addition of the Hermitian conjugate (denoted by H.c.). As seen in Sec. II B, the mixed terms involving one electron and one positron state give $O(\alpha^2)$ contributions, whereas the terms involving two positron states are singular leading to $O(\alpha^3)$ contributions rather than $O(\alpha^4)$, as might be expected from a direct power counting.

An alternative, manifestly Hermitian, form of the difference is obtained by using $H\Lambda^{++}=\Lambda^{++}H\Lambda^{++} +[V_{12},\Lambda^{++}]\Lambda^{++}$ together with standard formulas for commutators with a product:

$$\delta E^{F-C}=\frac{-e^2}{2c^2}\langle\Psi|\{[V_{12},\Lambda^{++}],[h_1+h_2,F_0]\}|\Psi\rangle =\frac{-e^2}{2c^2}\langle\Psi|\{([V_{12},\lambda_1^+]+[V_{12},\lambda_2^+]),[h_1+h_2,F_0]\}|\Psi\rangle-\frac{-e^2}{2c^2}\langle\Psi|[([[V_{12},\lambda_2^+],\lambda_1^+]),[h_1+h_2,F_0]]|\Psi\rangle , \tag{8}$$

where the first part on the right-hand side (RHS) is an anticommutator (denoted by curly brackets) between anti-Hermitian single commutators and the second is a commutator between a Hermitian double commutator and an anti-Hermitian single commutator. The identity $\lambda^-O\lambda^+=[\lambda^-,O]\lambda^+ =-[\lambda^+,O]\lambda^+$ shows that the last term in (8) involves a projection in both coordinates onto negative-energy intermediate states. It has the opposite sign of the last term in (7); however, this is compensated by the two single commutators, which imply projection onto the negative-energy states of only one of the coordinates 1 and 2 and thus include both positive- and negative-energy states in the other coordinate. Using commutators, no errors can result from approximations in $\lambda^-\Psi$, which should be identically zero.

To analyze the contributions to the discrepancy, we need expressions for the projection operators. In earlier work[13,14] we obtained expressions for the projection operators for positive- and negative-energy states based on an operator R, which relates the upper and lower components of the wave function. To leading order in $(Z\alpha)^2$, they are identical to the Casimir free-particle projection operators (given, e.g., in Refs. 11 and 15), and after a slight rearrangement we get

$$\lambda_\pm\approx\frac{1}{2}\pm\frac{\beta/2+\boldsymbol{\alpha}\cdot\mathbf{p}/2mc}{(1+\mathbf{p}^2/m^2c^2)^{1/2}} \approx\frac{1\pm\beta}{2}\mp\frac{\beta p^2}{4m^2c^2}\pm\frac{\boldsymbol{\alpha}\cdot\mathbf{p}}{2mc} , \tag{9}$$

where the approximate last expressions contain only the lowest-order terms. Including the nuclear potential in the projection operators leads to a reduction in the size of all operators close to the nucleus [where $p^2\gg(mc)^2$], as discussed in Ref. 13.

B. Analysis of the relative importance of different terms

The expression (8) for the difference between the two gauges includes three types of terms. The anticommutator on the RHS involves the commutator with only one projection operator acting on the same and different coordinate, respectively, as the commutator with the one-particle Hamiltonian. To emphasize that the effect of a commutator with $\lambda^+=1-\lambda^-$ is a projection onto negative-energy states, we prefer to write the expressions using the commutator with λ^-. For the first term on the right-hand side of (8) we get

$$\delta E_1=\frac{e^2}{2c^2}\langle\Psi|\{[V_{12},\lambda_1^-],[h_1,F_0]\}|\Psi\rangle+1\leftrightarrow2 ,$$

where $1\leftrightarrow2$ denotes the analogous expression with the commutators in coordinate 2. The leading term in the commutator with $V_{12}=e^2/r_{12}$ comes from $-\boldsymbol{\alpha}\cdot\mathbf{p}/2mc$ in λ^-. In the commutator with F, only the term $c\boldsymbol{\alpha}\cdot\mathbf{p}$ in h contributes. The low-frequency limit, $F_0\approx-r_{12}/2$, leads to

$$\delta E_1^0=\frac{e^4}{2c^2}\left\langle\Psi\left|\left\{\left[\frac{\boldsymbol{\alpha}_1\cdot\mathbf{p}_1}{2mc},\frac{1}{r_{12}}\right],\left[c\boldsymbol{\alpha}_1\cdot\mathbf{p}_1,-\frac{r_{12}}{2}\right]\right\}\right|\Psi\right\rangle+1\leftrightarrow2=\frac{e^4}{2mc^2}\left\langle\Psi\left|\frac{1}{r_{12}^2}\right|\Psi\right\rangle , \tag{10}$$

which is $O(Z^2\alpha^2)$ (in Hartree atomic units). Here one factor of 2 arises from the anticommutator and one from the identical term with the commutators in coordinate 2. (A similar expression is given by Sucher[16] in his discussion of the gauge noninvariance in the low-frequency limit.) Clearly, the low-frequency limit is not valid between positive- and negative-energy states, but it may be surprising that this leads to a discrepancy also when the approximation is used explicitly only between the positive-energy states in Ψ, in spite of the fact that the $O(\alpha^2)$—or rather $O(Z^3\alpha^2)$—corrections to the matrix elements have been included. The term involves an intermediate sum in coordinate 2 over positive- as well as negative-energy states and we will investigate the separate contributions from these terms before turning to the single commutators involving different coordinates.

The part involving *two electron-positron pairs in (10)* cancel those from the last term in (8) and can be written as

$$\delta E_1^- \approx \frac{e^2}{2c^2}\langle\Psi|[[V_{12},\lambda_1^-],[\lambda_2^-,[h_1,F_0]]]|\Psi\rangle + 1\leftrightarrow 2 \ . \tag{11}$$

Moving the commutator with λ_2 to the right does not change the result. Evaluating the commutators for the low-frequency limit of F leads to the terms with the radial dependence $1/r_{12}^4$ but also to the singular term

$$\delta E_1^- \approx \frac{\pi e^4}{2m^3c^4}\left\langle\Psi\left|(\beta_1+\beta_2)\delta(\mathbf{r}_{12})\frac{1}{r_{12}}\right|\Psi\right\rangle \ . \tag{12}$$

However, as pointed out, e.g., by Brown,[17] the creation of virtual electron-positron pairs in effect makes each electron occupy a volume proportional to $(\hbar/mc)^3$. If we use $r_{12}\approx\hbar/2mc$, we see that the contributions involving two negative-energy states becomes $O(Z\alpha^3)$ and we conclude that the contribution involving one positive- and one negative-energy state is dominated by the $O(Z^2\alpha^2)$ contribution in (10). In Sec. II C, we investigate a few diagrams involving negative-energy states, and find that, although they remove the discrepancy to leading order if the low-frequency limit of the operator is used, this does *not* hold when the proper frequency dependence is introduced.

The second type of terms in (8) involves commutators with the Hamiltonian and the projection operators acting on different coordinates:

$$\delta E_2 = \frac{e^2}{2c^2}\langle\Psi|\{[V_{12},\lambda_1^-],[h_2,F_0]\}|\Psi\rangle + 1\leftrightarrow 2 \ . \tag{13}$$

To evaluate this expression, note first that the projection operator in coordinate 1 can equally well be applied to the commutator with F, giving

$$\delta E_2 = \frac{e^2}{2c^2}\langle\Psi|\{V_{12},[\lambda_1^-,[h_2,F_0]]\}|\Psi\rangle + 1\leftrightarrow 2 \approx \frac{-2e^4}{mc^2}\left\langle\Psi\left|\frac{1}{r_{12}}\left[\frac{-\boldsymbol{\alpha}_1\cdot\boldsymbol{\alpha}_2}{2r_{12}}+\frac{(\boldsymbol{\alpha}_1\cdot\mathbf{r}_{12})(\boldsymbol{\alpha}_2\cdot\mathbf{r}_{12})}{2r_{12}^3}\right]\right|\Psi\right\rangle \ . \tag{14}$$

Here, we recognize a cross term between the Coulomb interaction and the retardation part of the Breit interaction. The presence of $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$ makes this contribution $O(Z^2\alpha^2)$ smaller than (10), i.e., $O(Z^4\alpha^4)$. For positive-energy states in coordinate 2 the low-frequency limit should be valid. The leading contribution involving positive intermediate states arises from the Coulomb gauge expression. However, to get the contributions involving only positive-energy states, we must subtract the contributions from negative intermediate states for this limit, given in (11) and (12), which is $O(\alpha^3)$. In this case, both the contributions from positive- and negative-energy states in coordinate 2 are $O(\alpha^3)$, but their sum is $O(\alpha^4)$.

The last term in (8), which involves negative-energy intermediate states in both coordinates, has already been given, albeit with opposite sign, in (12), and we have

$$\delta E^{F-C} \approx \delta E_1 + \delta E_2 - \delta E_1^- \ ,$$

where the first term on the RHS is $O(Z^2\alpha^2)$, as shown in (10), and the other two terms are $O(Z^4\alpha^4)$ and $O(Z\alpha^3)$, given in (14) and (12), respectively. The total estimate of the difference between the two gauges is dominated by the $O(Z^2\alpha^2)$ term δE_1, given in (10), i.e.,

$$\delta E^{F-C} \approx \frac{e^4}{2mc^2}\left\langle\Psi\left|\frac{1}{r_{12}^2}\right|\Psi\right\rangle \ ,$$

where Ψ is obtained within the *no-pair* approximation. In Sec. III we investigate the gauge dependence of second-order terms involving negative-energy states, but first we discuss briefly the additional terms that may arise from commutators with the one-body potential v.

C. Noncommuting potentials

In the analysis above, only the $c\boldsymbol{\alpha}\cdot\mathbf{p}$ part of h was used in the inner commutators. To find the desired form of the outer commutators, we note that for a general single-particle potential, the total Hamiltonian is given by

$$H = h_1 + h_2 + V_{12} - v_1 - v_2 \ ,$$

where only h_1 and h_2 commute with the projection operators. After addition of the vanishing commutator with V_{12} the expression for the gauge discrepancy in the no-pair approximation can then be rearranged as

$$\delta E^{F-C}=\frac{e^2}{2c^2}\langle\Psi|[h_1+h_2-v_1-v_2+V_{12},[c\boldsymbol{\alpha}_1\cdot\mathbf{p}_1+c\boldsymbol{\alpha}_2\cdot\mathbf{p}_2,F_0]]|\Psi\rangle$$
$$+\frac{e^2}{2c^2}\langle\Psi|[h_1,[v_1,F_0]]+[v_1,[c\boldsymbol{\alpha}_1\cdot\mathbf{p}_1,F_0]]+1\leftrightarrow 2|\Psi\rangle\ , \tag{15}$$

where also the v_1 and v_2 terms give $O(Z^2\alpha^2)$ contributions from the projections onto negative-energy states in coordinate 1 and 2, respectively. For a state-independent local potential, the correction term [the second expectation value in (15)] vanishes, but in general gives contributions in $O(Z^2\alpha^2)$.

It is interesting to note in this context that a multiconfiguration Dirac-Fock wave function is a "no-pair" wave function in a stronger sense than usually implied. By construction, an MCDF wave function will not get any contributions from the terms above involving one electron and one positron, if λ^+ is approximated by a projection λ^+(MCDF) onto orbitals included in the MCDF procedure:

$$\lambda_1^+(\mathrm{MCDF})\lambda_2^- H|\Psi(\mathrm{MCDF})\rangle$$
$$=\lambda_1^+(\mathrm{MCDF})\lambda_2^-(h_1+h_2-v_1-v_2-V_{12})$$
$$\times|\Psi(\mathrm{MCDF})\rangle$$
$$=0\ ,$$

since these terms would, in effect, be single excitations which vanish according to Brillouin's theorem (see, e.g., Ref. 18). Alternatively, we might say that these single excitations into negative-energy states have been absorbed in the definition of the orbitals. (Single excitations involving the creation and subsequent deletion of an electron in a third coordinate may, of course, still occur.) In this case we could not have any contributions from δE_1 in (10), but on the other hand, the nonlocal potential gives rise to additional $O(\alpha^2)$ terms from the additional commutators in (15).

To analyze the situation for a frequency-dependent F, we may introduce an operator with a rudimentary frequency dependence,

$$X_\omega=[c\boldsymbol{\alpha}\cdot\mathbf{p},F_\omega]\ ,$$
$$X_\omega\approx X'=\lambda^+X_0\lambda^+ +\lambda^+X_{2mc}\lambda^- +\lambda^-X_{2mc}\lambda^+ +\lambda^-X_0\lambda^-$$
$$=X_0+\lambda^-(X_{2mc}-X_0)\lambda^+ +\lambda^+(X_{2mc}-X_0)\lambda^-\ . \tag{16}$$

Selecting the frequency outside the commutator is correct for the Coulomb gauge expression, but not for Feynman gauge. In the latter case the second part of the correction term $[v,X_0]$ in (15) must be modified:

$$[v,X_0]\rightarrow[v,X_0]+\lambda^-[v,X_{2mc}-X_0]\lambda^+$$
$$+\lambda^+[v,X_{2mc}-X_0]\lambda^-\ .$$

These commutators still vanish for a local potential. Note also that the additional commutators do not contribute to an expectation value between positive-energy states.

The contributions δE_1 in (10) are no longer of $O(\alpha^2)$ for the more physical frequency dependence in X', as seen in Sec. III. However, the gauge dependence for a calculation involving only positive-energy states can clearly not be affected by a change in the interaction involving only the negative-energy states: If X' rather than X is used we must also subtract the nonvanishing commutator with V_{12}:

$$[V_{12},X_i']=[V_{12},X_{0,i}]+[V_{12},\lambda_i^-(X_{2mc}-X_0)_i\lambda_i^+]$$
$$+[V_{12},\lambda_i^+(X_{2mc}-X_0)_i\lambda_i^-]\neq 0\ ,$$

which includes the leading contribution (10) to the discrepancy. The gauge dependence in the no-pair approximation is of course not changed by this frequency dependence, and inclusion of the full frequency dependence will only give effects of $O(Z^5\alpha^4)$ for the moderate energy transfers between positive-energy states.

III. RECOVERY OF DISCREPANCIES?

We can now compare the contributions from the expression for the gauge discrepancy to the differences in the crossed second-order contributions (with one Coulomb interaction and one Breit interaction, in the respective gauges), involving negative-energy states. We let the final state be represented by *cd* and the initial state by *ab*, and assume that $(\varepsilon_a+\varepsilon_b)=(\varepsilon_c+\varepsilon_d)$. For completeness, we give also the expression for the second-order no-pair contribution. We consider explicitly only the contributions from V_{12}, although the crossed second-order effect between v_i and the transverse interaction gives contributions of the same order, as noted also by Mittlemann.[19]

A. The no-pair contribution

The diagram in Fig. 1(a) involves two electrons in intermediate states, r and s. It contributes

$$\delta E^{++}=\sum_{r,s}\frac{\langle cd|V_{12}|rs\rangle\left\langle rs\left|\frac{e^2}{2c^2}[h_1+h_2,[h_1,F_{\omega_1}]+[h_2,F_{\omega_2}]]\right|ab\right\rangle}{\varepsilon_a+\varepsilon_b-\varepsilon_r-\varepsilon_s}+\mathrm{H.c.}\ ,$$

where H.c. denotes the corresponding terms with the interactions interchanged. Here, we have preferred to use the commutator with h_1+h_2, rather than with the full Hamiltonian. It is easily seen that this cancels exactly the energy denominator, giving an extra minus sign. If the frequency dependence in F is neglected, the summation over the positive-energy states r and s can be replaced by λ^+ for the respective coordinates, provided the summation is extended to include ab and ba as well, and it is easy to see that these terms vanish, since they lead to an expectation value of an anti-Hermitian operator. We thus get

$$\delta E^{++}=-\frac{e^2}{2c^2}\langle cd|V_{12}\lambda_1^+\lambda_2^+([h_1,F_0]+[h_2,F_0])|ab\rangle + \text{H.c.}\ , \tag{17}$$

which is analogous to the $E\Lambda^{++}\Psi$ terms in the full expectation value given in Sec. II A.

B. Contributions involving single electron-positron pairs

The diagram in Figs. 1(b) and 1(c) involves the creation of one electron-positron pair. Using standard MBPT rules we find

$$\delta E^{-+}=-\sum_{i^-,r}\frac{\langle i^-d|V_{12}|ar\rangle\left\langle cr\left|\frac{e^2}{2c^2}[h_1+h_2,[h_1,F_{\omega_1}]+[h_2,F_{\omega_2}]]\right|i^-b\right\rangle}{\varepsilon_i+\varepsilon_b-\varepsilon_c-\varepsilon_r}+\text{H.c.}$$

The extra minus sign enters because the diagram has one internal down-going line. Again, the commutator with h_1+h_2 cancels exactly the energy denominator (and the overall sign). After removing the outer commutator, we use the Hermitian and anti-Hermitian properties, respectively, of V_{12} and the inner commutator with h_1 and rewrite the expression as

$$\delta E^{-+}=\frac{e^2}{2c^2}\sum_{i^-,r}\langle i^-d|V_{12}|ar\rangle\langle cr|[h_1,F_{\omega_1}]+[h_2,F_{\omega_2}]|i^-b\rangle+\text{H.c.}$$

$$=\frac{e^2}{2c^2}\sum_{i^-,r}\langle a^*d|V_{12}|i^{-*}r\rangle\langle i^{-*}r|-[h_1,F_{\omega_1}]+[h_2,F_{\omega_2}]|c^*b\rangle+\text{H.c.}$$

$$\approx\frac{e^2}{2c^2}\sum_{i^-,r}\langle a^*d|V_{12}\lambda_1^-\lambda_2^+(-[h_1,F_{2mc}]+[h_2,F_0])|c^*b\rangle+\text{H.c.}$$

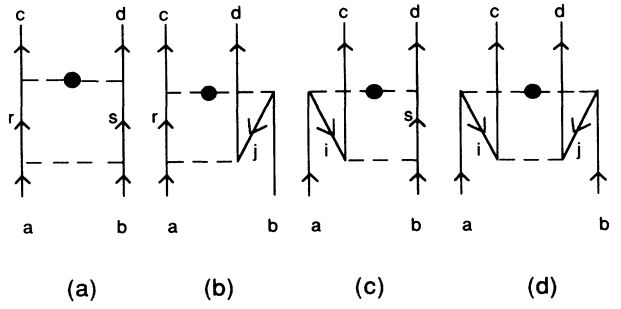

FIG. 1. Diagrammatic representation of the second-order energy contributions discussed in Sec. III. A dashed horizontal line represents the instantaneous Coulomb interaction and the horizontal line with dashes and dots represents the transverse interaction and retardation corrections. Up-going lines represent electronic states, whereas positron states are denoted by down-going lines. The letters a, b, c, and d denote the electronic states occupied in the unperturbed wave function. (a) is the no-pair contribution; (b) and (c) [(d)] contain excitations of single [double] virtual electron-positron pairs.

To get the last relation we have replaced the summations over positive (negative) intermediate states by the corresponding projection operators and approximated the frequency dependence in F. If we interchange a and c to get the same order as in the expression (17) for δE^{++} the energy resulting from the commutator with h_1 will involve ε_a rather than ε_c, but since it already involves the energy $\approx 2mc^2$ from a negative-energy state, this will only introduce a relative error $O(Z^2\alpha^2)$. Apart from this difference, the matrix element is unchanged and the contribution can thus be rewritten

$$\delta E^{-+}\approx\frac{e^2}{2c^2}(-\langle cd|V_{12}\lambda_1^-\lambda_2^+[h_1,F_{2mc}]|ab\rangle + \langle cd|V_{12}\lambda_1^-\lambda_2^+[h_2,F_0]|ab\rangle)+\text{H.c.} \tag{18}$$

Both contributions in (18) are dominated by h_1 in the external commutator used to remove the energy denominator. The first term is thus dominated by the Feynman gauge and the second by the Coulomb gauge contribu-

tions. As seen in Sec. II B [Eq. (10)], the first term is $O(Z^2\alpha^2)$ for the low-frequency limit (i.e., when F_{2mc} is replaced by F_0). Subtracting a one-body potential v_1 from V_{12} gives additional terms of the same order.

The low-frequency limit is, however, not a good approximation when the momentum transfer is $O(2mc)$. Using instead the approximate frequency-dependent form gives

$$\delta E_1^{\omega} \approx \frac{e^4}{2c^2}\left\langle \Psi \left| \left\{ \left[\frac{\boldsymbol{\alpha}_1\cdot\mathbf{p}_1}{2mc}, \frac{1}{r_{12}} \right], [c\boldsymbol{\alpha}_1\cdot\mathbf{p}_1, F_{2mc}] \right\} \right| \Psi \right\rangle + 1\leftrightarrow 2$$
$$= \frac{-e^4}{4m^3c^4}\left\langle \Psi \left| \frac{1-2mcr_{12}\sin(2mcr_{12})}{r_{12}^4} \right| \Psi \right\rangle \tag{19}$$

instead of (10). Here, the first term is $O(Z^4\alpha^4)$ and the second appears to be $O(Z^3\alpha^3)$ but is reduced by about one order $Z\alpha$ by the rapid oscillations of the sine function. The expression in (19) gives the sum of contributions from positive- and negative-energy states in coordinate 2. To estimate the part arising from positive-energy states, we must subtract the contribution involving negative-energy states in both coordinates. For the frequency-dependent form of F, they include instead of (12) the highly singular term

$$\delta E_1^{--}(2mc) \approx \frac{\pi^2 e^4}{m^5c^6}\langle \Psi | \delta^2(\mathbf{r}_{12}) | \Psi \rangle \tag{20}$$

[together with the expectation value of operators of order $O(\alpha^4)$ and $O(\alpha^5)$ times rapidly oscillating functions]. The square of the δ function would give an infinite integral, but since the δ function is actually spread out[16] over a volume proportional to $(\hbar/mc)^3$ the effect is instead $O(Z^3\alpha^3)$. The contributions from positive intermediate states in coordinate 2 must then also be of this order, but with opposite sign to reproduce the sum given in (19). The size of the second term in (18) is discussed in Sec. II B [Eq. (14)].

The terms in (18) are thus $O(Z^3\alpha^3)$ and $O(Z\alpha^3)$, respectively [with additional contributions in $O(Z^4\alpha^4)$], if the frequency-dependent form of F is used, whereas for the low-frequency limit, the first term is instead $O(Z^2\alpha^2)$, with contributions to this order only for the Feynman gauge. The minus sign for the first term in (18) makes it cancel corresponding terms in (7) if the low-frequency limit is used—the $O(\alpha^2)$ terms are thus recovered. The second term in (18) is instead added to that in (7), giving an extra factor of 2.

C. Contributions involving two electron-positron pairs

The contribution from the diagram in Fig. 1(d), which involves two negative-energy states is obtained in a similar way,

$$\sum_{i^-,j^-} \frac{\langle cd|V_{12}|i^-j^-\rangle \left\langle i^-j^- \left| \frac{e^2}{2c^2}[h_1+h_2,[h_1,F_{\omega_1}]+[h_2,F_{\omega_2}]] \right| ab \right\rangle}{\varepsilon_{i^-}+\varepsilon_{j^-}-\varepsilon_c-\varepsilon_d} + \text{H.c.}$$

Again, the commutator with (h_1+h_2) cancels the energy denominator since we assumed that $(\varepsilon_a+\varepsilon_b)=(\varepsilon_c+\varepsilon_d)$, but this time without giving an extra minus sign. With an approximation of the frequency dependence, the expression can then be rewritten as

$$\frac{e^2}{2c^2}\langle cd|V_{12}\lambda_1^-\lambda_2^-([h_1,F_{2mc}]+[h_2,F_{2mc}])|ab\rangle + \text{H.c.}$$

As seen above, this expression is $O(Z^3\alpha^3)$ for the frequency-dependent form of F and $O(Z\alpha^3)$ if the low-frequency limit is used. Again, the plus sign adds to the corresponding terms in the expression (7) for the gauge noninvariance.

D. Gauge dependence for the sum of the diagrams in Fig. 1

We can now sum all contributions to the gauge discrepancy. If we, albeit incorrectly, use the low-frequency limit $F=F_0$ everywhere in the expressions for $(\delta E^{++}+\delta E^{+-}+\delta E^{-+}+\delta E^{--})$ and add the vanishing expectation value of the commutator $[V_{12},[h_1+h_2,F_0]]$, we get the following noncancelling parts:

$$\frac{e^2}{c^2}\langle cd|V_{12}\{(\lambda_1^+ +\lambda_1^-)\lambda_2^-[h_1,F_0]+\lambda_1^-(\lambda_2^- +\lambda_2^+)[h_2,F_0]\}|ab\rangle + \text{H.c.}$$
$$= \frac{e^2}{c^2}\langle cd|\{[V_{12},\lambda_1^-],[h_2,F_0]\}+\{[V_{12},\lambda_2^-],[h_1,F_0]\}|ab\rangle \ . \tag{21}$$

An analogous expression can be obtained for matrix elements also between higher-order wave functions, rather than just the products of unperturbed orbitals. The commutator with (h_1+h_2) will then give only an approximate cancellation of the relevant energy denominators, correct to relative order $O(Z^2\alpha^2)$. An estimate of the expectation value in (21) was given in Sec. II. As seen in (14), this expectation value is $O(Z^4\alpha^4)$. If, instead, we keep the proper approximations of the frequency dependence in the expressions for $(\delta E^{+-}+\delta E^{-+}+\delta E^{--})$, then we get additional terms:

$$\frac{e^2}{2c^2}\langle cd|V_{12}\{(\lambda_1^+-\lambda_1^-)\lambda_2^-[h_1,F_0-F_{2mc}] + \lambda_1^-(\lambda_2^+-\lambda_2^-)[h_2,F_0-F_{2mc}]\}|ab\rangle+\text{H.c.}$$

The leading contributions from the commutator with F_0 are of order $O(Z^2\alpha^2)$, but the commutator with F_{2mc} contributes only in order $O(Z^3\alpha^3)$, as discussed in connection with Eqs. (10) and (18), respectively. Thus the diagrams discussed here cannot cure the $O(Z^2\alpha^2)$ discrepancy in the no-pair result.

E. Other gauge-dependent terms?

Since the diagrams considered above cure the gauge discrepancy *only* if the low-frequency limit of F is used also where it is not valid, we must question the gauge dependence in the terms not considered here. It is of course well known that the diagram in Fig. 1(d) gives spurious $O(\alpha^2)$ contributions when both interactions are the low-frequency Breit or Gaunt interaction, since, e.g.,

$$\lambda_1^+\lambda_2^+\frac{\boldsymbol{\alpha}_1\cdot\boldsymbol{\alpha}_2}{r_{12}}\lambda_1^-\lambda_2^-\frac{\boldsymbol{\alpha}_1\cdot\boldsymbol{\alpha}_2}{r_{12}}\lambda_1^+\lambda_2^+\approx\lambda_1^+\lambda_2^+\left[\frac{\boldsymbol{\alpha}_1\cdot\boldsymbol{\alpha}_2}{r_{12}}\right]^2\lambda_1^+\lambda_2^+ =\lambda_1^+\lambda_2^+\frac{3}{r_{12}^2}\lambda_1^+\lambda_2^+ ,$$

which is reduced to $O(Z^2\alpha^2)$ by the energy denominator $(\approx -4mc^2)$ in the diagram. However, the contribution is reduced by at least one order in α when the frequency dependence $\exp(i\omega r_{12})$ is included for both interactions. [When two frequency-dependent interactions are included, it is not sufficient to use only the real part, $\cos(\omega r_{12})$.]

There is a close relation between the self-energy and the Breit interaction. For example, terms which violate the exclusion principle arise in the exchange matrix element of the Breit interaction, when a photon leaves electron 1, changing it from state a to b before reaching electron 2, which is then changed from state b to a [Fig. 2(a)]. As pointed out by Feynman[20] these terms are cancelled by the term in the self-energy where the photon is instead absorbed by electron 1, changing it back to state a [Fig. 2(b)]. Similar terms arise also in higher order [see, e.g., Figs. 2(c) and 2(d)]. It is not clear to us that this cancellation still occurs when the Breit interaction is treated in Feynman gauge while the self-energy screening is treated in Coulomb gauge, as was done in the work by Indelicato and co-workers.[1-3] Further, in evaluating the expectation value of (2), terms where a Coulomb interaction occurs between the emission and absorption of the transverse photon are neglected and this approximation is especially unsatisfactory for photons of low momenta.

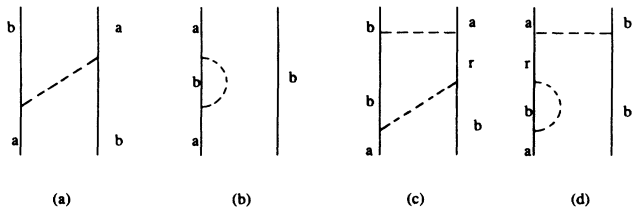

FIG. 2. Exclusion-principle violating diagrams arising from the Breit interaction [(a) and (c)], which should cancel the EPV diagrams arising from the self-energy [(b) and (d)]. Note that, e.g., (c) is not in itself gauge invariant, as discussed in Sec. II.

It is possible that there will not be an order-by-order cancellation of gauge-dependent terms. In their study or relativistic corrections to the Lamb shift, Baranger *et al.*[21] find a term in the "many-potential Lamb shift" (arising from at least two "scatterings" from the potential), which cancels exactly the terms proportional to V^2 (which are clearly gauge dependent) from the "one-potential Lamb shift." Possibly a similar cancellation will occur also for the e^4/r_{12}^2 term in (10) which describes the gauge discrepancy in the no-pair approximation. Salpeter[22] noted that for the Feynman gauge "to several orders in α, pairs of correction terms would be obtained which would finally cancel each other." It thus appears that it is often preferable to use the Coulomb gauge expression (1a). Within the no-pair approximation even the original frequency-independent form of the Breit interaction is more accurate than the frequency-dependent form (1b) in the Feynman gauge.

IV. CONCLUSION

Even when the $O(\alpha^2)$ [i.e., $O(Z^3\alpha^2)$] corrections to the matrix elements are included, a gauge discrepancy of $O(Z^2\alpha^2)$ remains in the no-pair approximation between the Coulomb and Feynman gauge forms of the electron-electron interaction. The discrepancy may be removed to leading order by including contributions in Figs. 1(b) and 1(c) from single virtual electron-positron pairs—but only if the low-frequency form $F_0=-r_{12}/2$, is used also where it is not valid.

Clearly, effects beyond the no-pair approximation are essential for heavy systems—not only to restore gauge invariance. The self-energy correction is a large effect for highly ionized systems and it may be necessary to use a gauge-dependent form to account for the screening of the other electrons. We expect that to leading order the gauge invariance will be restored when all terms involving two-photon exchange, including also the contribution involving negative-energy states, are accounted for. A possible alternative to the explicit inclusion of negative-energy states may be to use a gauge-dependent effective interaction to account approximately for the contributions from electron-positron pairs. The exact expression for the effective interaction in the different gauges requires a more detailed analysis than done here. The ac-

curate experiments for highly ionized systems make it important to resolve questions of gauge invariance and it appears that this is an intriguing question deserving more attention than it has hitherto received.

ACKNOWLEDGMENTS

Partial support for this work was provided by the Swedish Natural Science Research Council (NFR). Further, E.L. acknowledges support by the British Science and Engineering Research Council (SERC) and A.-M.M.-P. would like to express her appreciation of a stimulating time at the workshop for Quantum Electrodynamic, Relativistic and Weak Interaction effects in Atoms, Institute of Theoretical Physics, Santa Barbara, with partial support from the National Science Foundation under Grant No. PHY82-17853, supplemented by funds from the National Aeronautics and Space Administration. We are both happy to acknowledge enlightening discussions with Professor P. G. H. Sandars.

---

*Present address: Department of Physics, Chalmers University of Technology, S-412 96 Göteborg, Sweden.

[1]O. Gorceix and P. Indelicato, Phys. Rev. A **37**, 1087 (1988).

[2]O. Gorceix, P. Indelicato, and J.-P. Desclaux, J. Phys. B **20**, 639 (1987).

[3]P. Indelicato, O. Gorceix, and J.-P. Desclaux, J. Phys. B **20**, 651 (1987).

[4]I. P. Grant, J. Phys. B **20**, L735 (1987).

[5]H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of Two-Electron Atoms* (Springer, Berlin, 1957).

[6]J. B. Mann and W. R. Johnson, Phys. Rev. A **4**, 41 (1971).

[7]G. E. Brown, Philos. Mag. **XLIII**, 467 (1952).

[8]M. H. Mittlemann, Phys. Rev. A **4**, 893 (1971); **5**, 2395 (1972).

[9]I. Lindgren and J. Morrison, *Atomic Many-Body Theory,* Vol. 13 of *Springer Series in Chemical Physics,* 2nd ed. (Springer, Berlin, 1986).

[10]G. E. Brown and D. G. Ravenhall, Proc. R. Soc. London, Ser. A **208**, 552 (1951).

[11]J. Sucher, Phys. Rev. A **22**, 348 (1980); Int. J. Quant. Chem. **XXV**, 3 (1984).

[12]It is possible to add a local potential v_i to $c\boldsymbol{\alpha}_i\cdot\mathbf{p}_i$ in the commutator with F_ω since ω is defined by the states *surrounding* the commutator and v_i commutes with F_ω for any *given* ω. When studying the commuting properties of F_ω it is important to recall where ω is defined. The addition of a potential V_{12} to (h_1+h_2) in the outer commutator in (5) is discussed in Sec. II A below. Such an addition is possible as long as ω is defined outside the commutator, but $H=(h_1+h_2+V_{12})$ *cannot* be moved to the left or to the right to work directly on a possible eigenfunction Ψ since that would imply that ω was chosen inside the outer commutator in (5).

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

[14]E. Lindroth, J.-L. Heully, I. Lindgren, and A.-M. Mårtensson-Pendrill, J. Phys. B **20**, 1679 (1987).

[15]J. Sucher, Phys. Scr. **36**, 271 (1987).

[16]J. Sucher, J. Phys. B **21**, L585 (1988).

[17]G. E. Brown, Phys. Scr. **36**, 71 (1987).

[18]C. Froese Fischer, *The Hartree-Fock Method for Atoms—A Numerical Approach* (Wiley, New York, 1976) (see in particular Chap. 3.6).

[19]M. H. Mittlemann, Phys. Rev. A **5**, 2395 (1972).

[20]R. P. Feynman, Phys. Rev. **76**, 769 (1949); see also Ref. 5, Sec. 43 and P. K. Kabir and E. E. Salpeter, Phys. Rev. **108**, 1256 (1957).

[21]M. Baranger, H. A. Bethe, and R. P. Feynman, Phys. Rev. **92**, 482 (1953).

[22]E. E. Salpeter, Phys. Rev. **87**, 328 (1952).