Range of validity for perturbative treatments of relativistic sum rules

Scott M. Cohen
Department of Physics, Duquesne University, Pittsburgh, Pennsylvania 15282-0321, USA
(Received 28 January 2003; published 9 October 2003)

The range of validity of perturbative calculations of relativistic sum rules is investigated by calculating the second-order relativistic corrections to the Bethe sum rule and its small momentum limit, the Thomas-Reiche-Kuhn (TRK) sum rule. For the TRK sum rule and atomic systems, the second-order correction is found to be less than 0.5% up to about $Z=70$. The total relativistic corrections should then be accurate at least through this range of Z, and probably beyond this range if the second-order terms are included. For Rn ($Z=86$), however, the second-order corrections are nearly 1%. The total corrections to the Bethe sum rule are largest at small momentum, never being significantly larger than the corresponding corrections to the TRK sum rule. The first-order corrections to the Bethe sum rule also give better than 0.5% accuracy for $Z<70$, and inclusion of the second-order corrections should extend this range, as well.

DOI: 10.1103/PhysRevA.68.042704 PACS number(s): 34.80.Dp, 34.50.Bw

I. INTRODUCTION

Oscillator strength sum rules have long played an important role in studies of atomic and molecular systems. They are often used to check the accuracy and completeness of approximate basis sets, or to avoid calculation of slowly convergent series [1,2]. These sum rules also play a central role in the theoretical description of inelastic scattering of charged particles from atoms [3], particularly in the evaluation of integrated cross sections such as the stopping power [4–6].

Evaluation of these sum rules involves the use of closure to sum over a complete set of final states of the system under study. When this system consists of atoms with high atomic number Z a relativistic treatment of the electronic states is necessary. In such a treatment, and from the point of view of Dirac's hole theory, the sea of negative energy states is filled in the initial, ground state of the atom. According to Pauli's exclusion principle, transitions into these states are then forbidden. If the normally small contributions from pair production are neglected, only the positive energy states should be included in the theoretical description. However, in the Dirac representation of the electrons, these positive energy states do not by themselves constitute a complete set for the full space of four-component spinors. Therefore, a direct evaluation of these sum rules by the use of closure is not possible due to the lack of completeness of the available final states.

In order to overcome this difficulty, generalizations of these sum rules have utilized various methods of projecting onto the subspace of positive energy states [7–14]. These methods are perturbative, and the first-order corrections are found to be proportional to the ground-state expectation value of K_{NR}/mc^2, with K_{NR} the nonrelativistic kinetic energy operator of the electrons in the target and mc^2 the electron rest energy. For a single target electron within an atom, this result is simply proportional to α^2Z^2, which is recognized as an expansion parameter for these calculations (α is the fine-structure constant). As such, one might question the accuracy of such a perturbative approach, especially if only the first-order term is retained, unless $Z \ll 1/\alpha$. However, for a many-electron system, the kinetic energy of each electron is altered by its interaction with the other electrons, including the effects of screening of the nuclear charge. So, particularly for the outer shell electrons, the effective value of Z will be significantly reduced. Furthermore, though significant for the analysis of experiments [15,16], the first-order corrections are already small, contributing no more than a few percent of the total. For these reasons, one might instead expect that the first-order corrections are adequate for a description of even the heaviest atoms. Given these conflicting indicators and the absence of any studies that extend beyond the first-order, the question remains: What systems, and in particular, what range of values of Z, can be described by these calculations? The answer to this question is the primary aim of the present paper.

In a recent paper [17], we have presented an approach to these sum rules, which can readily be extended to find the corrections at the next order. These results will then yield direct insight into the range of validity of the first-order corrections. We show below that the second-order corrections may also be obtained in terms of ground-state expectation values, where the operators that appear at this level of approximation are those of the one-body relativistic terms in the Breit-Pauli Hamiltonian H_{BP}. These are the mass variation,

$$H_{mv} = -\frac{1}{8m^3c^2}\sum_{k=1}^{Z} p_k^4, \tag{1.1}$$

the first Darwin,

$$H_{d1} = -\frac{e\hbar^2}{8m^2c^2}\sum_{k=1}^{Z} \vec{\nabla}\cdot\vec{E}_k, \tag{1.2}$$

and the spin-orbit terms,

$$H_{so} = -\frac{e\hbar}{4m^2c^2}\sum_{k=1}^{Z} \vec{\sigma}_k\cdot(\vec{E}_k\times\vec{p}_k), \tag{1.3}$$

where $\vec{p}_k$ and $\vec{\sigma}_k$ are the momentum operator and the Pauli spin matrices [18] for the kth electron, and $\vec{E}_k = -\nabla V_k/e$ is

the field of the nucleus at that electron's position. There will also be contributions from the two-body terms in H_{BP}, those arising from the electron-electron interactions, but we have found (see Appendix A for an outline of this calculation) that these contributions appear as ground-state expectation values of precisely these same two-body operators, and are therefore two orders of magnitude smaller than those of the one-body terms even for the largest values of Z [19–21]. Therefore, these contributions should be small and will not be considered in detail in the main part of this paper.

Expectation values of the above one-body terms are no larger than about 20% that of the nonrelativistic kinetic energy operator

$$K_{NR}=\frac{1}{2m}\sum_k p_k^2. \tag{1.4}$$

Furthermore, expectation values of H_{mv} and H_{d1} are opposite in sign so that their contributions to the energy tend to cancel each other, leaving a total that is well below 10% of the kinetic energy. Thus, one might expect that their contribution to other quantities, such as the sum rules being considered here, will also be small relative to the corresponding first-order contribution involving K_{NR}. As shall be seen below, the first-order terms are adequate even for fairly large Z, but additional terms are needed as Z becomes larger than about 70, at least if accuracy of 0.5% or better is desired. Nonetheless, the perturbative approach to these sum rules should be accurate for a wide range of targets.

In the following section, we discuss the Bethe sum rule, and in Sec. III we describe our approach. Then, in Sec. IV the results are discussed, followed by our conclusions.

II. THE BETHE SUM RULE

Let us consider the Bethe sum, defined as,

$$S_1(Q)=Z^{-1}\sum_{all} E_{n0}\sum_{j=1}^{Z}|\langle n|e^{i\vec{q}\cdot\vec{r}_j/\hbar}|0\rangle|^2, \tag{2.1}$$

where $E_{n0}=E_n-E_0$ is the difference between the energy of the ground $|0\rangle$ and excited states $|n\rangle$ of the atom with Z electrons. The recoil energy, $Q=\sqrt{q^2c^2+m^2c^4}-mc^2$, is the kinetic energy of an electron with momentum $\vec{q}$. The summation in Eq. (2.1) is to be taken over all positive energy states of the many-electron atom.

Though other methods are available, such as the introduction of projection operators to exclude contributions to the sum from the negative energy states, we have found that it is advantageous to use the Foldy-Wouthuysen transformation (FWT) [22] for this purpose. In Refs. [14] and [17], a formally exact expression in terms of an initial-state expectation value has been derived for $S_1(Q)$, using an FWT to a representation in which the positive energy states are by themselves a complete set. The result found there may be written as

$$S_1(Q)=\frac{1}{2Z}\sum_{j=1}^{Z}\langle\tilde{0}|(2R^{j\dagger}\tilde{H}(\vec{q})R^j-\{R^{j\dagger}R^j,\tilde{H}\})|\tilde{0}\rangle, \tag{2.2}$$

where $\{A,B\}=AB+BA$, $|\tilde{0}\rangle=e^{iU}|0\rangle$ is the initial state in the new representation, and R^j and $\tilde{H}$ are the projections of $e^{-i\vec{q}\cdot\vec{r}_j/\hbar}e^{iU}e^{i\vec{q}\cdot\vec{r}_j/\hbar}e^{-iU}=e^{iU(\vec{q})}e^{-iU}$ and $e^{iU}He^{-iU}$ onto the subspace of positive energy eigenstates, respectively. An operator written explicitly as a function of $\vec{q}$ represents a momentum boost, such as $\tilde{H}(\vec{q})=e^{-i\vec{q}\cdot\vec{r}_j/\hbar}\tilde{H}e^{i\vec{q}\cdot\vec{r}_j/\hbar}$ and the FWT operator $U(\vec{q})$ introduced in the preceding sentence in the definition of R^j.

For H, we use a many-electron Dirac Hamiltonian,

$$H=\sum_{k=1}^{Z}H_D^k+V_{ee}, \tag{2.3}$$

where

$$H_D^k=\beta_k mc^2+c\vec{\alpha}_k\cdot\vec{p}_k+V_k \tag{2.4}$$

is the single-electron Dirac Hamiltonian, β_k and $\vec{\alpha}_k$ are the usual Dirac matrices [23], and V_k and V_{ee} represent the nuclear and electron-electron interactions, respectively [24]. Following Sucher [25], we take,

$$V_{ee}=\mathcal{L}_+\left(\sum_{k>l}^{Z}V_{kl}\right)\mathcal{L}_+, \tag{2.5}$$

with $\mathcal{L}_+=\Pi_{k=1}^{Z}\mathcal{L}_+(k)$, and $\mathcal{L}_+(k)$ is the projection operator onto the space of positive energy states of H_D^k. As has been mentioned previously, the contributions from the two-body terms arising from V_{ee} are negligible, so the specific form chosen for V_{kl} is unimportant for us here (a common choice would be the sum of Coulomb and Breit interactions; see Appendix A for further details).

We will write,

$$S_1(Q)=S_1^0(Q)[1-\Delta(Q)], \tag{2.6}$$

with $S_1^0(Q)=Q(1+Q/2mc^2)/(1+Q/mc^2)$ the result originally found by Bethe for a free electron at rest. $\Delta(Q)$ then represents the corrections arising from a treatment of the electrons as bound, relativistic particles.

For the purpose of calculation, it is convenient to separate out the various parts which are $\mathcal{O}(V)$. The transformed Hamiltonian will have the form (see following section),

$$\tilde{H}=e^{iU}He^{-iU}=\sum_{k=1}^{Z}\epsilon_k+\tilde{H}_V, \tag{2.7}$$

with $\epsilon_k=\sqrt{p_k^2c^2+m^2c^4}$. We may also write

$$R^j=R_f^j+R_V^j, \tag{2.8}$$

yielding

$$2R^{j\dagger}\tilde{H}(\vec{q})R^{j}-\{R^{j\dagger}R^{j},\tilde{H}\}=\{R_f^{j\dagger}R_f^{j},\epsilon_j(\vec{q})-\epsilon_j\}+\{R_f^{j\dagger}R_V^{j}+R_V^{j\dagger}R_f^{j},\epsilon_j(\vec{q})-\epsilon_j\}+\{R_f^{j\dagger}R_f^{j},\tilde{H}_V(\vec{q})-\tilde{H}_V\}+R_f^{j\dagger}([\epsilon_j(\vec{q}),R_V^{j}]+[H_V(\vec{q}),R_f^{j}])+([R_V^{j\dagger},\epsilon_j(\vec{q})]+[R_f^{j\dagger},H_V(\vec{q})])R_V^{j},\tag{2.9}$$

where R_f^j and the first term in the latter equation are $V=0$ or ''free'' parts, and the other terms are $\mathcal{O}(V)$. After averaging over the direction of $\vec{q}$, or equivalently, the orientation of the target, the calculations described in the following section yield

$$\Delta(Q)=\Delta_f(Q)+\Delta_V(Q),\tag{2.10}$$

with

$$\Delta_f(Q)=\frac{2}{3mc^2Z(1+Q/mc^2)^2}\left[1+\frac{3}{2(1+Q/mc^2)^2}\right]\times\langle\tilde{0}|K_{NR}|\tilde{0}\rangle+\frac{7}{mc^2Z(1+Q/mc^2)^8}\langle\tilde{0}|H_{mv}|\tilde{0}\rangle,\tag{2.11}$$

and

$$\Delta_V(Q)=-\frac{2(1+Q/3mc^2)}{mc^2Z(1+Q/mc^2)^4(1+Q/2mc^2)}\langle\tilde{0}|H_{so}|\tilde{0}\rangle.\tag{2.12}$$

Note that in $\Delta_V(Q)$, the coefficient of H_{d1} vanishes identically. This is somewhat surprising, given the complicated nature of the calculations leading to this result. It is also important, since the expectation value of H_{d1} is ordinarily significantly larger than that of H_{so}. In fact, expectation values of the latter are usually smaller than some of the two-body terms in H_{BP} that have already been neglected. Therefore, to the present level of approximation, we have

$$\Delta_V(Q)=0,\tag{2.13}$$

and the potential energy of the electrons does not appear explicitly in the result, entering only through its effects on the ground-state wavefunction of the target system.

These results give $\Delta(Q)$ to $\mathcal{O}(p^4)$, $\mathcal{O}(p^2V)$, and to all orders in Q. The methods used to obtain them are described in the following section. Those readers not interested in specifics concerning the methods may skip directly to Sec. IV.

III. METHOD OF CALCULATION

The object of the FWT is to eliminate by a canonical transformation all operators in the Hamiltonian that couple upper and lower components of the four-component, single-particle spinors. These types of operators are denoted as odd operators, whereas even operators do not couple the different components. Apart from the case of free electrons, where the exact FWT operator U is known, the approach is to follow a step-by-step procedure, yielding a suitable approximation to the exact Hamiltonian. The approximation used in Ref. [17] to obtain results valid for all Q is an expansion [26–28] in powers of the potential energy operator $V=\Sigma_k V_k+V_{ee}$. This may be written as

$$e^{iU}=\cdots e^{iU_3}e^{iU_2}e^{iU_1},\tag{3.1}$$

where U_n is $\mathcal{O}(V^{n-1})$. For the many-electron treatment being used here, we choose

$$e^{iU_1}=\prod_{k=1}^{Z}e^{iU_f^k},\tag{3.2}$$

where U_f^k is the free-electron FW operator [22]

$$e^{\pm i\hat{U}_k}=\frac{1}{\sqrt{2\epsilon_k(\epsilon_k+mc^2)}}(\epsilon_k+mc^2\pm\beta_k c\vec{\alpha}_k\cdot\hat{p}_k).\tag{3.3}$$

Note that to obtain detailed results, an expansion in powers of the momentum operator is helpful in the sense that it leads to results in the form of expectation values that are available in the literature (see Sec. IV, below). However, if the $\vec{p}_j$ expansion is implemented prior to the boost transformation of the operators, $\mathcal{M}\rightarrow\mathcal{M}(\vec{q})=e^{-i\vec{q}\cdot\vec{r}_j/\hbar}\mathcal{M}e^{i\vec{q}\cdot\vec{r}_j/\hbar}=\mathcal{M}|_{\vec{p}_j\rightarrow\vec{p}_j+\vec{q}}$, then the replacement of $\vec{p}_j$ by $\vec{p}_j+\vec{q}$ will transform the expansion $a_0+\vec{a}_1\cdot\vec{p}_j+\cdots$ directly into an expansion in powers of q [17]. Since we wish to include a description of the large-q behavior of these sums, the $\vec{p}_j$ expansion must be delayed until after the boost is introduced.

The Hamiltonian after the first step in the FWT is found to be

$$H_1=e^{iU_1}He^{-iU_1}=\sum_{k=1}^{Z}\left(\beta_k\epsilon_k+V_k+\frac{1}{2mc}[\beta_k\vec{\alpha}_k\cdot\vec{p}_k,V_k]+\frac{1}{8m^2c^2}[\beta_k\vec{\alpha}_k\cdot\vec{p}_k,[\beta_k\vec{\alpha}_k\cdot\vec{p}_k,V_k]]\right)+e^{iU_1}V_{ee}e^{-iU_1}.\tag{3.4}$$

The next step is to identify the FW operator U_2. This is to be chosen such that it eliminates odd operators from H_1 to $\mathcal{O}(V)$. Thus, with $U_2=\Sigma_k U_2^k$, we require

$$e^{iU_2^k}\beta_k\epsilon_k e^{-iU_2^k}=\beta_k\epsilon_k-t_o^k,\tag{3.5}$$

where

$$t_o^k=\frac{ie\hbar}{2mc}\beta_k\vec{\alpha}_k\cdot\vec{E}_k\tag{3.6}$$

is the odd part of the kth-electron terms in H_1 (we have assumed that V_k is itself an even operator, such as the unretarded Coulomb interaction with the nucleus or a similar effective potential). To first-order in U_2^k, this is

$$[iU_2^k,\beta_k\epsilon_k]=-t_o^k. \tag{3.7}$$

The fact that t_o^k is $\mathcal{O}(pV)$ implies that U_2^k is also, so that to $\mathcal{O}(p^2V)$, ϵ_k may be approximated as mc^2 in this equation. Noting that U_2^k is odd and hence anticommutes with β_k, this yields

$$U_2^k=\frac{e\hbar}{4m^2c^3}\vec{\alpha}_k\cdot\vec{E}_k, \tag{3.8}$$

which is in agreement with the original work of Foldy & Wouthuysen [22]. Using this result, one finds $\tilde{H}=\Sigma_{k=1}^Z\tilde{H}_k$, where to the desired level of approximation,

$$\tilde{H}_k=\epsilon_k+V_k-\frac{e\hbar^2}{8m^2c^2}\vec{\nabla}\cdot\vec{E}_k-\frac{e\hbar}{4m^2c^2}\vec{\sigma}_k\cdot(\vec{E}_k\times\vec{p}_k)+\tilde{V}_{ee}, \tag{3.9}$$

with $\tilde{V}_{ee}=e^{iU}V_{ee}e^{-iU}$.

The above results cannot be used to obtain $\tilde{H}(\vec{q})=e^{-i\vec{q}\cdot\vec{r}/\hbar}\tilde{H}e^{i\vec{q}\cdot\vec{r}/\hbar}$, since an expansion in powers of p_k has already been taken. Instead, with $\tilde{H}(\vec{q})=\Sigma_{k\neq j}^Z\tilde{H}_k+\tilde{H}_j(\vec{q})$ being the even part of $H_1(\vec{q})=e^{iU_1(\vec{q})}H(\vec{q})e^{-iU_1(\vec{q})}$, one finds directly that

$$\begin{aligned}\tilde{H}_j(\vec{q})=&\epsilon_j(\vec{q})+V_j+e\hbar^2\kappa(\vec{q}\cdot\vec{\nabla}_j)\vec{q}\cdot\vec{E}_j\\&-\frac{e\hbar^2Q}{4q^2\epsilon_q}\left(\vec{\nabla}_j\cdot\vec{E}_j+\frac{2}{\hbar}\vec{\sigma}_j\cdot[\vec{E}_j\times(\vec{p}_j+\vec{q})]-2ia_1\vec{\sigma}_j\right.\\&\left.\cdot(\vec{q}\times\nabla_j)\left[\vec{q}\cdot\vec{E}_j-\frac{2i}{e\hbar}V_j\vec{q}\cdot\vec{p}_j\right]\right)+\tilde{V}_{ee}(\vec{q}),\end{aligned} \tag{3.10}$$

with $\epsilon_q=\sqrt{q^2c^2+m^2c^4}=Q+mc^2$,

$$\kappa=\frac{Q}{4q^4\epsilon_q}-\frac{m^2c^6}{8q^2\epsilon_q^4}, \tag{3.11}$$

and

$$a_1=\frac{mc^4}{2Q\epsilon_q^2}-\frac{1}{q^2}. \tag{3.12}$$

To find $U_2(\vec{q})=U_2^j(\vec{q})+\Sigma_{k\neq j}U_2^k$, which is needed to obtain R^j, we start with the analog of Eq. (3.7),

$$[iU_2^j(\vec{q}),\beta_j\epsilon_j(\vec{q})]=-t_o^j(\vec{q}). \tag{3.13}$$

Here $t_o^j(\vec{q})$ is the odd part of the jth-electron potential term in $H_1(\vec{q})$, which is given by

$$\begin{aligned}t_o^j(\vec{q})=&\frac{c}{2}\frac{1}{\sqrt{\epsilon_j(\vec{q})}}\left[\frac{\beta_j\vec{\alpha}_j\cdot(\vec{p}_j+\vec{q})}{\sqrt{\epsilon_j(\vec{q})+mc^2}}V_j\sqrt{\epsilon_j(\vec{q})+mc^2}\right.\\&\left.-\sqrt{\epsilon_j(\vec{q})+mc^2}\,V_j\frac{\beta_j\vec{\alpha}_j\cdot(\vec{p}_j+\vec{q})}{\sqrt{\epsilon_j(\vec{q})+mc^2}}\right]\frac{1}{\sqrt{\epsilon_j(\vec{q})}}.\end{aligned} \tag{3.14}$$

This is seen to vanish when $p_j=0$, so that $U_2^j(\vec{q})$, like U_2^j itself, is $\mathcal{O}(pV)$. Expanding in powers of p, Eq. (3.13) may be solved to yield

$$U_2^j(\vec{q})=\frac{\beta_j}{2i\epsilon_q}t_o^j(\vec{q})+\frac{ic^2\beta_j}{4\epsilon_q^3}\{t_o^j(\vec{q}),\vec{p}_j\cdot\vec{q}\}. \tag{3.15}$$

An explicit expression for $U_2^j(\vec{q})$, along with expressions for other relevant quantities needed for the Bethe sum rule, Eq. (2.9), may be found in Appendix B.

With this result for $U_2^j(\vec{q})$, we may find R^j, which is the even part of $e^{iU(\vec{q})}e^{-iU}$ and may be written,

$$\begin{aligned}R^j=&\tfrac{1}{2}(e^{iU_2^j(\vec{q})}e^{iU_f^j(\vec{q})}e^{-iU_f^j}e^{-iU_2^j}\\&+\beta_je^{iU_2^j(\vec{q})}e^{iU_f^j(\vec{q})}e^{-iU_f^j}e^{-iU_2^j}\beta_j).\end{aligned} \tag{3.16}$$

Given that $\beta_je^{\pm iU_f^j}\beta_j=e^{\mp iU_f^j}$ and similarly for $e^{iU_f^j(\vec{q})}$; and that we need keep only first order in U_2^j, we find

$$R^j=R_f^j+R_V^j, \tag{3.17}$$

with the $V=0$ part of R^j given by the expression

$$\begin{aligned}\sqrt{\frac{\epsilon_q+mc^2}{2\epsilon_q}}\;R_f^j=&\frac{\epsilon_q+mc^2}{2\epsilon_q}+\frac{q^2c^2}{4m\epsilon_q^3}\vec{q}\cdot\vec{p}_j\\&+\frac{i}{4m\epsilon_q}\vec{\sigma}_j\cdot(\vec{q}\times\vec{p}_j),\end{aligned} \tag{3.18}$$

while the $\mathcal{O}(V)$ part is given to $\mathcal{O}(pV)$ as

$$\begin{aligned}\sqrt{\frac{2\epsilon_q}{Q}}\;R_V^j=&-\frac{ie\hbar}{4m^2c^3q\epsilon_q^3}[(\epsilon_q^3+m^3c^6)\vec{q}\cdot\vec{E}_j\\&+iq^2c^2\epsilon_q\vec{\sigma}_j\cdot(\vec{q}\times\vec{E}_j)].\end{aligned} \tag{3.19}$$

One may note that R_V^j is needed to $\mathcal{O}(p^2V)$ in the second term on the right-hand side of Eq. (2.9); the required expression is given in Appendix B. We also point out that

$$R_f^{j\dagger}R_f^j=\frac{1}{2}+\frac{\epsilon_j}{2\epsilon_j(\vec{q})}+\frac{c^2\vec{q}\cdot\vec{p}_j}{2\epsilon_j\epsilon_j(\vec{q})} \tag{3.20}$$

is an exact expression for the free part of $R^{j\dagger}R^j$ [14]. Combining all these results yields Eqs. (2.10)–(2.13).

IV. DISCUSSION

In this section, we consider these results to ascertain the range of Z over which they will be valid. We first note that Eqs. (2.10)–(2.13) involve expectation values of K_{NR} and H_{mv} (omitting the small contributions from H_{so} and V_{ee}) with respect to the ground state in the FWT-transformed representation. Since e^{iU_1} commutes with these operators and U_2 is $\mathcal{O}(pV)$, we see that $e^{-iU}p^n e^{iU} = e^{-iU_2}e^{-iU_1}p^n e^{iU_1}e^{iU_2} = p^n + \mathcal{O}(p^{n+1}V)$. Therefore, since we work only to $\mathcal{O}(p^2V)$,

$$\langle \tilde{0}|K_{NR}|\tilde{0}\rangle = \langle \tilde{0}|e^{iU}(e^{-iU}K_{NR}e^{iU})e^{-iU}|\tilde{0}\rangle = \langle 0|K_{NR}|0\rangle, \tag{4.1}$$

plus small corrections, and similarly for H_{mv}. The required expectation values may be evaluated with $|0\rangle = e^{-iU}|\tilde{0}\rangle$ instead of the transformed ground state, the differences being higher-order in the relativistic corrections. This will allow us to use data calculated in the Dirac representation.

Next, we separate out the first- and second-order relativistic corrections in a way that allows use of data available in the literature. Identifying the nonrelativistic part of the ground-state wavefunction as $|0_{NR}\rangle$, and writing

$$\langle 0|K_{NR}|0\rangle = \langle 0_{NR}|K_{NR}|0_{NR}\rangle + (\langle 0|K_{NR}|0\rangle - \langle 0_{NR}|K_{NR}|0_{NR}\rangle), \tag{4.2}$$

the first term on the right-hand side divided by mc^2 is seen to be $\mathcal{O}(\alpha^2)$, while the terms in parentheses give higher-order corrections. The latter may be rewritten by approximating the nonrelativistic kinetic energy operator as $K_{NR} = K - H_{mv}$ in the first term, so that

$$\langle 0|K_{NR}|0\rangle = \langle 0_{NR}|K_{NR}|0_{NR}\rangle + (\langle 0|(K - H_{mv})|0\rangle - \langle 0_{NR}|K_{NR}|0_{NR}\rangle). \tag{4.3}$$

This accomplishes the separation of the first-order relative corrections from higher corrections, and is in terms of expectation values available in the literature. Thus, from Eqs. (2.10)–(2.13), we have

$$\Delta(Q) = \Delta_1(Q) + \Delta_2(Q) \tag{4.4}$$

with the first-order contributions given by

$$\Delta_1(Q) = \frac{2}{3mc^2Z(1+Q/mc^2)^2}\left[1 + \frac{3}{2(1+Q/mc^2)^2}\right] \times \langle 0_{NR}|K_{NR}|0_{NR}\rangle, \tag{4.5}$$

while the second-order corrections are

$$\Delta_2(Q) = \frac{2}{3mc^2Z(1+Q/mc^2)^2}\left[1 + \frac{3}{2(1+Q/mc^2)^2}\right] \times (\langle 0|(K - H_{mv})|0\rangle - \langle 0_{NR}|K_{NR}|0_{NR}\rangle) + \frac{7}{mc^2Z(1+Q/mc^2)^8}\langle 0|H_{mv}|0\rangle. \tag{4.6}$$

also note that the TRK sum, defined as

$$S_1 = \frac{2m}{\hbar^2 Z}\sum_{all} E_{n0}\Big|\sum_{j=1}^{Z}\langle n|z_j|0\rangle\Big|^2, \tag{4.7}$$

may be obtained directly from our results for the Bethe sum, as

$$S_1 = \lim_{q\to 0}\frac{2m}{q^2}S_1(Q) = 1 - \Delta, \tag{4.8}$$

where $\Delta = \Delta_1 + \Delta_2$, with

$$\Delta_1 = \frac{5}{3mc^2Z}\langle 0_{NR}|K_{NR}|0_{NR}\rangle, \tag{4.9}$$

and

$$\Delta_2 = \frac{5}{3mc^2Z}(\langle 0|K|0\rangle - \langle 0_{NR}|K_{NR}|0_{NR}\rangle) + \frac{16}{3mc^2Z}\langle 0|H_{mv}|0\rangle. \tag{4.10}$$

One may note that the above result for Δ_1 has long been known [7–11], but that Δ_2 does not seem to have been given in the literature, previously. Also, while there have been a few studies of the corrections to the Bethe sum rule that have given results as an expansion in powers of Q [12–14], the exact Q dependence for $\Delta_1(Q)$ in Eq. (4.5) was only recently obtained [17]. The small-Q limit of Eq. (4.5) is in agreement with the earlier works utilizing the FWT approach [12,14]. It is not easy, however, to make a direct comparison to the $\Delta_1(Q)$ derived in Ref. [13], since their result is still in the form of a sum over final states, while all our results are given simply as ground-state expectation values. The expression for $\Delta_2(Q)$ is new.

Data for all expectation values appearing in these equations may be found in the literature for a wide range of atoms. For $\langle 0_{NR}|K_{NR}|0_{NR}\rangle$, values are obtained as the difference between the total nonrelativistic energy and the sum of the individual electron potential energies, as given in Ref. [29]; for $\langle 0|K|0\rangle$, the results of Ref. [30] will be used; and values for $\langle 0|H_{mv}|0\rangle$ are calculated from the data of Ref. [31]. In Fig. 1, our results in Eqs. (4.9) and (4.10) are plotted as a function of Z to illustrate the relative magnitudes of the two correction terms in the TRK sum rule (the equations have been evaluated for $Z = 2$, integer multiples of 10 up to 80, and 86; the lines in the figure are interpolated from these values). One may note that Δ_1 and Δ_2 are opposite in sign,

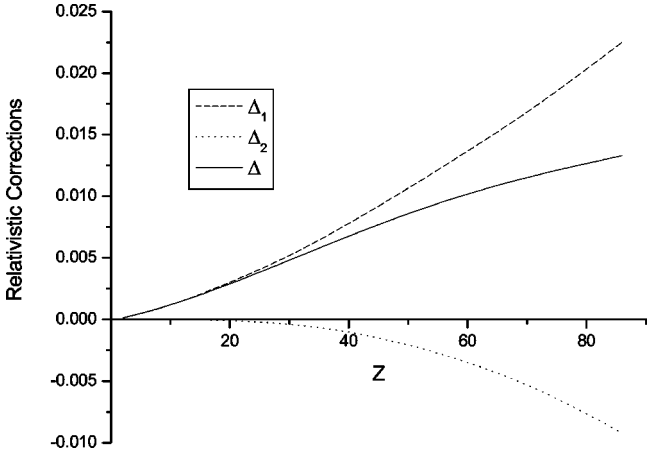

FIG. 1. The relativistic corrections to the TRK sum rule, plotted as a function of Z to illustrate the relative magnitudes of the two correction terms (the equations have been evaluated for $Z=2$, integer multiples of 10 up to 80, and 86; the lines in the figure are interpolated from these values).

so tend to cancel, making the total correction always smaller than the first-order corrections. It is seen that Δ_1 is less than 0.5% for $Z<30$ and the total correction Δ differs little from this in that range. Δ_2 is smaller than 0.5% in magnitude out almost to $Z=70$. Therefore, our results for Δ should offer reasonable accuracy at least through this range of Z, and probably beyond if the second-order corrections are included.

In Fig. 2, we plot the relativistic corrections to the Bethe sum rule as a function of Q/mc^2 for Ytterbium, $Z=70$. We see that $\Delta_1(Q)$ decreases monotonically with increasing Q, while $\Delta_2(Q)$ has its largest magnitude at $Q=0$ and changes sign when Q is a relatively small fraction of mc^2. The total correction $\Delta(Q)$ is very well approximated by $\Delta_1(Q)$, except perhaps at the smallest values of Q. Similar trends are observed for other atoms, as well. In Fig. 3, the second-order corrections to the Bethe sum rule are plotted as a function of Q/mc^2 for a few of the heavier atoms. One may note that,

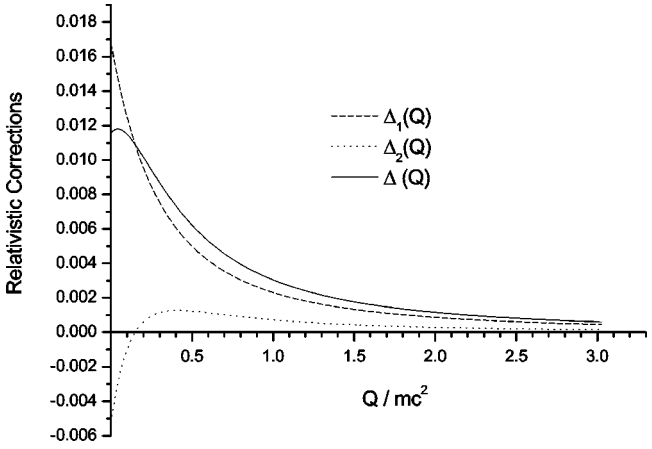

FIG. 2. Comparison of the first- and second-order and total relativistic corrections to the Bethe sum rule, plotted as a function of Q/mc^2, for Ytterbium ($Z=70$).

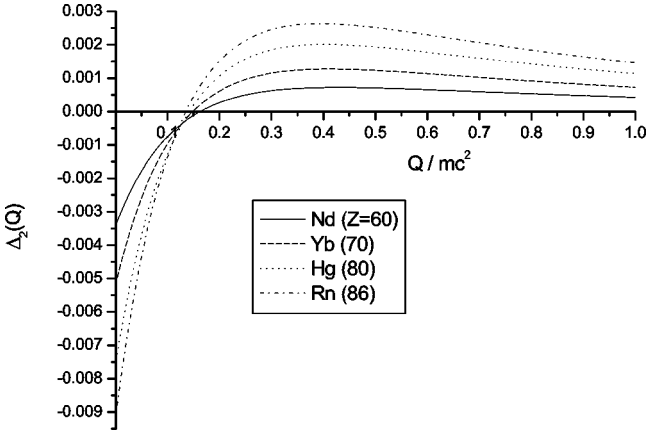

FIG. 3. The second-order corrections to the Bethe sum rule, plotted as a function of Q/mc^2, for $Z=60$, 70, 80, 86.

even for Rn ($Z=86$), these corrections are small, except at the lower range of Q, say $Q<mc^2/25$. The perturbative results for the Bethe sum rule should therefore be quite accurate, except for large Z and small Q.

V. CONCLUSION

In this paper, we have extended previous calculations of the Bethe and TRK sum rules to include second-order relativistic corrections from an expansion in powers of the binding energy of the target electrons. A many-electron treatment of the atoms has been used in an approach that is valid irrespective of the size of the recoil energy Q. The results presented here indicate that perturbative calculations of these sum rules yield accurate results over a wide range of Z. Further work is needed, however, when $Z>70$ and Q is small, if one wishes to achieve an accuracy of 0.5% or better.

We now close with some observations about electron-electron interactions in many-electron systems, and their contributions to these calculations. It is well known that in a nonrelativistic treatment of the target, the Bethe sum is found to be proportional to a ground-state expectation value of a constant

$$S_1(Q)=Z^{-1}\sum_{all}\sum_{j=1}^{Z}\langle 0|e^{-i\vec{q}\cdot\vec{r}_j/\hbar}|n\rangle\langle n|[H_{NR},e^{i\vec{q}\cdot\vec{r}_j/\hbar}]|0\rangle$$

$$=\langle 0|Q(1+Q/2mc^2)|0\rangle=Q(1+Q/2mc^2). \qquad (5.1)$$

This means, of course, that the result is independent of the ground state of the target, and different systems—free electrons, a single bound electron, or a system of many interacting, bound electrons—all yield the same $S_1(Q)$. Thus one concludes that electron-electron interactions, if present, do not alter the result at all. As is discussed in the following paragraph, however, this conclusion no longer holds for a relativistic treatment.

Earlier studies of the relativistic case have generally relied upon an independent-particle approximation (IPA) [12–14,32], for which the electron-electron interactions [V_{ee} in Eq. (2.3)] are dropped from the Hamiltonian of the system (though these interactions have previously been discussed in

the context of an effective potential [13]). The approach presented here formally avoids use of the IPA and offers a clear picture of the effects of including V_{ee} in the Hamiltonian H. We have indicated in the Introduction that, apart from their effects on the ground-state wavefunction (which are, of course, important) these two-body interactions lead to sum rule corrections which are much smaller than those from the one-body terms included in our calculations. In fact, by including the two-body terms explicitly in the calculation of these sum rules, we have been able to show (see Appendix A) that the additional terms that arise are all proportional to expectation values of the same two-body operators that appear in the Breit-Pauli Hamiltonian H_{BP}, such as the second Darwin, orbit-orbit, and other operators (the operator V_{ee}, by itself, does not appear). Since it is known that expectation values of these two-body operators in H_{BP} are quite small relative to those of the corresponding one-body operators [19–21], their contributions to the sum rules should then also be quite small. Therefore, our results would be unchanged if H were replaced by an independent-particle Hamiltonian, as long as the electron-electron interactions are still included in the ground-state wavefunction. One can therefore see why the IPA leads to good agreement with experiment in Ref. [32]. Nonetheless, it is important to remember that the effects of these interactions on the wavefunctions are important in the relativistic case considered in this paper. Though these effects are important, it should also be noted that different choices for V_{ee} are not expected to alter our numerical results significantly.

ACKNOWLEDGMENTS

The author wishes to acknowledge the donors of The Petroleum Research Fund, administered by the ACS, for support of this research.

APPENDIX A

In this Appendix, we outline how contributions from the two-body terms in the Hamiltonian may be calculated. For this discussion, we take [see Eq. (2.5)]

$$V_{kl}=V_{kl}^{C}+V_{kl}^{B}, \tag{A1}$$

where $V_{kl}^{C}=e^2/r_{kl}$ is the Coulomb interaction between electrons k and l, and

$$V_{kl}^{B}=-\frac{e^2}{2r_{kl}}(\vec{\alpha}_k\cdot\vec{\alpha}_l+\vec{\alpha}_k\cdot\hat{r}_{kl}\vec{\alpha}_l\cdot\hat{r}_{kl}), \tag{A2}$$

is the Breit operator.

We now wish to introduce the FWT. Once again we write

$$e^{iU}=\cdots e^{iU_3}e^{iU_2}e^{iU_1}, \tag{A3}$$

and choose $U_1=\Pi_k U_f^k$ and U_2 to eliminate the odd parts from the V_k terms in $H_1=e^{iU_1}H_1e^{-iU_1}$, as before [see Eqs. (3.2)–(3.7)]. One might expect that this will leave odd operators that arise from V_{kl} that would then also need to be eliminated. However, it turns out that U_2 along with the presence of the projection operator $\mathcal{L}_+$ suffices to eliminate all such odd operators.

This latter conclusion may be seen as follows: First, note that since $\mathcal{L}_+(k)$ are projection operators onto the space of positive energy states of the single-electron Hamiltonians, H_D^k, then the FWT of these projection operators must be equal to the projection operators in the transformed representation. That is,

$$e^{iU_2^k}e^{iU_f^k}\mathcal{L}_+(k)e^{-iU_f^k}e^{-iU_2^k}=\tfrac{1}{2}(1+\beta_k). \tag{A4}$$

Of course, this is only completely correct if the exact U_2 is used, but the error incurred by approximating U_2 as we have done throughout this paper will be higher order in small quantities, and so may be ignored. Therefore,

$$\begin{aligned}e^{iU_2}e^{iU_1}\mathcal{L}_+(k)\mathcal{L}_+(l)V_{kl}\mathcal{L}_+(k)\mathcal{L}_+(l)e^{-iU_1}e^{-iU_2}&\\ =\tfrac{1}{16}(1+\beta_k)(1+\beta_l)(e^{iU_2}e^{iU_1}V_{kl}e^{-iU_1}e^{-iU_2})&\\ \times(1+\beta_l)(1+\beta_k).&\end{aligned} \tag{A5}$$

Thus, since $(1+\beta_l)\mathcal{M}_l(1+\beta_l)=0$ when $\mathcal{M}_l$ is an odd operator, no odd operators from V_{kl} survive. This is, of course, not surprising. The projection operators serve the function of eliminating these odd operators, so there is no need to do so within the FWT, itself.

What this means, then, is that to obtain the relativistic corrections to the Bethe sum rule for this no-pair model, the choice of FWT operator U is identical to that which was used to transform the Hamiltonian H in the main part of this paper. Then, since U has no V_{kl} parts, neither will R^j, which is therefore still given by Eqs. (3.18) and (3.19). The transformed Hamiltonians, $\tilde{H}$ and $\tilde{H}(\vec{q})$, will now have additional terms arising from V_{kl} added to the results found above in Eqs. (3.9) and (3.10). Then, consideration of Eq. (2.2) leads to the conclusion that $S_1(Q)$ for the no-pair model will be given by the results found in the main text plus additional terms arising from the V_{kl} terms in $\tilde{H}$ and $\tilde{H}(\vec{q})$. With a fair amount of effort, one can show that these additional terms are proportional to expectation values of the various two-body operators appearing in the Breit-Pauli Hamiltonian (excluding V_{kl} itself), as has been noted elsewhere in this paper.

There is one final, rather technical point that needs to be mentioned to forestall questions that might be raised by those readers familiar with the work of Chraplyvy [33,34] on FWT-type transformations for Hamiltonians including two-body operators. He found that when two interacting particles have equal masses, it is not generally possible to transform the Hamiltonian into an even operator (actually, the desired transformation for us here is to an ''even-even'' form; that is, such that $\tilde{H}$ commutes with the Dirac β operators of both particles). He then introduced more general transformations, which leave the Hamiltonian in a form that he refers to as ''uU separating'' [34], where even-even operators are a subset of the uU-separating ones. At first glance, it would seem that the case discussed in this Appendix, where the two-body terms are treated explicitly, must be treated using Chraply-

vy's methods. However, since the presence of the projection operators in H already serves to eliminate all but the even-even operators in V_{ee}, there is in fact no need to introduce his generalization of the FWT.

APPENDIX B

In this Appendix, we present explicit expressions for some of the quantities used in the calculation of the Bethe sum rule.

For Eq. (3.15), $U_2^j(\vec{q})$ is given as

$$\begin{aligned}\frac{4\epsilon_q^2}{\hbar c}U_2^j(\vec{q}) &= \vec{\alpha}_j\cdot\vec{E}_j - \frac{Q}{q^2\epsilon_q}\vec{q}\cdot\vec{\alpha}_j\vec{q}\cdot\vec{E}_j + a_1\{\vec{p}_j\cdot\vec{\alpha}_j,\vec{q}\cdot\vec{E}_j\}\\ &+\left(a_1-\frac{c^2}{2\epsilon_q^2}\right)\{\vec{q}\cdot\vec{p}_j,\vec{\alpha}_j\cdot\vec{E}_j\}\\ &+\frac{c^4(4\epsilon_q+3mc^2)}{2\epsilon_q^3(\epsilon_q+mc^2)^2}\vec{q}\cdot\vec{\alpha}_j\{\vec{q}\cdot\vec{p}_j,\vec{q}\cdot\vec{E}_j\}\\ &-\frac{Q}{2q^2\epsilon_q}\vec{q}\cdot\vec{\alpha}_j\{\vec{p}_j;\vec{E}_j\},\end{aligned} \tag{B1}$$

where $\{\vec{A};\vec{B}\}=\vec{A}\cdot\vec{B}+\vec{B}\cdot\vec{A}$ and a_1 is given in Eq. (3.12).

For Eq. (2.9), the following expressions are used:

$$\begin{aligned}R_f^{j\dagger}R_V^j+R_V^{j\dagger}R_f^j &= \frac{e\hbar q^2}{4m^2\epsilon_q^3}\vec{\sigma}_j\cdot(\vec{q}\times\vec{E}_j)+\frac{e\hbar c^4}{8\epsilon_q^5}\{\vec{\sigma}_j\cdot(\vec{q}\times\vec{p}_j),\vec{q}\cdot\vec{E}_j\}-\frac{e\hbar^2q^2c^2}{8m^2\epsilon_q^5}(\vec{q}\cdot\nabla_j)\vec{q}\cdot\vec{E}_j\\ &-\frac{e\hbar}{8m^2\epsilon_q^5}\left[\epsilon_q^2-\frac{2m^2c^2}{q^2}(\epsilon_q^2+q^2c^2)+\frac{m^3c^6}{Q}\right]\{\vec{\sigma}_j\cdot(\vec{q}\times\vec{E}_j),\vec{q}\cdot\vec{p}_j\}.\end{aligned} \tag{B2}$$

$$R_f^{j\dagger}[\epsilon_j(\vec{q}),R_V^j]+[R_f^{j\dagger},\epsilon_j(\vec{q})]R_V^j=-\frac{e\hbar^2}{4m^2\epsilon_q^5}(\epsilon_q^3+m^3c^6)(\vec{q}\cdot\vec{\nabla}_j)\vec{q}\cdot\vec{E}_j. \tag{B3}$$

$$\begin{aligned}R_f^{j\dagger}[H_V(\vec{q}),R_f^j]+[R_f^{j\dagger},H_V(\vec{q})]R_f^j &= -\frac{e\hbar^2Q}{8m^2c^2\epsilon_q^2}(Q+3mc^2)\vec{\nabla}_j\cdot\vec{E}_j+\frac{e\hbar^2c^2Q}{8q^2\epsilon_q^5}(\epsilon_q^2+q^2c^2+2\epsilon_q^3/mc^2)(\vec{q}\cdot\vec{\nabla}_j)\vec{q}\cdot\vec{E}_j\\ &-\frac{e\hbar}{2m\epsilon_q}\vec{\sigma}_j\cdot(\vec{q}\times\vec{E}_j)-\frac{e\hbar Q}{4m^2q^2c^2\epsilon_q^2}(Q+3mc^2)\vec{q}\cdot(\vec{E}_j\times\vec{p}_j)\sigma_j\cdot\vec{q}\\ &-\frac{e\hbar Q}{8m^2q^2c^2\epsilon_q^2}(Q+3mc^2)\{\vec{\sigma}_j\cdot(\vec{q}\times\vec{p}_j),\vec{q}\cdot\vec{E}_j\}\\ &-\frac{e\hbar Q}{8m^2q^2c^2\epsilon_q^3}(Q^2-3m^2c^4)\{\vec{\sigma}_j\cdot(\vec{q}\times\vec{E}_j),\vec{q}\cdot\vec{p}_j\}.\end{aligned} \tag{B4}$$

[1] H. Ågren *et al.*, Phys. Rev. A **47**, 3810 (1993).
[2] Z.-C. Yan, J.-M. Zhu, and B.-L. Zhou, Phys. Rev. A **62**, 034501 (2000).
[3] M. Inokuti, Rev. Mod. Phys. **43**, 297 (1971).
[4] H. Bethe, Ann. Phys. (Paris) **5**, 325 (1930).
[5] H. Bethe, in *Handbuch der Physik*, edited by H. Geiger and K. Scheel (Springer, Berlin, 1933), Vol. 24, p. 273.
[6] U. Fano, Annu. Rev. Nucl. Sci. **13**, 1 (1963).
[7] J.S. Levinger, M.L. Rustgi, and K. Okamoto, Phys. Rev. **106**, 1191 (1957).
[8] H.O. Dogliani and W.F. Bailey, J. Quant. Spectrosc. Radiat. Transf. **9**, 1643 (1969).
[9] T. Matsuura and K. Yazaki, Phys. Lett. **46B**, 17 (1973).
[10] J.L. Friar and S. Fallieros, Phys. Rev. C **11**, 274 (1975).
[11] K.-M. Schmitt and H. Arenhövel, Z. Phys. A **320**, 311 (1985).
[12] P.T. Leung, M.L. Rustgi, and S.A.T. Long, Phys. Rev. A **33**, 2827 (1986).
[13] G.A. Aucar, J. Oddershede, and J.R. Sabin, Phys. Rev. A **52**, 1054 (1995).
[14] S.M. Cohen and P.T. Leung, Phys. Rev. A **57**, 4994 (1998).
[15] H. Bichsel, Radiat. Res. **153**, 208 (2000).
[16] J. Ziegler, J. Appl. Phys. **85**, 1249 (1999).
[17] S.M. Cohen, Phys. Rev. A **68**, 012720 (2003).
[18] E. Merzbacher, *Quantum Mechanics*, 2nd ed. (Wiley, New York, 1970).
[19] B. Lo, K. Saxena, and S. Fraga, Theor. Chim. Acta **25**, 97 (1972).

[20] K. Saxena, B. Lo, and S. Fraga, J. Phys. B **5**, 768 (1972).
[21] B. Lo, K. Saxena, and S. Fraga, Theor. Chim. Acta **24**, 300 (1972).
[22] L.L. Foldy and S.A. Wouthuysen, Phys. Rev. **78**, 29 (1950).
[23] J. Sakurai, *Advanced Quantum Mechanics* (Addison-Wesley, Reading, MA, 1982).
[24] Various approximate models could be considered for V_{ee} and our whole procedure involving the FWT can still be carried through [33,34]. As has been mentioned previously, however, these terms should lead to small corrections (see our Appendix A). Therefore, discussion of such issues will not be entered into here. We keep V_{ee} in H for the sake of consistency, since our final results will be found using data from calculations involving eigenstates of the full, many-electron Hamiltonian. This is important, since expectation values using many-electron eigenstates can differ significantly from those using states that do not include the electron-electron interactions. Nonetheless, expectation values of operators arising from V_{ee} through the application of the FWT, such as the second Darwin and orbit-orbit terms, are small and will be dropped throughout the paper.
[25] J. Sucher, Phys. Rev. A **22**, 348 (1980).
[26] M. Douglas and N. Kroll, Ann. Phys. (N.Y.) **82**, 89 (1974).
[27] B.A. Hess, Phys. Rev. A **33**, 3742 (1986).
[28] G. Jansen and B.A. Hess, Phys. Rev. A **39**, 6016 (1989).
[29] C.F. Fischer, At. Data **4**, 301 (1972).
[30] K. Huang *et al.*, At. Data **18**, 243 (1976).
[31] F. Herman and S. Skillman, *Atomic Structure Calculations* (Prentice-Hall, Englewood Cliffs, NJ, 1963).
[32] D.Y. Smith, Phys. Rev. A **35**, 3381 (1987).
[33] Z. Chraplyvy, Phys. Rev. **91**, 388 (1953).
[34] Z. Chraplyvy, Phys. Rev. **92**, 1310 (1953).