Rayleigh-Ritz approximation of the Dirac operator in atomic and molecular physics

I. P. Grant*
Mathematical Institute, 24/29 St. Giles, Oxford OX1 3LB, United Kingdom

H. M. Quiney
School of Chemistry, University of Melbourne, AUS-3052 Parkville, Victoria, Australia
(Received 8 February 2000; published 18 July 2000)

Four-component (spinor) solutions of the Dirac equation may be approximated by *L*-spinor expansions. We discuss their orthogonality and completeness and relate *L*-spinor properties to those of the Coulomb Sturmians. The mathematics of Rayleigh-Ritz approximations for one-electron Schrödinger and Dirac operators provides a rigorous setting for applying finite *L*-spinor matrix approximations to the relativistic hydrogenic atom. Convergence of eigenvalues and eigenvectors with respect to the size of the *L*-spinor set, of expectation values of quantum-mechanical operators, sum rules, and perturbation expansions is examined. The contribution to perturbation sums over states from solutions with eigenvalues in the continuum range (mc^2,∞) (electronic scattering states) and $(-\infty,-mc^2)$ (positronic scattering states) is shown to be essential to get accurate results.

PACS number(s): 31.10.+z, 31.15.Pf, 31.30.Jv

I. INTRODUCTION

L spinors were first mentioned in [1], although they were introduced with a different name in an earlier paper ([2], [Eq. (71)]) (see also [3], p. 240 and [4], Sec. 22.6.3). Although they have played a major role in our development of techniques for solving Dirac-Fock(-Breit) equations for atoms [2,4–6] and molecules [8–11], relativistic correlation studies [2,7] and related problems in quantum electrodynamics [12–15], we have not previously given a full description in print. Recently, Szmytkowski [16] introduced what he called relativistic Coulomb Sturmians, which somewhat resemble *L* spinors, though there are quite profound differences of detail, and his description of our work, based on the limited published material, is misleading. This paper aims to give a definitive account of *L* spinors and their applications to enable readers to evaluate their usefulness. A secondary, but no less important, aim is to show that despite the common belief that it cannot be done, the Rayleigh-Ritz method can be applied to Dirac-Coulomb and related operators as long as the trial solutions are expanded in spinor basis sets that span the correct operator domain.

The Dirac Hamiltonian for a single particle (we only consider electrons and/or positrons in this paper) is

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

where $\boldsymbol{\alpha},\beta$ are the usual 4×4 Dirac matrices, c is the speed of light, $\boldsymbol{p}$ is the three-momentum operator, and $V(\mathbf{r})$ is the potential energy of interaction of the particle with an external electric field, the nuclear Coulomb attraction $V(\mathbf{r})=-Z/r$ in this paper. We shall always use atomic (Hartree) units, so that $m=1$ and $c\approx137$. Dirac solutions therefore are four component spinors, whose properties are defined in standard texts. We shall adhere in this paper to the conventions of [4].

The stationary state solutions in such a spherically symmetric potential for energy E have the form

$$\psi_{E\kappa m}(\mathbf{r}) = \frac{1}{r}\begin{bmatrix} P_{E\kappa}(r)\chi_{\kappa m}(\theta,\varphi) \\ iQ_{E\kappa}(r)\chi_{-\kappa m}(\theta,\varphi)\end{bmatrix}. \tag{2}$$

The two-component spinor $\chi_{\kappa m}(\theta,\varphi)$ is a coupled spin-orbit function (see, for example, Ref. [4], Eq. (22.92) for a full description). The angular quantum number κ is related to the total angular-momentum quantum number j by $|\kappa|=j+\frac{1}{2}$, with $j=\frac{1}{2},\frac{3}{2},\ldots$ and the quantum number m, representing the component of total angular momentum on the quantization axis Oz, can take values in the range $-j,-j+1,\ldots,j-1,j$. The spinor $\chi_{\kappa m}(\theta,\varphi)$ is associated with an orbital angular-momentum quantum number l having $l=j+1/2$ if κ is positive and $l=j-1/2$ if κ is negative. These functions form a basis for the irreducible representation $\mathcal{D}^j$ of SO(3) for both signs of κ and are the fundamental building blocks for constructing atomic and molecular wave functions. A major difference from Schrödinger wave functions is the coupling of spatial and spin degrees of freedom, expressed by the four components of the Dirac spinors. The coupling of these components by the Dirac operator determines the asymptotic behavior of the wave functions as $r\rightarrow0$, the region in which the dynamic effects of relativity largely originate. An understanding of the structure and symmetry properties of Dirac spinors is therefore essential for the effective use of relativistic wave functions [4].

Sturmian functions have long been advocated in nonrelativistic quantum mechanics because they allow wave functions, for bound states or for scattering, to be expanded in a complete countable set of eigenfunctions of a convenient form [17,18] and the range of applications is now quite wide. We therefore start with a brief account of the properties of Coulomb Sturmians, introducing ideas and relations that serve in part to motivate the definition of *L* spinors that follows. We connect the orthogonality and completeness properties of *L* spinors to those of Coulomb Sturmians by exam-

*Electronic address: ipg@maths.ox.ac.uk

ining the nonrelativistic limiting behavior as $c\to\infty$. The emphasis in this paper will be on using L spinors in the context of the Rayleigh-Ritz procedure for Dirac Hamiltonians. As many have claimed that this approach is doomed to failure, we next give an account of the relevant mathematics that is sufficiently rigorous to support this work and related applications in relativistic quantum mechanics and quantum electrodynamics. Some of the mathematical background was previously summarized in [3,4]; the present paper covers aspects not previously discussed.

The theory is illustrated with several applications, starting with electronic states of the hydrogenic atom for which exact analytic results are available. The large and small components are each expanded in matched L-spinor subsets of the same dimension, using the radial amplitudes $\{f^L_{n\kappa}(r)\}^N_{n=1}$ and $\{f^S_{n\kappa}(r)\}^N_{n=1}$, respectively, and we examine the convergence of bound-state solutions for different values of the symmetry quantum number κ as N increases, and examine the influence of the tuning parameter λ. The completeness of the L-spinor expansions is important for calculating sum-over-states expressions in perturbation theory. We illustrate this by way of calculations on the static polarizability of hydrogenic atoms, and through the very instructive model in which the nuclear charge of a hydrogenic atom is changed from Z to $Z+Z'$, which we studied earlier [19,20]. Here, we know the exact answer, and the calculations provide strong evidence of the completeness of L-spinor expansions provided the negative-energy states are included. More importantly, it also shows that the partition between positive- and negative-energy states depends on the choice of potential, exposing the major flaw in proposals to eliminate the influence of negative-energy states with projection operators.

II. THE COULOMB STURMIAN FUNCTIONS

A. Definition and properties of Coulomb Sturmians

The nonrelativistic Sturmian functions are defined by [17,18] as the normalized solutions of the differential equation

$$\left[-\frac{d^2}{dr^2}+\frac{l(l+1)}{r^2}-2E_0+2\alpha_{nl}V(r)\right]S_{nl}(r)=0,\quad 0<r<\infty \tag{3}$$

vanishing at the end points $r=0$ and $r=\infty$. The integers $n=1,2,3,\ldots$ and $l=0,1,\ldots,n-1$ correspond to the usual nonrelativistic quantum numbers and E_0 is a fixed, negative number. The parameter α_{nl} must be adjusted to ensure that the boundary conditions are satisfied. The functions are orthonormal with respect to the weight function $V(r)$ (which must be strictly of one sign, usually negative) so that

$$\int_0^\infty S_{nl}(r)S_{n'l}(r)V(r)dr=-\delta_{nn'}. \tag{4}$$

The most important case is that in which $V(r)$ is a Coulomb potential

$$V(r)=-\frac{Z}{r},\quad 0<r<\infty.$$

We set $E_0=-\lambda^2/2$, and rewrite Eq. (3) in terms of the independent variable $x=2\lambda r$, so that

$$\left[-\frac{d^2}{dx^2}+\frac{l(l+1)}{x^2}+\frac{1}{4}-\frac{\alpha_{nl}Z}{\lambda x}\right]S_{nl}(x)=0, \tag{5}$$

having the solutions

$$S_{nl}(x):=\mathcal{N}_{nl}s_{nl}(x),$$
$$s_{nl}(x)=e^{-x/2}x^{l+1}L^{2l+1}_{n-l-1}(x), \tag{6}$$
$$n=l+1,l+2,\ldots,$$

which vanish at $x=0$ and $x=\infty$ provided

$$n=\alpha_{nl}Z/\lambda.$$

The $L^\alpha_k(x)$ are Laguerre polynomials [21] and $\mathcal{N}_{nl}$ is a normalization constant. We recover the standard solutions for the radial hydrogenic eigenfunctions when $\alpha_{nl}=1$, giving $E_0=-Z^2/2n^2$. Notice that λ (and therefore E_0) is fixed for the Coulomb Sturmians whereas $\lambda=\sqrt{-2E_0}$ depends upon n for the Schrödinger eigenfunctions.

The properties of the orthogonal polynomials $L^\alpha_k(x)$ are listed in many compilations such as [21]. When $\alpha\geqslant 1$ they are orthogonal on $(0,\infty)$ with weight function $w(x)=e^{-x}x^\alpha$, such that

$$\int_0^\infty e^{-x}x^\alpha L^\alpha_k(x)L^\alpha_{k'}(x)dx=\frac{\Gamma(\alpha+k+1)}{k!}\delta_{k,k'}. \tag{7}$$

The generating function

$$\Phi^{(\alpha)}(t,s):=L^\alpha_k(s)t^k=(1-t)^{-\alpha-1}\exp\left(\frac{ts}{t-1}\right),\quad |t|<1 \tag{8}$$

can be used to write down explicit representations of the polynomials. It also provides an economical means of evaluating integrals of the form

$$\langle kl|x^p|k'l\rangle=\int_0^\infty s_{kl}(x)x^p s_{k'l}(x)dx$$

for integer values of p for which this integral exists by identifying the coefficient of $t^k u^{k'}$ in the integral

$$I^{(p)}_l(t,u)=\int_0^\infty e^{-x}x^{2l+2+p}\Phi^{(2l+1)}(t,x)\Phi^{(2l+1)}(u,x)dx$$
$$=\frac{(2l+2+p)!}{(1-tu)^{2l+3+p}}[(1-t)(1-u)]^{1+p}. \tag{9}$$

Two cases have immediate application

Case A. $p=-1$: then

$$I_l^{(-1)}(t,u)=\frac{(2l+1)!}{(1-tu)^{2l+2}}=\sum_{k=0}^{\infty}\frac{(2l+k+1)!}{k!}(tu)^k$$

from which we obtain

$$\langle kl|x^{-1}|k'l\rangle=\frac{(2l+k+1)!}{k!}\delta_{k,k'}, \tag{10}$$

which agrees with Eq. (7) if we put $\alpha=2l+1$ and $k=n-l-1$.

Case B. $p=0$: This gives the Gram (overlap) matrix G of Coulomb Sturmians. In this case

$$I_l^{(0)}(t,u)=\frac{(2l+2)!}{(1-tu)^{2l+3}}[(1-t)(1-u)]$$

so that there are nonvanishing matrix elements for $k'=k\pm1$ as well as for $k'=k$. We shall normalize the Sturmians so that

$$\langle kl|k'l\rangle=\delta_{k,k'}, \tag{11}$$

so that, remembering $k=n-l-1$, Eq. (10) gives

$$\mathcal{N}_{nl}=\left[\frac{(n-l-1)!}{2n(n+l)!}\right]^{1/2}. \tag{12}$$

The nonvanishing elements of the Gram matrix, are thus

$$g_{nn}^l=1,$$

$$g_{n,n+1}^l=g_{n+1,n}^l=-\frac{1}{2}\sqrt{1-\frac{l(l+1)}{n(n+1)}},\quad n=l+1,l+2,\ldots. \tag{13}$$

B. Completeness and linear independence

The classical proof that the orthonormal system $(2/x)^{1/2}S_{nl}(x)_{n=l+1}^{\infty}$ is complete in $L^2(\boldsymbol{R}_+)$ is given by ([22], p. 170 and [23], p. 95). It follows that if we define the functions on $\boldsymbol{R}^3\to\boldsymbol{R}$ by

$$\phi_{nlm}(\boldsymbol{r})=\frac{S_{nl}(r)}{r}Y_{lm}(\theta,\varphi),\quad n\geqslant l+1,\quad |m|\leqslant l,$$

where (r,θ,φ) are spherical polar coordinates of the position $\mathbf{r}$ and $Y_{lm}(\theta,\varphi)$ are spherical harmonics, then the set $\{(2/r)^{1/2}\phi_{nlm}(\boldsymbol{r})\}$ is a complete orthonormal system in $L^2(\boldsymbol{R}^3)$ ([24], Lemma 6, p. 31).

Although the weighted Sturmians are therefore a complete orthonormal set in $L^2(\boldsymbol{R}^3)$, it is often more convenient to use the unweighted Sturmians, so that the Gram matrix of the set $\{\phi_{nlm}(\boldsymbol{r})\}$ is tridiagonal in each infinite lm subset. The completeness and linear independence of the Sturmians has to be reexamined in this context.

In a finite set of dimension N, positivity of the Gram determinant, $\det G^{(N)}>0$, is sufficient to ensure that a set is linearly independent. This is equivalent to the statement that for every finite set $\{\chi_n\}_{n=1}^N$, there exists a *biorthonormal set* (BOS) $\{\chi_n^*\}_{n=1}^N$ such that, for all m,n, $(\chi_m^*,\chi_n)=\delta_{mn}$, where $(,)$ is the appropriate inner product. For any Ψ in the span of the set $\{\chi_n\}_{n=1}^N$, then we can decompose Ψ in the form

$$\Psi=\sum_{m=1}^{N}(\Psi,\chi_m^*)\chi_m,$$

so that in particular

$$\chi_n^*=\sum_{m=1}^{N}(\chi_n^*,\chi_m^*)\chi_m=\sum_{m=1}^{N}G_{mn}^{(N)*}\chi_m,$$

where $G^{(N)*}$ is the Gram matrix of the set $\{\chi_n^*\}_{n=1}^N$. Similarly

$$\chi_n=\sum_{m=1}^{N}(\chi_m,\chi_n)\chi_m^*=\sum_{n=1}^{N}G_{mn}\chi_m^*,$$

where $G^{(N)}$ is the Gram matrix of the set $\{\chi_n\}_{n=1}^N$. Clearly this construction ensures that each element χ_n^* is in the span of $\{\chi_n\}_{n=1}^N$ and each element χ_m is in the span of $\{\chi_m^*\}_{m=1}^N$, and

$$G^{(N)*}G^{(N)}=G^{(N)}G^{(N)*}=I^{(N)},$$

where $I^{(N)}$ is the N-dimensional identity matrix.

Things are not quite so simple for infinite sets, and there are several inequivalent criteria for linear independence in Hilbert spaces. For our purposes, the *minimality* property is the most useful: an infinite set of functions is said to be *minimal* if no single element χ_k can be approximated with arbitrary precision in the linear span of the remaining elements of the set. A minimal set has a BOS, so that the expansion coefficients c_k^n of an element $P_n\Psi=\Psi^{(n)}$ in the span of the finite subset

$$\{\chi_1,\ldots,\chi_n\}$$

can be determined from the equation

$$c_k^n=(\chi_k^*,\Psi^{(n)})=\left(\chi_k^*,\sum_{i=1}^{n}c_i^n\chi_i\right),\quad k=1,2,\ldots,n,$$

independently of the way in which c_k^n is calculated. Moreover, the expansion coefficients converge to the expansion coefficients of Ψ itself as $n\to\infty$.

In this setting, it is easy to see that the Coulomb Sturmians have BOS with elements $S_{nl}^*(x)=(2/x)S_{nl}(x)$ and $\phi_{nlm}^*(x)=[2S_{nl}(x)/x^2]Y_{lm}(\theta,\varphi)$, respectively. Thus every element $\Psi\in L^2(\boldsymbol{R}^3)$ has a formal expansion in terms of this set. The set $\{\phi_{nlm}(x)\}$ is also minimal and complete on the Sobolev spaces $W_2^{(p)}(\boldsymbol{R}^3)$ for $p=1,2$, which proves to be exactly what is needed for the convergence of the Rayleigh-Ritz method in atomic and molecular problems.

Gerschgorin's circle theorem [25,26] allows us to verify that any finite subset of the Coulomb Sturmians has a positive definite Gram matrix when $l>0$. Since each eigenvalue

μ of the tridiagonal matrix $G^{(N)}$ is located in the union of intervals $1-\rho_n \leqslant \mu \leqslant 1+\rho_n$, where

$$\rho_n = |g_{n,n-1}| + |g_{n,n+1}|, \quad n = l+1, l+2, \ldots, l+N$$

so that

$$\rho_{l+1} = |g_{l+1,l+2}| = \frac{1}{\sqrt{2l+4}}$$

and

$$\rho_n \sim 1 - \frac{l(l+1)}{n^2} + O(n^{-3}), \quad n \gg l+1.$$

Thus all eigenvalues lie in the interior of the interval (0,2) so that every finite subset of the Coulomb Sturmians has a positive definite Gram matrix. The condition number of the matrix $G^{(N)}$ is $k_N = (1+\rho_N)/(1-\rho_N) \sim 2N^2/l(l+1)$ when N is large, so that the system is very well conditioned and there is little danger of linear dependence problems in practice.

This estimate fails when $l=0$, as $g_{n,n+1} = -1/2$ independent of the value of n. However, since $G^{(N)}$ is a tridiagonal matrix with diagonal elements 1 and subdiagonal and superdiagonal elements $-\frac{1}{2}$, it is easy to show that $\det G^{(N)}$ never vanishes, so that we still have linear independence.

III. DIRAC *L* SPINORS

The Dirac L spinors are defined as relativistic analogs of the Coulomb Sturmians. We envisage representing Dirac four-component wave functions as linear combinations

$$\begin{bmatrix} \frac{1}{r}\sum_{n_r} c^L_{n_r} f^L_{n_r\kappa}(r)\chi_{\kappa m}(\theta,\varphi) \\ i\frac{1}{r}\sum_{n_r} c^S_{n_r} f^S_{n_r\kappa}(r)\chi_{-\kappa m}(\theta,\varphi) \end{bmatrix} \tag{14}$$

of a form similar to that of Eq. (2).

In terms of the independent variable $x = 2\lambda r$, $\lambda > 0$ constant, the L-spinor amplitudes are given by the formulas ([4] Eqs. (22.146) and (22.147))

$$f^L_{n_r\kappa}(x) = \mathcal{N}_{n_r\kappa} x^\gamma e^{-x/2}\left\{ -(1-\delta_{n_r,0}) L^{2\gamma}_{n_r-1}(x) + \frac{N_{n_r\kappa}-\kappa}{n_r+2\gamma} L^{2\gamma}_{n_r}(x)\right\}, \tag{15}$$

$$f^S_{n_r\kappa}(x) = \mathcal{N}_{n_r\kappa} x^\gamma e^{-x/2}\left\{ -(1-\delta_{n_r,0}) L^{2\gamma}_{n_r-1}(x) - \frac{N_{n_r\kappa}-\kappa}{n_r+2\gamma} L^{2\gamma}_{n_r}(x)\right\}. \tag{16}$$

The labels L,S identify the ''large'' and ''small'' components of the Dirac spinor in a conventional way, n_r is a non-negative integer, and

$$\gamma = +\sqrt{\kappa^2 - Z^2/c^2}, \quad N_{n_r\kappa} = +\sqrt{n_r^2 + 2n_r\gamma + \kappa^2}, \tag{17}$$

are, respectively, the leading exponent of the power series expansion of the functions about $x=0$ and the *apparent principal quantum number*. The L spinors are solutions of the differential equation system[1]

$$\begin{bmatrix} \frac{1}{2} - \frac{\alpha_{n_r\kappa} Z\mu^2}{c}\frac{1}{x} & -\frac{d}{dx} + \frac{\kappa}{x} \\ \frac{d}{dx} + \frac{\kappa}{x} & -\frac{1}{2} - \frac{Z}{\alpha_{n_r\kappa}\mu^2 c}\frac{1}{x} \end{bmatrix} \begin{bmatrix} \mu^{-1} f^L_{n_r\kappa}(x) \\ \mu f^S_{n_r\kappa}(x) \end{bmatrix} = 0, \tag{18}$$

where c is the speed of light ($c \approx 137$ in atomic units), and μ^2 is a root of the equation

$$\mu^4 - \frac{2c}{\lambda}\mu^2 + 1 = 0. \tag{19}$$

We choose

$$\mu^2 = \frac{c}{\lambda}\left(1 + \sqrt{1 - \frac{\lambda^2}{c^2}}\right), \quad \mu^{-2} = \frac{c}{\lambda}\left(1 - \sqrt{1 - \frac{\lambda^2}{c^2}}\right),$$

which ensures that $f^L_{n_r\kappa}(x) \to \text{const}\, S_{n_r l}(x)$ in the nonrelativistic limit $c \to \infty$ which we study below.

The analog of the nonrelativistic energy parameter $E_0 = -\lambda^2/2$ is given by

$$E^R_0 = c^2\sqrt{1 - \lambda^2/c^2} = c^2 + E_0 + O(1/c^2). \tag{20}$$

The boundary conditions as $r \to 0$ and $r \to \infty$ are satisfied when

$$\alpha_{n_r\kappa} = N_{n_r\kappa}\lambda/Z$$

and the L-spinor amplitudes coincide with Dirac-Coulomb eigenfunctions $P_{n\kappa}(r)$ and $Q_{n\kappa}(r)$ having principal quantum number $n = n_r + |\kappa|$ when $\alpha_{n_r\kappa} = 1$.

A. *L* spinors in the nonrelativistic limit

It is essential that our construction should coincide with the nonrelativistic Coulomb Sturmians in the limit $c \to \infty$, corresponding to instantaneous propagation of electromagnetic disturbances, and this can be done directly from the definitions (15) and (16). It is easy to see that $\gamma \to |\kappa|$, $N_{n_r\kappa} \to n$, so that for negative values of $\kappa = -l-1$, we have

[1]The brief summary of L spinors in ([4], Sec. (22.6.3)) is correct save for Eq. (22.145), which should be replaced by the present Eq. (18).

$$f^L(x)\to \text{const}\times x^{l+1}e^{-x/2}\{-(1-\delta_{n_r0})L^{2l+2}_{n_r-1}(x)+L^{2l+2}_{n_r}(x)\}$$
$$=\text{const}\times x^{l+1}e^{-x/2}L^{2l+1}_{n_r}(x) \tag{21}$$

using ([21], Eq. (22.7.30)) in the second line. Similarly for positive $\kappa=l$, remembering that $n_r\geqslant 1$ in this case, we have

$$f^L(x)\to \text{const}\times x^{l}e^{-x/2}\left\{-(1-\delta_{n_r0})L^{2l}_{n_r-1}(x)+\frac{n_r}{n_r+2l}L^{2l}_{n_r}(x)\right\}$$
$$=\text{const}\times x^{l}e^{-x/2}\{-(n_r+2l)L^{2l}_{n_r-1}(x)+n_rL^{2l}_{n_r}(x)\}$$
$$=\text{const}\times x^{l+1}e^{-x/2}L^{2l+1}_{n_r}(x), \tag{22}$$

using ([19] Eq. (22.7.31)) in the third line. Since $n_r=n-l-1$, we have proved that

$$f^L(x)\to \text{const}\times S_{nl}(x)$$

for both signs of κ in the nonrelativistic limit.

A similar analysis shows that for both signs of κ,

$$f^S(x)\to \text{const}\left(\frac{d}{dx}+\frac{\kappa}{x}\right)x^{l+1}e^{-x/2}L^{2l+1}_{n_r}(x),$$

which may be expressed in terms of the criterion of *strict kinetic balance*

$$\lim_{c\to\infty} f^S(x)=\text{const}\times\left(\frac{d}{dx}+\frac{\kappa}{x}\right)f^L(x).$$

The kinetic balance criterion simply states that the large and small component amplitudes should be related according to the Pauli approximation, guaranteeing consistency with nonrelativistic equations in the limit $c\to\infty$ [27].

B. Orthogonality properties

The standard orthogonality properties of Laguerre polynomials can be used to write down L-spinor generalizations of Sturmian properties. However, an orthogonality relation with respect to the weight function $1/x$ can be written down in an elementary way from Eq. (18). Multiplying from the left by the adjoint vector $[\mu^{-1}f^L_{n_r'\kappa},\mu f^S_{n_r'\kappa}]$ and subtracting the result from the corresponding equation with n_r and n_r' interchanged gives

$$(\alpha_{n_r\kappa}-\alpha_{n_r'\kappa})\int_0^\infty\{f^L_{n_r'\kappa}(x)f^L_{n_r\kappa}(x)-(\alpha_{n_r\kappa}\alpha_{n_r'\kappa})^{-1}f^S_{n_r'\kappa}(x)f^S_{n_r\kappa}(x)\}\frac{dx}{x}=0.$$

Thus the integral vanishes if the eigenvalues $\alpha_{n_r\kappa}$ and $\alpha_{n_r'\kappa}$ are different. Although this reduces to the Sturmian orthogonality relation (7) in the nonrelativistic limit, the integrand is not obviously positive definite and the result is not very useful.

However, it is easy to use the elementary results of the previous section to write down the elements of the Gram matrix. The normalization factor $\mathcal{N}_{n_r\kappa}$ is the same for both $f^L_{n_r\kappa}(x)$ and $f^S_{n_r\kappa}(x)$, and is given by the equation

$$1=g^{(\kappa)}_{n_rn_r}$$
$$=\mathcal{N}^2_{n_r\kappa}\left\{(1-\delta_{n_r0})\frac{\Gamma(2\gamma+n_r)}{(n_r-1)!}+\left(\frac{N_{n_r\kappa}-\kappa}{2\gamma+n_r}\right)^2\frac{\Gamma(2\gamma+n_r+1)}{(n_r)!}\right\}$$
$$=\mathcal{N}^2_{n_r\kappa}2N_{n_r\kappa}(N_{n_r\kappa}-\kappa)\frac{\Gamma(2\gamma+n_r)}{n_r!(2\gamma+n_r)},$$

so that

$$\mathcal{N}_{n_r\kappa}=\left[\frac{n_r!(2\gamma+n_r)}{2N_{n_r\kappa}(N_{n_r\kappa}-\kappa)\Gamma(2\gamma+n_r)}\right]^{1/2}. \tag{23}$$

In a similar fashion we can easily show that

$$g^{\kappa}_{n_r,n_r}=1,$$
$$g^{(\kappa)}_{n_r,(n_r+1)}=g^{(\kappa)}_{(n_r+1),n_r}=\frac{\eta^T}{2}\left[\frac{(n_r+1)(2\gamma+n_r+1)(N_{n_r\kappa}-\kappa)}{N_{n_r\kappa}N_{(n_r+1),\kappa}(N_{(n_r+1),\kappa}-\kappa)}\right]^{1/2}, \tag{24}$$
$$T=L,S$$

where $\eta^L=-1$ and $\eta^S=+1$.

It is straightforward to show that the L-spinor Gram matrices reduce to the Gram matrices for Coulomb Sturmians (apart from the sign of the off-diagonal elements) in the nonrelativistic limit. Writing $g^{(N)}=G^{(N)}-I^{(N)}$, we see, by expanding with respect to the last row, that $f^{(N)}(\sigma)=\det(g^{(N)}-\sigma I^{(N)})$ satisfies

$$f^{(N)}(\sigma)=-\sigma f^{(N-1)}(\sigma)-g^2_{N,N-1}f^{(N-2)}(\sigma)$$

with $f^{(1)}(\sigma)=-\sigma$ and $f^{(2)}(\sigma)=\sigma^2-g^2_{12}$. We conclude inductively that $f^{(2k)}(\sigma)$ and $f^{(2k+1)}(\sigma)/\sigma$ are polynomials in σ^2 of degree k, so that the eigenvalues of $G^{(N)}$ are in the interval $(1-\rho_N,1+\rho_N)$, where

$$\rho_N=1-\frac{C}{N^2}+O(N^{-3}), \tag{25}$$

where C is a positive constant. The eigenvalues of $f^{(N)}$ are distributed symmetrically about $\sigma=0$ when N is even, and there is an additional zero eigenvalue when N is odd. Thus $G^{(N)}$ has condition number $k_N=(1+\rho_N)/(1-\rho_N)\sim 2N^2/C$

when N is large, so that the linear independence behavior is very similar to the Coulomb Sturmians.

C. Completeness of *L* spinors

We can establish completeness of L spinors in a variety of Hilbert spaces by exploiting the following ([24], Lemma 5).

Lemma 1 (Klahn). Let $\{\varphi_n\}_{n=1}^{\infty}$ be a complete system in a Hilbert space H. Moreover, let $a_{n\mu}$, $(1\leqslant\mu\leqslant n)$ be arbitrary complex numbers with $a_{nn}\neq 0$. Then the system

$$\left\{\psi_n=\sum_{\mu=1}^{n} a_{n\mu}\varphi_\mu\right\}_{n=1}^{\infty}$$

is also complete in H.

To apply this to the L spinors, we note that Eqs. (15) and (16) can be written

$$f^T_{n_r\kappa}(x)=a_{n_r,n_r-1}x^\gamma e^{-x/2}L^{2\gamma}_{n_r-1}(x) + \eta^T a_{n_r,n_r}x^\gamma e^{-x/2}L^{2\gamma}_{n_r}(x),\quad T=L,S$$

with $\eta^L=+1,\eta^S=-1$. Since $a_{n_r,n_r-1}=-(1-\delta_{n_r,0})$, only the second term contributes when $n_r=0$ for both signs of κ. Also since $N_{0,\kappa}=|\kappa|$, the first nonvanishing L spinor for $\kappa>0$ has $n_r=1$. We infer that the radial amplitudes appearing in Eq. (14) have formal L-spinor expansions and that the L spinors are both complete and minimal on the Sobolev spaces $[W_2^{(p)}(\boldsymbol{R}^3)]^2$ for $p=1,2$. We shall see that this is exactly what we need for constructing trial wave functions for the Rayleigh-Ritz approximation of Dirac four-component wave functions.

D. Charge conjugation and *L* spinors

One of the most important symmetries of the Dirac equation is charge conjugation which, loosely speaking, sets up a correspondence between electron and positron states. Under charge conjugation, Dirac four-spinors transform like

$$\psi\rightarrow\psi_c=C\overline{\psi}^t \tag{26}$$

where the superscript t denotes transposition and $\overline{\psi}=\psi^*\gamma^0$ is Dirac conjugation. The matrix C is given by

$$C=i\gamma^2\gamma^0=\begin{bmatrix}0 & -i\sigma^2\\ -i\sigma^2 & 0\end{bmatrix}.$$

When the radial amplitudes $P(r),Q(r)$ are real, it is easy to show that if

$$\psi=\frac{1}{r}\begin{bmatrix}P(r)\chi_{\kappa,m}\\ iQ(r)\chi_{-\kappa,m}\end{bmatrix}$$

then

$$\psi_c=-i(-1)^{m+1/2}\begin{bmatrix}Q(r)\chi_{-\kappa,-m}\\ iP(r)\chi_{\kappa,-m}\end{bmatrix}.$$

Under this transformation, expectation values of the position variable and the charge-current vector remain invariant, whilst those of spin, orbital, and total angular momentum change sign, as does the sign of the energy parameter E and the sign of Z coupling the electron to the external Coulomb potential.

By making the corresponding changes

$$Z\leftrightarrow -Z,\quad f^L_{n_r\kappa}(x)\leftrightarrow f^S_{n_r\kappa}(x),\quad \kappa\leftrightarrow-\kappa,\quad \mu\leftrightarrow\mu^{-1}$$

in Eq. (18) we see that L spinors retain the charge conjugation symmetries of the Dirac eigenfunctions on which they are modeled. Since the mapping $\mu\leftrightarrow\mu^{-1}$ is equivalent to changing the sign of the energy parameter $E_0^R=+\sqrt{1-\lambda^2/c^2}$, Eq. (20), we infer that L-spinor expansions will be able correctly to represent positron (negative-energy electron) states as well as bound states. This assertion will be clarified in the discussion of L-spinor applications that follows.

IV. THE RAYLEIGH-RITZ METHOD FOR SCHRÖDINGER AND DIRAC OPERATORS

In view of the misunderstanding of the status of the Rayleigh-Ritz methods for Dirac problems in atomic and molecular physics, it is useful to begin with a review of the known mathematical results and to see how they can be applied to both nonrelativistic and relativistic cases.

A. The Rayleigh quotient

Applications of the Rayleigh-Ritz method in quantum mechanics usually assume that one is dealing with a *self-adjoint, non-negative, compact* operator T, defined on a domain $\mathcal{D}$ contained in a Hilbert space $\mathcal{H}$. Such an operator has an ordered set of non-negative eigenvalues

$$0\leqslant\lambda_1\leqslant\lambda_2\leqslant\cdots$$

with due regard for multiplicity, with corresponding eigenfunctions $\psi_1,\psi_2,\ldots$, respectively. The *Rayleigh quotient*, defined (in Dirac bra-ket notation) by

$$R[\psi]:=\langle\psi|T|\psi\rangle/\langle\psi|\psi\rangle, \tag{27}$$

evidently exists for all nontrivial $\psi\in\mathcal{D}$. The Rayleigh-Ritz method assumes that we can approximate every such ψ by its projection, $P_n\psi$, on a finite dimensional subspace $\mathcal{W}_n:=\{\chi_1,\chi_2,\ldots,\chi_n\}\subset\mathcal{D}$, so that there exists some set of coefficients (in general complex) for which

$$P_n\psi=c_1^n\chi_1+c_2^n\chi_2+\cdots+c_n^n\chi_n. \tag{28}$$

A simple calculation shows that

$$F^n[\mathbf{c^n}]:=R[P_n\psi]=\mathbf{c}^{n\dagger}\mathbf{T}^n\mathbf{c}^n/\mathbf{c}^{n\dagger}\mathbf{S}^n\mathbf{c}^n, \tag{29}$$

where $\mathbf{c}^n=(c_1^n,c_2^n,\ldots,c_n^n)^t$, $\mathbf{c}^{n\dagger}$ is its Hermitian conjugate, and $\mathbf{T}^n,\mathbf{S}^n$ are $n\times n$ Hermitian matrices with elements

$$T^n_{ij}=\langle\chi_i|T|\chi_j\rangle,\quad S^n_{ij}=\langle\chi_i|\chi_j\rangle,\quad 1\leqslant i,j\leqslant n \tag{30}$$

respectively. The Gram matrix $\mathbf{S}^n$ must be positive definite if Eq. (29) is to have a meaning, so that the set $\mathcal{W}_n$ must be linearly independent. We shall henceforth require that this be true for every set of functions meriting consideration. The Gram matrix is the n-dimensional identity matrix if the functions χ_i are orthonormal, but we have already seen that Coulomb Sturmians and L spinors are not of this type.

This manipulation approximates the Rayleigh quotient by an algebraic function $F^n[\mathbf{c^n}]$. Elementary calculus tells us how to find its stationary points as the roots of the Galerkin equation ([28], p. 395)

$$\mathbf{T}^n\mathbf{c}^n=\Lambda^n\mathbf{S}^n\mathbf{c}^n, \tag{31}$$

where Λ^n is a diagonal eigenvalue matrix whose elements can be ordered so that

$$0<\Lambda_1^n\leqslant\Lambda_2^n\leqslant\cdots\leqslant\Lambda_n^n,$$

with due regard for multiplicity. We denote the corresponding eigenvectors, the columns of $\mathbf{c}^n$, by $\mathbf{c}_1^n,\mathbf{c}_2^n,\ldots,\mathbf{c}_n^n$.

It is customary in the quantum-mechanical literature to show at this point that the lowest eigenvalue Λ_1^n is an upper bound to the true eigenvalue λ_1^n. This is attributed to the fact that T is a positive definite operator, and the conclusion is easily generalized to encompass all operators T which are bounded below. However, this needs to be seen in a more general setting given by the following theorem ([28], pp. 397–398).[2]

Theorem 1 (Poincaré). Let T be as defined above. Then the eigenvalues of the Galerkin equation (31) are upper bounds to the target eigenvalues

$$0<\lambda_1\leqslant\Lambda_1^n\leqslant\cdots\leqslant\lambda_n\leqslant\Lambda_n^n.$$

Stakgold [28] lists a number of comments, in particular the following.

(1) If T is *nonpositive* then $-T$ is non-negative, and all the above holds with the inequalities reversed.

(2) If T is *indefinite* then the inequalities for non-negative T hold for the upper part of the spectrum whilst those for nonpositive T hold for the lower eigenvalues.

(3) Increasing the size of the set $\mathcal{W}_n$ generally improves the approximation. In practice, one would like to use a complete set of functions $\mathcal{W}\coloneqq\{\chi_i\}_{i=1}^{\infty}$ spanning the domain $\mathcal{D}$, so that $\mathcal{W}_n$ spans an increasing subspace of $\mathcal{D}$ as n increases. Then $\Lambda_i^n\rightarrow\lambda_i$ as $n\rightarrow\infty$ because the compact operator $P_nT\rightarrow T$ in the operator norm.

B. Convergence of Rayleigh-Ritz approximations

We need rather more precise criteria of convergence if we are to rely on the Rayleigh-Ritz approximations on finite basis sets. Not only do we need to have some idea how close our estimates of eigenvalues are to the exact values, we should like to know in what sense the approximate eigenfunction $P_n\psi$ approximates the true eigenfunction ψ and whether approximate expectation values and transition matrix elements constructed from approximate wave functions are sufficiently close to the true values to be of practical use.

The convergence of Rayleigh-Ritz eigenvalues and eigenfunctions of nonrelativistic quantum Hamiltonian operators was analyzed in [24]. Let $T=\boldsymbol{p}^2/2$ be the usual particle kinetic energy, and consider the Sobolev space $W_2^{(1)}\subset L^2$ equipped with the norm

$$\|\psi\|_{W_2^{(1)}}^2=\langle u|1+T|u\rangle=\|(1+T)^{1/2}u\|_{L^2}^2. \tag{32}$$

Klahn and Bingel [22] establish that the Rayleigh-Ritz approximate eigenvalues converge to the eigenvalues of the target Hamiltonian (which they term *E convergence*) if the set $\mathcal{W}$ is complete in $W_2^{(1)}$. Moreover, the approximation $\psi_i^n\coloneqq P_n\psi_i$ to the ith bound state converges in the mean to the eigenfunction ψ_i whenever the corresponding eigenvalue, E_i^n, converges to E_i as n increases. Thus we have only to construct a basis set that is complete in $W_2^{(1)}$ to be certain of E convergence both to eigenvalues and eigenfunctions of the Hamiltonian.

This analysis has been extended by Klahn and Morgan [29] to the convergence of expectation values and transition matrix elements. Let A be a strictly positive operator, self-adjoint on a domain $\mathcal{D}(A)$, and define a new A norm by

$$\|\psi\|_A^2=\langle\psi|A|\psi\rangle=\|A^{1/2}\psi\|_{L^2}^2. \tag{33}$$

With the related scalar product, this induces a new Hilbert space with the A norm. We say that a set of functions is A-complete if it is complete in this space. The Sobolev space $W_2^{(1)}$ is a special case of this construction having $A=1+T$.

Lemma 2. The sequence

$$\langle A\rangle^n\coloneqq\langle\psi^n|A|\psi^n\rangle$$

converges to $\langle A\rangle=\langle\psi|A|\psi\rangle$ if and only if $\psi^n\rightarrow\psi$ as $n\rightarrow\infty$ in the A norm.

This is a direct consequence of two inequalities

$$|\langle A\rangle^n-\langle A\rangle|\leqslant\|\psi^n-\psi\|_A^2+2\langle A\rangle^{1/2}\|\psi^n-\psi\|_A$$

and

$$\|\psi^n-\psi\|_A^2\leqslant|\langle A\rangle^n-\langle A\rangle|+2\|A\psi\|\,\|\psi^n-\psi\|.$$

Thus the set $\mathcal{W}$ must be A complete if a sequence of eigenfunctions $\{\psi^n\}$ generated by the Rayleigh-Ritz method is also to give a convergent sequence of approximations $\langle A\rangle^n$.

We can avoid having to deal with A-completeness if the operator A is *relatively form bounded* by T: that is, there exists a pair of non-negative numbers a,b such that

$$|\langle\psi|A|\psi\rangle|<a\langle\psi|\psi\rangle+b\langle\psi|T|\psi\rangle,\quad\forall\psi\in\mathcal{D}(T). \tag{34}$$

This includes a wide range of operators: *bounded* operators, for which we can set $b=0$; Coulomb potentials; T itself (with $a=0$ and $b=1$); components of the momentum operator $\boldsymbol{p}$; and nonrelativistic atomic and molecular Hamilto-

[2]Stakgold's inequalities [28] are reversed as he lists the eigenvalues in reverse order.

nians, say H_S. Clearly, T can be relatively form bounded by H_S+k, where $k>0$ is chosen large enough that H_S+k has a purely positive spectrum. Then if $\langle\psi|A|\psi\rangle$ satisfies Eq. (34), we choose $k>0$ so that $T+k$ is strictly positive, and the sequence ψ^n is E convergent to ψ in the $T+k$ norm, then

$$\|\psi^n-\psi\|_A^2\leqslant\max(a/k,b)\|\psi^n-\psi\|_{T+k}^2$$

so that $\{\psi^n\}$ is also A convergent to ψ. This means that it is sufficient for $\mathcal{W}$ to be complete in the Sobolev space $\mathcal{W}_2^{(1)}$.

Finally, it is straightforward to show that transition matrix elements of the form $\langle\psi_i^n|A|\psi_j^n\rangle$ converge to the desired limit $\langle\psi_i|A|\psi_j\rangle$ as $n\to\infty$ provided the sequences $\{\psi_i^n\}$ and $\{\psi_j^n\}$ are also A convergent.

C. Extension to Dirac operators

1. Operator domains

The Dirac operators occurring in atomic and molecular physics calculations have many features in common with Schrödinger operators but the analysis reveals some essential differences. The first step is to identify a domain on which the Dirac operator is essentially self-adjoint. The books of [30] (Chap. V.5), [31] (Chap. X), and [32] have a common strategy, which is to identify a suitable domain in the relevant Hilbert space in which the *free particle* Hamiltonian H_0 is essentially self-adjoint and then to establish the largest class of potentials V such that $H:=H_0+V$ is essentially self-adjoint on $\mathcal{D}(H_0)$.

In the nonrelativistic case, $H_0=\boldsymbol{p}^2/2$, H_0 is *essentially self-adjoint on* $C_0^\infty(\boldsymbol{R}^3)$, the space of infinitely differentiable functions with compact support in $\boldsymbol{R}^3$. The eigenfunctions of H_0 have square integrable partial derivatives of order 2 so that H_0 is self-adjoint on the Sobolev space $W_2(\boldsymbol{R}^3)$ ([31], p. 54, Theorem IX.27; p. 166, Ex. 1) demonstrating that when $V=-Z/r$, where r is the radial distance from the point charge nucleus of a hydrogenic atom, then $H:=H_0+V$ is essentially self-adjoint on $\mathcal{D}(H_0)$. Kato's theorem ([31] p. 166, Theorem X.16) extends this to the full nonrelativistic Hamiltonian for atoms and molecules.

The analogs for the relativistic Dirac Hamiltonian are more messy. When

$$H_0:=c\,\boldsymbol{\alpha}\cdot\boldsymbol{p}+c^2\beta \tag{35}$$

(using Hartree atomic units), then H_0 is essentially self-adjoint on $C_0^\infty(\boldsymbol{R}^3\backslash 0)^4$ and self-adjoint on the Sobolev space $\mathcal{D}(H_0)=W_1(\boldsymbol{R}^3)^4\subset L^2(\boldsymbol{R}^3)^4$ ([32], Theorem 1.1). In the case of the hydrogenic atom with a point nucleus, $V=\pm Z/r$, then $H=H_0+V$ is well defined and essentially self-adjoint on $C_0^\infty(\boldsymbol{R}^3\backslash 0)^4$ and self-adjoint on $\mathcal{D}(H_0)$ only if $Z/c<\sqrt{3}/2$ ([32], Theorem 4.4). Since $c\approx 137$ in atomic units, this restricts the potentials to $Z<118$, which satisfactorily covers all elements in the Periodic Table. However, the standard textbook analytic eigenfunctions for the relativistic hydrogenic atom are well behaved for $118<Z\leqslant 137$. There are two analytic solutions of the Dirac Coulomb problem: the solution *regular* at $r=0$ having amplitudes of order $O(r^\gamma)$ there (recall $\gamma=+\sqrt{\kappa^2-Z^2/c^2}$) and an *irregular* solution of order $O(r^{-\gamma})$. When $\kappa=\pm 1$ then $\gamma=1/2$ at $Z/c=\sqrt{3}/2$. So for $Z<118$, only the regular solution is square integrable near $r=0$, whereas both solutions are square integrable for $Z>118$, and a general square integrable eigensolution of the Dirac Coulomb problem is an arbitrary linear combination of the two. Order can be restored by noting that any solution containing a multiple of the irregular solution makes $\langle\psi|V|\psi\rangle$ infinite ([4], Sec. 22.5), so that the physics requires exclusion of the irregular solution to make the expectation of V finite. This behavior can also be interpreted in the language of the theory of differential equations either as a transition of the behavior at the origin from a limit-point description for $\gamma>1/2$ to a limit-circle description for $\gamma<\frac{1}{2}$ ([33], Chap. 11), or in terms of the theory of deficiency indices ([31], Chap. X).

The idealized point charge nucleus is not a good enough model for the heavier elements, or for very accurate work anywhere in the Periodic Table, and one can then replace the Coulomb potential by one which is infinitely differentiable. In this case H is again essentially self-adjoint on $\mathcal{D}(H_0)$ ([33], Theorem 4.3). The extension to many-electron atoms or molecules is not considered in [33], but there seem to be no insuperable difficulties once the single particle case has been understood.

2. Rayleigh-Ritz method for the Dirac operator

We now have a mathematical framework in which we can study the Rayleigh-Ritz method for Dirac operators. For simplicity, consider the Dirac Coulomb atomic Hamiltonian

$$H(\nu):=c\,\boldsymbol{\alpha}\cdot\boldsymbol{p}+mc^2\beta-\nu\frac{Z}{r} \tag{36}$$

with $Z/c<\sqrt{3}/2$ on $\mathcal{D}(H_0)=W_1(\boldsymbol{R}^3)^4$, where

$$\boldsymbol{\alpha}=\begin{bmatrix}\mathbf{0} & \boldsymbol{\sigma}\\ \boldsymbol{\sigma} & \mathbf{0}\end{bmatrix},\quad \beta=\begin{bmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & -\mathbf{I}\end{bmatrix},$$

and $\mathbf{I}$ is the two-dimensional identity matrix. The real number ν defines a family of Dirac Hamiltonians interpolating smoothly between $H(0)=H_0$ and $H(1)$ which incorporates the full strength Coulomb field. We assume a trial wave function of the general form

$$\phi:=\begin{bmatrix}\displaystyle\sum_{n=1}^{N} c_n^L M[L,n\kappa m,\boldsymbol{r}]\\ \\ \displaystyle i\sum_{n=1}^{N} c_n^S M[S,n\kappa m,\boldsymbol{r}]\end{bmatrix}. \tag{37}$$

Later we shall identify the expansion functions $M[T,n\kappa m,\boldsymbol{r}],T=L,S$ with the L spinors (15) and (16), but at this stage we merely assume that the set is complete in $\mathcal{D}(H_0)$. The Galerkin equations take the form

$$\begin{bmatrix} c^2\mathbf{S}^{LL}+\nu\mathbf{V}^{LL} & c\mathbf{\Pi}^{LS} \\ c\mathbf{\Pi}^{SL} & -c^2\mathbf{S}^{SS}+\nu\mathbf{V}^{SS} \end{bmatrix}\begin{bmatrix} \mathbf{c}^L \\ \mathbf{c}^S \end{bmatrix}=E\begin{bmatrix} \mathbf{S}^{LL} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}^{SS} \end{bmatrix}\begin{bmatrix} \mathbf{c}^L \\ \mathbf{c}^S \end{bmatrix} \tag{38}$$

in which all the submatrices are of dimension $N\times N$. The Gram matrices denoted by $\mathbf{S}^{TT}$, $T=L,S$, and $\mathbf{V}^{TT}$ are the corresponding matrices of the potential V. The matrices $\mathbf{\Pi}^{T\overline{T}}$ (where $\overline{T}=S$ when $T=L$ and vice versa) are given by

$$\mathbf{\Pi}^{LS}_{nn'} := \int M^{\dagger}[L,n\kappa m,\boldsymbol{r}]\boldsymbol{\sigma}\cdot\boldsymbol{p}\,M[S,n'\kappa m,\boldsymbol{r}]d\boldsymbol{r},$$

$$\mathbf{\Pi}^{SL}_{nn'} := \int M^{\dagger}[S,n\kappa m,\boldsymbol{r}]\boldsymbol{\sigma}\cdot\boldsymbol{p}\,M[L,n'\kappa m,\boldsymbol{r}]d\boldsymbol{r}.$$

The system is Hermitian and has real eigenvalues.

The original operator $H(0)$ has a pure absolutely continuous spectrum consisting of two disjoint intervals $\sigma_- =(-\infty,-mc^2]$ and $\sigma_+=[mc^2,\infty)$. Suppose that ϕ is an arbitrary trial function in $\mathcal{D}(H_0)$, and let

$$E_\phi(\nu)=\langle\phi|H(\nu)|\phi\rangle=E_\phi(0)+\nu V_\phi,$$

where $V_\phi=\langle\phi|V|\phi\rangle$. Suppose that ϕ has been chosen so that $E_\phi(0)>mc^2$. Since V has, by hypothesis, a strictly negative expectation, $V_\phi<0$, then $E_\phi(\nu)<E_\phi(0)$ decreases monotonically as ν increases from 0 to 1, and we shall have $mc^2>E_\phi(\nu)>-mc^2$ if $E_\phi(0)$ is not too large.[3] As only point eigenvalues of bound states can lie in $(-mc^2,mc^2)$, $E_\phi(\nu)$ will approximate a bound-state eigenvalue for values of ν in some interval $0<\nu_0\leqslant\nu\leqslant 1$. A sufficient condition that $E_\phi(1)>-mc^2$ is that $0>V_\phi\geqslant V_{min}>-2mc^2$, so that no state of the positive-energy spectrum can ever enter the ''negative-energy sea'' as ν approaches unity. Of course, a state with $E_\phi(0)<-mc^2$ will have $E_\phi(\nu)<-mc^2$ for all values $0\leqslant\nu\leqslant 1$. Another way of putting this is to say that the positive-energy eigenvalues are bounded below (by $mc^2+V_{min}>-mc^2$), and the negative-energy eigenvalues are bounded above (by $-mc^2$), in the manner suggested by Stakgold [28].

This behavior applies also to the eigensolutions of the Galerkin equations (38). We shall see that with L spinors (and with other admissible sets of expansion functions discussed by [4], with matched sets of N functions for both large and small components, we obtain a discrete pseudospectrum having N eigensolutions in the interval $-\infty<E_i^N<-mc^2$ $(i=1,2,\ldots,N)$ and N in $-mc^2<mc^2+V_{min}<E_i^N<\infty$ $(i=N+1,N+2,\ldots,2N)$. As N increases, the lower eigenvalues of the upper set, $E_{N+1}^N, E_{N+2}^N,\ldots$, converge to the lowest bound-state eigenvalues from above, exactly as they do in nonrelativistic calculations. The solutions with eigenvalues in the continuum regions, $E>mc^2$ above and $E<-mc^2$ below, represent continuum (standing-wave scattering solutions) in a region $0<r<R_{n\kappa}(\lambda)$ of size depending on the state and on the arbitrary parameter λ before decaying exponentially as $r\to\infty$. This behavior is consistent with the discussion following the Poincaré theorem.

Provided the Coulomb potential is not too strong, we see that the spectrum has two disjoint parts, and that the domain $\mathcal{D}(H)$ is the union of two disjoint subspaces, $\mathcal{D}_+$ and $\mathcal{D}_-$. This suggests we define a new form of convergence, D convergence (D for Dirac) on $\mathcal{D}_+$ with the norm

$$\|\psi\|_D=\langle\psi|H(0)|\psi\rangle,$$

$$\forall\,\psi \text{ such that } \langle\psi|H(0)|\psi\rangle\geqslant mc^2. \tag{39}$$

For basis sets that are complete with respect to this norm we can infer [3] the following.

(i) Rayleigh-Ritz convergence in $\mathcal{D}_+$ for *bound-state* energies and eigensolutions of the one-electron Dirac operator for atomic potentials. This may be extended to many-electron atoms and molecules (in the Born-Oppenheimer approximation with the nuclei in fixed positions) without difficulty.

(ii) A version of A convergence for operators that are relatively form bounded by $H(0)$ on $\mathcal{D}(H)$. Fortunately, this class includes most of the interesting operators in atomic and molecular physics:

(1) Bounded operators: constants, Dirac $\boldsymbol{\alpha}$ and β matrices, and operators such as $\boldsymbol{\alpha}\times\boldsymbol{r}$ and $\boldsymbol{\alpha}\cdot\boldsymbol{A}$, where $\boldsymbol{A}$ is the vector potential of some external field.

(2) Powers $r^\lambda,\lambda\geqslant -1$.

(3) Components of momentum $\boldsymbol{p}$ and combinations such as $\boldsymbol{\alpha}\cdot\boldsymbol{p}$.

(4) Other pieces of the Dirac operator itself.

V. APPLICATIONS TO RELATIVISTIC ATOMS

This section reviews some applications of the preceding theory of Rayleigh-Ritz methods to a range of physical problems. Much of the power of L spinors derives from the ability to write down simple algebraic expressions for the effective Hamiltonian in hydrogenic problems. The difficulty of evaluating matrix elements for the electron-electron interaction with L spinors effectively rules out their use in many-electron systems. Fortunately other sets of functions are available for that purpose ([4], Sec. 22.6.3).

For simplicity, we shall define $a:=2\gamma$, and exploit Pochhammer's symbol

$$(a)_0=1,$$

$$(a)_k=(a+k-1)\,(a)_{k-1}=a(a+1)\cdots(a+k-1),\quad k\geqslant 1.$$

We also use the notations

$$G_k(a)=\frac{(a+1)_k}{k!},\quad H_{mn}(a)=\sum_{k=0}^{min\ m,n} G_k(a-1). \tag{40}$$

Then

[3]For examples of this behavior see Ref. [34], Figs. 9.11 and 9.13.

$$V^T_{mn} = -Z\lambda\left[\frac{m!n!(a+m)(a+n)}{N_{m\kappa}N_{n\kappa}(N_{m\kappa}-\kappa)(N_{n\kappa}-\kappa)(a)_m(a)_n}\right]^{1/2}\left\{H_{m-1,n-1}(a)+\eta^T\frac{N_{n\kappa}-\kappa}{n+a}H_{m-1,n}(a)\right.$$
$$\left.+\eta^T\frac{N_{m\kappa}-\kappa}{m+a}H_{m,n-1}(a)+\frac{N_{m\kappa}-\kappa}{m+a}\frac{N_{n\kappa}-\kappa}{n+a}H_{mn}(a)\right\}, \tag{41}$$

where $\eta^L=-1$ and $\eta^S=+1$. The kinetic energy matrices are

$$\Pi^{SL}_{mn}=\Pi^{LS}_{nm}=\frac{\lambda}{2}\left[\frac{m!n!(a+m)(a+n)}{N_{m\kappa}N_{n\kappa}(N_{m\kappa}-\kappa)(N_{n\kappa}-\kappa)(a)_m(a)_n}\right]^{1/2}\left\{(2N_{n\kappa}+2n-2+a)\left[H_{m-1,n-1}(a)+\frac{N_{m\kappa}-\kappa}{m+a}H_{m,n-1}(a)\right]\right.$$
$$-\frac{N_{n\kappa}-\kappa}{n+a}(2n+2\kappa+a)\left[H_{m-1,n}(a)+\frac{N_{m\kappa}-\kappa}{m+a}H_{m,n}(a)\right]-2(n+a-1)\left[H_{m-1,n-2}(a)-\frac{N_{m\kappa}-\kappa}{m+a}H_{m,n-2}(a)\right]$$
$$\left.-G_{m-1,n-1}(a)-\frac{N_{m\kappa}-\kappa}{m+a}G_{m,n-1}(a)+\frac{N_{n\kappa}-\kappa}{n+a}G_{m-1,n}(a)+\frac{N_{m\kappa}-\kappa}{m+a}\frac{N_{n\kappa}-\kappa}{n+a}G_{mn}(a)\right\}, \tag{42}$$

whilst the symmetric tridiagonal Gram matrices $\boldsymbol{S}^{TT}$ are related to the expressions (24) by

$$S^{TT}_{ij}=g^{\kappa}_{ij}/2\lambda,$$

where the additional factor 2λ arises from the change of an independent variable from x to r. It is convenient that the parameter λ only enters these matrix definitions as a constant multiplier, so that it is easy to assess the effect of making changes to its value.

A. Hydrogenic atoms

Matrix diagonalizations were done using standard EISPACK routines with Fortran double precision arithmetic taking the speed of light $c=137.035\,989\,5$ a.u. (the currently accepted value of the reciprocal fine-structure constant).

The numerical method appears to be very stable. The eigenvalues of the Gram matrices lie, as predicted (25), in the range $(1-\rho_N,1+\rho_N)$, where $0<\rho_N<1$. Typical values of ρ_N and the condition number k_N appear in Table I. These are much smaller than the condition numbers, of order 10^8, reported for methods that use unrestricted kinetic balance [35]. Typical values of the lowest eigenvalue of the system

$$\boldsymbol{V}^{TT}\boldsymbol{c}=\upsilon\,\boldsymbol{S}^{TT}\boldsymbol{c},$$

the same for $T=L$ and $T=S$, estimating the lower bound V_{min} (Sec. IV C), are shown in Table II. This lower bound is in the gap $(0,-2mc^2)$ for $Z<135$, well beyond the atomic number $Z=118$ which is usually taken as the limit of self-adjointness of the Dirac Coulomb operator. The L spinors are constructed from functions that have a finite expectation of $1/r$, so that the condition to extend the range of Z from $Z=118$ to 137 is satisfied. We have successfully generated solutions when $|\kappa|=1$ even for values of Z as close to critical as 137.035 989, where $V_{min}<-2mc^2$ though the accuracy is no longer very good. This should not cause surprise, as the condition $V_{min}>-2mc^2$ is sufficient, but not necessary. We can generate solutions for larger values of $|\kappa|$ up to $Z=c|\kappa|$ in the same way.

Convergence of the solution as the basis-set dimension N is increased follows the expected pattern. The matrix diagonalizer produces ordered eigenvalues, in which those numbered $N+1,N+2,\ldots,2N$ correspond to the positive spectrum, and those numbered $1,\ldots,N$ lie below $-2mc^2$. The positive eigenvalues converge to the correct values from above, exactly as in nonrelativistic calculations, and for exactly the same reasons. A sample for $Z=50$, large enough to show some relativistic effects, appears in Table III, showing rapid stabilization of the lowest eigenvalues as N increases with $\lambda=50.0$. The special choice $\lambda=Z/N_{n_r\kappa}$ ensures that the $n_r\kappa$ eigenstate is represented by a single L spinor, although

TABLE I. Gram matrix conditioning.

Z	N	κ	$1-\rho_N$	k_N
10	100	−1	4.837×10^{-4}	4134
		1	9.702×10^{-4}	2061
		−2	9.702×10^{-4}	2061
		2	1.565×10^{-3}	1277
		−3	1.565×10^{-3}	1277
		3	2.257×10^{-3}	885
		−4	2.257×10^{-3}	885
		4	3.036×10^{-3}	658
		−5	3.036×10^{-3}	658
100	100	−1	2.173×10^{-4}	9203
		1	4.387×10^{-4}	4558
		−2	4.377×10^{-4}	4568
		2	7.103×10^{-4}	2815
		−3	7.107×10^{-4}	2813
		3	2.261×10^{-3}	884
		−4	2.260×10^{-3}	884
		4	1.395×10^{-3}	143
		−5	1.394×10^{-3}	143

TABLE II. V_{min} in atomic units. $N=100$.

Z	$\kappa=-1$	$\kappa=1$	$\kappa=-2$	$\kappa=2$	$\kappa=-3$
10	−552	−201	−201	−107	−107
20	−1117	−406	−402	−215	−215
30	−1708	−618	−607	−324	−323
40	−2342	−843	−815	−434	−432
50	−3039	−1086	−1028	−547	−542
60	−3827	−1355	−1248	−664	−653
70	−4748	−1663	−1476	−785	−767
80	−5868	−2026	−1715	−910	−883
90	−7296	−2476	−1966	−1042	−1002
100	−9243	−3069	−2232	−1181	−1123
110	−12176	−3928	−2517	−1328	−1249
120	−17440	−5407	−2823	−1487	−1378
130	−31921	−9286	−3157	−1657	−1512
$(-2mc^2=-37\,557.7248$ a.u.)					

other eigenstates will be represented by a nontrivial linear combination. Because the basis is complete, we expect that stable eigenvalues are insensitive to the choice of λ, and this is what we observe. Thus the eigenvalues ϵ_{ns}, $n=1-8$, are unchanged to seven decimal places, when λ is reduced from 50.0 to 30.0. The same behavior is observed with other symmetries and it is gratifying that, even though the matrix Hamiltonians are different, the stabilized eigenvalues display the correct $|\kappa|$ degeneracy, $\epsilon_{2s}=\epsilon_{2p_{1/2}}$, $\epsilon_{3p_{3/2}}=\epsilon_{3d_{3/2}}$, . . . , and the correct fine-structure ordering. The $(N+1)$th eigenvalue is always the lowest bound state of its symmetry and there are no spurious states, interlopers, or pathological behavior characterized as ''finite basis-set disease'' and ''continuum dissolution'' in such papers as [36].

The Nth eigenvalue of each symmetry is always the highest in the ''negative-energy sea'' and is always safely below the upper bound $-2mc^2$. Since there are no bound states, we expect to see no convergent sequences of eigenvalues as N increases nor do we find any. The sensitivity of the highest negative eigenvalue to both N and λ, Table IV, illustrates these conclusions.

B. Static dipole polarizability of relativistic hydrogenic atoms

This much studied problem (see [16] for the extensive literature) can be used to illustrate two uses of L-spinor completeness: evaluation of matrices of simple functions of coordinates, and the evaluation of perturbation sums. The second-order static dipole polarizability α_{zz} for the state $|\psi_0\rangle$, energy ϵ_0, of a hydrogenic atom is given by the formula

$$\alpha_{zz}=2\sum_{n\neq 0}\frac{\langle\psi_0|z|\psi_n\rangle\langle\psi_n|z|\psi_0\rangle}{\epsilon_n-\epsilon_0}, \tag{43}$$

where the restriction $n\neq 0$ excludes ψ_0 from the sum over states. In the Dirac case, the sum includes both positive- and negative-energy eigenstates. For brevity, we treat only the

TABLE III. Convergence with respect to the matrix block dimension N for positive-energy states of a hydrogenic atom with $Z=50$.

N	ϵ_{1s}	ϵ_{2s}	ϵ_{3s}	ϵ_{4s}	ϵ_{5s}
			λ = 50.0		
20	−1294.62616	−326.494806	−143.829353	−79.5730938	−35.1391668
40	−1294.62616	−326.494806	−143.829802	−80.3703311	−51.1923424
60	−1294.62616	−326.494806	−143.829802	−80.3703316	−51.1977244
80	−1294.62616	−326.494806	−143.829802	−80.3703316	−51.1977244
	$\epsilon_{2p_{1/2}}$	$\epsilon_{3p_{1/2}}$	$\epsilon_{4p_{1/2}}$	$\epsilon_{5p_{1/2}}$	$\epsilon_{6p_{1/2}}$
			λ = 25.0		
20	−326.494806	−143.829807	−80.3703372	−51.1972465	−35.2027152
40	−326.494806	−143.829803	−80.3703331	−51.1977253	−35.433571
60	−326.494806	−143.829802	−80.3703323	−51.1977248	−35.4335707
80	−326.494806	−143.829802	−80.370332	−51.1977247	−35.4335706
100	−326.494806	−143.829802	−80.3703319	−51.1977246	−35.4335705
	$\epsilon_{2p_{3/2}}$	$\epsilon_{3p_{3/2}}$	$\epsilon_{4p_{3/2}}$	$\epsilon_{5p_{3/2}}$	$\epsilon_{6p_{3/2}}$
			λ = 25.0		
20	−315.144355	−140.457874	−78.952058	−50.4731861	−34.7554737
40	−315.144355	−140.457874	−78.952058	−50.4738674	−35.0157937
60	−315.144355	−140.457874	−78.952058	−50.4738674	−35.0157937
	$\epsilon_{3d_{3/2}}$	$\epsilon_{4d_{3/2}}$	$\epsilon_{5d_{3/2}}$	$\epsilon_{6d_{3/2}}$	$\epsilon_{7d_{3/2}}$
			λ = 15.0		
20	−140.457874	−78.9520581	−50.4738674	−35.0157937	−25.7034854
40	−140.457874	−78.952058	−50.4738674	−35.0157937	−25.7037387
60	−140.457874	−78.952058	−50.4738674	−35.0157937	−25.7037387

TABLE IV. Highest negative-energy eigenvalues for $Z=50$ (a.u.) relative to $-2mc^2$.

λ \ $N=$	20	40	60	80	100
			$\kappa=-1$		
30	−57.6	−26.6	−17.0	−12.4	−9.7
40	−80.3	−36.7	−23.4	−17.0	−13.3
50	−104.3	−47.4	−30.1	−21.8	−17.1
			$\kappa=+1$		
20	−46.6	−20.7	−13.0	−9.4	−7.3
25	−61.1	−27.0	−16.9	−12.1	−9.4
30	−76.4	−33.6	−20.9	−15.0	−11.6
			$\kappa=-2$		
20	−34.8	−16.6	−10.7	−7.9	−6.2
25	−44.9	−21.2	−13.7	−10.1	−7.9
30	−55.4	−26.0	−16.8	−12.3	−9.7
			$\kappa=+2$		
10	−19.6	−9.1	−5.9	−4.3	−3.3
15	−31.4	−14.6	−9.3	−6.7	−5.3
20	−44.2	−20.4	−12.9	−9.4	−7.3

TABLE V. Contributions κ, Eq. (45), computed with the convergence criterion, Eq. (46), $\eta=10^{-6}$.

Z	Z^4+1	N_{+1}	Z^4-2	N_{-2}
1	6.749531	2	6.749676	2
5	6.738272	2	6.741888	2
10	6.703128	4	6.717556	2
15	6.644694	6	6.677018	2
20	6.563177	10	6.620296	3
30	6.332152	18	6.458439	3
40	6.013420	26	6.232346	3
50	5.611749	36	5.942529	4
60	5.133374	46	5.589629	4
70	4.586085	56	5.174405	3
80	3.979358	66	4.697663	5
90	3.324546	76	4.160097	6
100	2.635150	86	3.561882	6
110	1.927202	96	2.901526	5
120	1.219632	102	2.171964	10
130	0.532359	108	1.339899	16
135	0.192185	103	0.819494	22

case in which ψ_0 is the $1s_{1/2}$ ground state with angular-momentum projection $m_j=+1/2$; angular-momentum selection rules then restrict the intermediate sum to states with symmetry types $\kappa=+1,\kappa=-2$, both with $m_j=+\frac{1}{2}$. We can therefore ignore the quantum number m_j in what follows. After performing angular integrations, we obtain the atomic dipole polarizability of the $1s$ state in the form

$$\alpha_d=\frac{2}{9}(\Delta_{+1}+2\Delta_{-2}), \tag{44}$$

Szmytkowski has presented analytical expressions for the quantities Δ_{+1} and Δ_{-2} ([16], Eqs. (182)–(184)) with which we have compared our numerical results in Table V. These have been obtained by computing the sums

$$\Delta\kappa=\sum_{n\neq 0}^{+}\frac{(0|r|n^+\kappa)(n^+\kappa|r|0)}{\epsilon_n^+-\epsilon_0^+}+\sum_{n}^{-}\frac{(0|r|n^-\kappa)(n^-\kappa|r|0)}{\epsilon_n^--\epsilon_0^+}, \tag{45}$$

where superscripts $+/-$ designate the two branches of the pseudospectrum and the matrix elements are now purely radial. It is convenient to choose the tuning parameter to have the value $\lambda=Z$, so that the $1s$ reference state ψ_0 is represented exactly by the L spinors with $n_r=0,\kappa=-1$. The basis-set dimensions N_κ have been adjusted so that

$$Z^4|\Delta_\kappa-\Delta_\kappa^{analytic}|\leqslant\eta. \tag{46}$$

The value $\eta=10^{-6}$ used in Table V ensures agreement of our numerical values with the analytic values to six significant figures. We have done similar calculations at higher precision, but these take longer to run.

The results are relatively insensitive to the choice of λ over a wide range, consistent with the completeness of L-spinor sets. However, it also influences the rate of convergence of perturbation expansions making it essential to examine the sensitivity to λ in each application.

C. A simple perturbation problem

Perturbation calculations show the completeness properties of L spinors to advantage, in this case to study the convergence of the perturbation expansion of the energy of a hydrogenic atom in which the nuclear charge is perturbed from Z to $Z+Z'$ in powers of Z'. This model was first studied nonrelativistically by Rossky and Karplus [37]; it has the advantage that the states of the unperturbed system are known exactly, as is the final answer, so that the error at each order of perturbation is easy to establish. In the nonrelativistic case, the sum over all diagrams contributing to the energy of order k vanishes for $k>2$, since $\epsilon_{nl}=-(Z+Z')^2/2n^2\,E_h$. However, there are usually several diagrams of order $k\geqslant 3$ which should sum to zero for each value of k, but Rossky and Karplus found that this was only true approximately. Although summing over the discrete spectrum is straightforward, integrating over continuum states is troublesome. Whilst in principle this is just a quadrature, integrals over the perturbation $-Z'/r$ which are diagonal in energy diverge, so that the energy integration needs to be done with care. We showed [19] that all these difficulties could be avoided if we solved the Schrödinger equation for the nonrelativistic hydrogenic atom using simple finite basis sets of either exponential or Gaussian form. The more difficult Dirac case was studied in [20]; the higher-order diagrams of order $k\geqslant 3$ no longer sum to zero, and we found that it was essential to include negative-energy sums in order to get the analytic results obtained by expanding the Sommerfeld formula for charge $Z+Z'$ in powers of Z'.

TABLE VI. Perturbed hydrogenic 1*s* state.

Z	ϵ_0	ϵ_2^+	ϵ_2^-	ϵ_2	ϵ_2 [Eq. (48)]
10	−50.066742	−0.504124	0.000103	−0.504021	−0.504021
20	−201.076523	−0.517050	0.000639	−0.516412	−0.516412
30	−455.524907	−0.539986	0.001761	−0.538225	−0.538225
40	−817.807498	−0.575014	0.003560	−0.571454	−0.571455
50	−1294.626156	−0.625644	0.006153	−0.619491	−0.619495
60	−1895.682356	−0.697763	0.009731	−0.688032	−0.688042
70	−2634.846565	−0.801543	0.014631	−0.786912	−0.786943
80	−3532.192151	−0.955598	0.021459	−0.934139	−0.934228
90	−4617.757654	−1.197084	0.031394	−1.165689	−1.165967
100	−5939.195384	−1.610319	0.047022	−1.563297	−1.564287

Here we treat the perturbed 1*s* level, for which we have the simple formula

$$\epsilon(Z)=(\gamma-1)c^2, \quad \gamma=\sqrt{1-Z^2\alpha^2}, \tag{47}$$

so that the perturbation series takes the form

$$\epsilon(Z+Z')=\epsilon(Z)+\epsilon_1(Z)Z'+\epsilon_2(Z)Z'^2+\cdots \tag{48}$$

where, using Taylor's theorem,

$$\epsilon_1(Z)=\epsilon'(Z)=-Z/\gamma,$$

$$\epsilon_2(Z)=\frac{1}{2}\epsilon''(Z)=-1/2\gamma^3, \ldots$$

which clearly agrees with the nonrelativistic result in the limit $c\to\infty$. In the finite dimensional formulation, we have to evaluate the expressions

$$\epsilon_1(Z)=\langle\psi_0|1/r|\psi_0\rangle,$$
$$\epsilon_2(Z)=\sum_{n\neq 0}\frac{\langle\psi_0|1/r|\psi_n\rangle\langle\psi_n|1/r|\psi_0\rangle}{\epsilon(Z)-\epsilon_n(Z)}, \tag{49}$$

where $\epsilon_n(Z)$ is the eigenvalue associated with $|\psi_n\rangle$, the sum extending over the complete spectrum of positive- and negative-energy states. Table VI displays the results computed with block dimension $N=100$ taking $\lambda=Z$ to be fixed. This is not sufficient for the highest values of Z; the difference between the sum over states in the penultimate column and the exact value in the last column gives some idea of the error. The negative-energy state contribution grows roughly like Z^3, and is clearly non-negligible for high Z, demonstrating that any perturbation of the Dirac Hamiltonian, whether one-electron or two-electron, contains contributions from unperturbed negative-energy states. Thus attempts to solve many-electron problems using methods that attempt to project out negative-energy contributions [36] always introduce unquantifiable errors as well as formidable technical complications.

VI. DISCUSSION

We have set out the theory of *L* spinors, and demonstrated their orthonormality properties. The theory of Raleigh-Ritz approximation has been formulated for Dirac operators, and we have illustrated it with applications to a range of problems in the theory of hydrogenic atoms using a basis of *L* spinors. We have verified the existence of a lower bound to the positive branch of the spectrum in the bound-state gap, and shown that in a basis of dimension *N*, the *N* lowest eigenvalues lie in the lower continuum region whilst the *N* highest approximate bound states and the positive continuum. The completeness properties of *L* spinors in the Hilbert spaces of Sec. IV C are essential for the convergence of perturbation sums in Secs. V B and V C. The numerical accuracy of these calculations depends upon the numerical stability of *L*-spinor expansions for large *N*, and we have found that the Gram matrices with $N\sim 100$ have condition numbers of order 10^3, comparable to those of nonrelativistic Sturmians, rather than the order 10^8 quoted for extended small component basis sets in [35].

The applications discussed in this paper by no means exhaust applications of *L* spinors, although they are largely limited to problems involving hydrogenic atoms as it is extremely difficult to evaluate electron-electron interactions in a computationally practical way. For this reason, we have introduced other types of basis function, *S* spinors and *G* spinors, see for example [2–4] which have similar properties making them suitable for the Rayleigh-Ritz approximation of Dirac many-electron atomic and molecular problems. However, nonrelativistic Sturmians were originally introduced as a way to approximate the effect of continuum states [17,18] and these ideas can be extended naturally to the Dirac case.

The Dirac-Coulomb Sturmians defined by Szmytkowski [16] are closely related to the *L* spinors presented in this paper, although there are major differences. *L* spinors involve paired *two*-component basis sets for large and small components, whereas Szmytkowski deals only with *four*-component objects and focuses on the construction of the Dirac-Coulomb Green's function. He has correctly noted that *L* spinors do not satisfy the differential equation of Ref. [4] [(Eq. (22.145)], an error corrected in Sec. III of this paper. However, his suggestion that the *L*-spinor basis sets are incomplete ([16], p. 837) is clearly wrong (see Sec. III C). It is

unlikely that we could have reproduced exactly the same results as Szmytkowski for relativistic hydrogenic polarizabilties had there been any substance to this claim. It is indeed gratifying that these two very different computational approaches to this problem agree so well.

ACKNOWLEDGMENTS

This work was supported by EPSRC Research Grant No. GR/L55155. H.Q. wishes to acknowledge further support from the Australian Research Council.

[1] H. M. Quiney, I. P. Grant, and S. Wilson, in *Lecture Notes in Quantum Chemistry*, edited by U. Kaldor (Springer-Verlag, Berlin, 1989), Vol. 52, pp. 307–344.
[2] I. P. Grant and H. M. Quiney, Adv. At. Mol. Phys. **23**, 37 (1988).
[3] I. P. Grant, in *Relativistic, Quantum Electrodynamic and Weak Interaction Effects in Atoms*, edited by P. J. Mohr, W. R. Johnson, and J. Sucher, AIP Conf. Proc. No. 189 (AIP, New York, 1989), pp. 235–253.
[4] I. P. Grant, in *Atomic, Molecular and Optical Physics Reference Book*, edited by G. W. F. Drake (AIP, New York, 1996), Chap. 22.
[5] H. M. Quiney, I. P. Grant, and S. Wilson, J. Phys. B **20**, 1412 (1987).
[6] I. P. Grant, in *Methods of Computational Chemistry, Vol. 2*, edited by S. Wilson (Plenum Press, New York, 1988), pp. 1–71.
[7] H. M. Quiney, in *Methods of Computational Chemistry, Vol. 2*, edited by S. Wilson (Plenum Press, New York, 1988), pp. 227–278.
[8] H. M. Quiney, H. Skaane, and I. P. Grant, Adv. Quantum Chem. **32**, 1 (1997).
[9] H. M. Quiney, H. Skaane, and I. P. Grant, J. Phys. B **30**, L829 (1997).
[10] H. M. Quiney, H. Skaane, and I. P. Grant, J. Phys. B **31**, L85 (1998).
[11] H. M. Quiney, H. Skaane, and I. P. Grant, Chem. Phys. Lett. **290**, 473 (1998).
[12] I. P. Grant and H. M. Quiney, J. Phys. A **26**, 7547 (1993).
[13] I. P. Grant and H. M. Quiney, Phys. Scr. **T46**, 132 (1993).
[14] I. P. Grant, Adv. At., Mol., Opt. Phys. **32**, 169 (1994).
[15] H. M. Quiney and I. P. Grant, J. Phys. B **27**, L299 (1994).
[16] R. Szmytkowski, J. Phys. B **30**, 825 (1997).
[17] M. Rotenberg, Ann. Phys. (N.Y.) **19**, 262 (1962).
[18] M. Rotenberg, Adv. At. Mol. Phys. **6**, 233 (1970).
[19] H. M. Quiney, I. P. Grant, and S. Wilson, J. Phys. B **18**, 577 (1985).
[20] H. M. Quiney, I. P. Grant, and S. Wilson, J. Phys. B **18**, 2805 (1985).
[21] *Handbook of Mathematical Functions*, edited by M. Abramowitz and I. A. Stegun (Dover, New York, 1970).
[22] G. Szeg'ö, *Orthogonal Polynomials*, American Mathematical Society Colloquium Publication No. 23 (American Mathematical Society, Providence, 1959).
[23] R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience, New York, 1953), Vol. 1.
[24] B. Klahn and W. A. Bingel, Theor. Chim. Acta **44**, 9,27 (1977).
[25] S. Gerschgorin, Izv. Akad. Nauk. USSR, Ser. Math. **7**, 749 (1931).
[26] J. H. Wilkinson, *The Algebraic Eigenvalue Problem* (Clarendon Press, Oxford, 1965).
[27] I. P. Grant, J. Phys. B **19**, 3187 (1986).
[28] I. Stakgold, *Green's Functions and Boundary Value Problems* (Wiley, New York, 1979).
[29] B. Klahn and J. D. Morgan III, J. Chem. Phys. **81**, 410 (1984).
[30] T. Kato, *Perturbation Theory for Linear Operators*, 2nd ed. (Springer-Verlag, Berlin, 1976).
[31] M. Reed and B. Simon, *Methods of Modern Mathematical Physics, II. Fourier Analysis and Self-Adjointness* (Academic, New York, 1975).
[32] B. Thaller, *The Dirac Equation* (Springer-Verlag, Berlin, 1992).
[33] R. D. Richtmyer, *Principles of Advanced Mathematical Physics* (Springer-Verlag, New York, 1978), Vol. 1.
[34] W. Greiner, *Relativistic Quantum Mechanics*, 2nd ed. (Springer-Verlag, Berlin, 1985).
[35] T. Saue, K. Fægri, Jr., T. Helgaker, and O. Gropen, Mol. Phys. **91**, 937 (1997).
[36] W. Kutzelnigg, Int. J. Quantum Chem. **25**, 107 (1984).
[37] P. Rossky and M. Karplus, J. Chem. Phys. **67**, 5419 (1977).