Conditions for convergence of variational solutions of Dirac's equation in a finite basis

Ian P. Grant*
Instituto de Física, Universidad Nacional Autónoma de México,
Apartado Postal 20-364, Delegación Alvaro Obregón, 01000 México, D.F., Mexico
(Received 23 June 1981; revised manuscript received 14 September 1981)

A recent variational study of the Dirac Coulomb problem using a discrete basis set leads to a conjecture of a type of Hylleraas-Undheim theorem for "positive energy" solutions. Analysis of the matrices involved suggests convergence criteria for the method which may be of use in self-consistent-field calculations.

Of the many recent attempts to perform relativistic molecular structure calculations using variational methods with discrete basis sets, three recent papers[1–3] on the hydrogenlike atom are particularly instructive. We describe here a matrix analysis of the calculation of Drake and Goldman[3] which may also have relevance to relativistic molecular (self-consistent-field) SCF calculations.

Drake and Goldman[3] separated angular and radial parts of the Dirac operator in the usual way and solved the radial equations

$$H_\kappa\Phi(r)\equiv\begin{bmatrix}-Z/r & c\left(-\frac{d}{dr}+\frac{\kappa}{r}\right)\\ c\left(\frac{d}{dr}+\frac{\kappa}{r}\right) & -2c^2-Z/r\end{bmatrix}\Phi(r)$$
$$=\epsilon\Phi(r)\ ,$$
$$\Phi(r)=\begin{bmatrix}g(r)\\ f(r)\end{bmatrix}\ , \qquad (1)$$
$$\int_0^\infty\{[f(r)]^2+[g(r)]^2\}\,dr=1\ .$$

Here c is the velocity of light (α^{-1} a.u.), and the zero of energy is shifted by subtracting the electron rest energy (c^2) so that, in their notation,

$$H_\kappa=cH_r-c^2,\quad \epsilon=E-c^2\ .$$

We need some mathematical properties[4] of the operator H_κ. This is defined on a domain which it is natural[5] to take as a dense subset of the Hilbert space $\mathfrak{K}\equiv[L^2(\mathfrak{R})]^2$ of two-component functions $\Phi(r)$, of which each component is a distribution in $L^2(\mathfrak{R})$. The free-particle operator

$$H_\kappa^{(0)}=\begin{bmatrix}0 & c\left(-\frac{d}{dr}+\frac{\kappa}{r}\right)\\ c\left(\frac{d}{dr}+\frac{\kappa}{r}\right) & -2c^2\end{bmatrix} \qquad (2)$$

is self-adjoint on the domain, $\mathfrak{D}(H_\kappa^{(0)})=\{\Phi\epsilon\mathfrak{K}:H_\kappa^{(0)}\Phi\epsilon\mathfrak{K}\}$. Moreover, $H_\kappa=H_\kappa^{(0)}-Z/r$ (Ref. 6) is self-adjoint on $\mathfrak{D}(H_\kappa^{(0)})$ provided $(Z/c)^2<\frac{3}{4}$ ($Z<118$). This condition appears because $f(r)$ and $g(r)$, the components of Φ, both behave like r^γ or like $r^{-\gamma}$ near the origin, $\gamma=[\kappa^2-Z^2/c^2]^{1/2}$, $|\kappa|$ is an integer. When $\gamma>\frac{1}{2}$, only the r^γ solution is in $\mathfrak{K}$(the limit-point case). When $0<\gamma<\frac{1}{2}$ both r^γ and $r^{-\gamma}$ solutions are in $\mathfrak{K}$ (the limit circle case), and a further condition, which is automatically satisfied for $\gamma>\frac{1}{2}$,

$$\int_0^\infty\Phi^\dagger(r)\Phi(r)\frac{dr}{r}<\infty,\quad \Phi\in\mathfrak{K} \qquad (3)$$

(finite expectation of the potential) is needed to ensure uniqueness, eliminating the $r^{-\gamma}$ solution.

In these conditions H_κ has a discrete spectrum (the Sommerfeld eigenvalues) in $-2c^2<\epsilon<0$; the corresponding eigensolutions lie in $\mathfrak{K}$,[4] having real exponential behavior like $e^{\lambda r}$ or $e^{-\lambda r}$, $\lambda=+[-2\epsilon(1+\epsilon/2c^2)]^{1/2}$, as $r\to\infty$.[7] On the rest of the real line excluding $(-2c^2, 0)$, λ is pure imaginary and all eigensolutions lie in the continuum. Further details can be found in Richtmyer's book.[4]

Drake and Goldman[3] expanded $\Phi(r)$ in a set $\{\chi_j:j=1,\ldots,N\}$ of Slater type orbitals (STO) on $[0,\infty)$, $\chi_j=e^{-\mu r}\times r^{\gamma+j-1}$, $\gamma=(\kappa^2-Z^2/c^2)^{1/2}$; for $N\to\infty$ and any real $\mu>0$, the set $\{A\chi_j:\ A=\binom{1}{0},\binom{0}{1}\}$ is dense in $\mathfrak{K}$. They determined estimates of eigenfunctions $\Phi_i^{(N)}$ and eigenvalues $\epsilon_i^{(N)}$ variationally for $\kappa=-1,1,-2,2$ ($s_{1/2}$, $p_{1/2}$, $d_{3/2}$, $d_{5/2}$ symmetries) and $N=1,2,\ldots,14$ in the usual way. The pertinent results are (a) the eigenvalues $\epsilon_i^{(N)}$ of the discrete problem divide into two classes "positive energy" and "negative energy" each having N members.[8] (b) There appears to be a sort of Hylleraas-Undheim theorem for the "positive" eigenvalues $\epsilon_i^{(N)+}$, which appear to converge from above to the corresponding Sommerfeld eigenvalues. The lowest positive root for $\kappa>0$ is degenerate with the lowest root for the case $-\kappa$ and is clearly spuri-

ous. (c) The N negative energy eigenvalues all lie below $-2c^2$ in the continuous spectrum of H_κ, and do not converge as $N\to\infty$. These correspond to square integrable pseudostates which may be important in some applications, where a complete (finite) basis set is needed.[3] An analysis of the process[3] allows us to interpret these observations. It is convenient to replace the STO set $\{\chi_j\}$ by an equivalent orthonormal set[9] $\{\phi_j\}$, $\phi_j(r) = [(n-1)!2\mu/\Gamma(2\gamma+n)]^{1/2}\exp(-\mu r)\cdot(2\mu r)^\gamma L_{j-1}^{(2\gamma)}(2\mu r)$, $0<r<\infty$, $j=1,2,\ldots,$ where $L_n^{(\alpha)}(\rho)$ is a generalized Laguerre polynomial. The variational procedure[3] requires us to solve a matrix eigenvalue problem

$$\begin{bmatrix} -ZW & c\Pi_\kappa \\ c\Pi_\kappa^\dagger & -2c^2 I - ZW \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \epsilon \begin{bmatrix} a \\ b \end{bmatrix} , \tag{4}$$

where a, b are N vectors of expansion coefficients for $f(r)$ and $g(r)$, respectively, and I, W, Π_κ, and its adjoint $\Pi_\kappa^\dagger$ are $N\times N$ matrices. By virtue of (5), W is bounded, and it is also symmetric and positive definite, with elements

$$W_{ij} = \int_0^\infty \phi_i(r)\phi_j(r)\frac{dr}{r} , \tag{5}$$

$$i,j = 1,2,\ldots,N .$$

If D is the antisymmetric matrix with matrix elements

$$D_{ij} = -D_{ji} = \int_0^\infty \phi_i(r)\frac{d}{dr}\phi_j(r)\,dr , \tag{6}$$

$$i,j, = 1,2,\ldots,N ,$$

then

$$\Pi_\kappa = \kappa W - D, \quad \Pi_\kappa^\dagger = -\Pi_{-\kappa} = \kappa W + D . \tag{7}$$

We observe that

$$\begin{aligned} \Pi_\kappa\Pi_\kappa^\dagger &= \kappa^2 W^2 - \kappa[D,W] - D^2 , \\ \Pi_\kappa\Pi_\kappa^\dagger &= \kappa^2 W^2 + \kappa[D,W] - D^2 . \end{aligned} \tag{8}$$

The formal commutation relation $[d/dr, r^{-1}] = -r^{-2}$ corresponds to a matrix equivalence of $[D,W]$ and $-W^2$. This cannot be exact for any finite value of N, and the rate of convergence of $[D,W]$ to $-W^2$, if any, as $N\to\infty$ will depend crucially on the choice of basis set. Nevertheless, we can identify the matrices (8) as (possibly poor) matrix representations of the radial Schrödinger kinetic energy operators $-\frac{1}{2}d^2/dr^2 + l(l+1)/2r^2$, and we write

$$T_l = \tfrac{1}{2}\Pi_\kappa\Pi_\kappa^\dagger, \quad T_{\bar l} = \tfrac{1}{2}\Pi_\kappa^\dagger\Pi_\kappa , \tag{9}$$

where $l,\bar l$ are the usual orbital angular quantum numbers corresponding to $g(r)$ and $f(r)$, defined as non-negative solutions of $l(l+1) = \kappa(\kappa+1)$, and $l(l+1) = (-\kappa)(-\kappa+1)$, respectively. The matrix equivalent of the textbook "elimination of small components" gives

$$[H_l^{(Z)} - \epsilon I - K_\kappa^{(Z)}(\epsilon)]a = 0 , \tag{10}$$

where $H_l^{(Z)} = T_l - ZW$ is a matrix representation of the radial Schrödinger Coulomb Hamiltonian and

$$K_\kappa^{(Z)}(\epsilon) = \tfrac{1}{2}\Pi_\kappa\left(I + \frac{ZW+\epsilon I}{2c^2}\right)^{-1}\left(\frac{ZW+\epsilon I}{2c^2}\right)\Pi_\kappa^\dagger . \tag{11}$$

It is clear from this that the operator $K_\kappa^{(Z)}(\epsilon)$, and hence a, ϵ (and, by implication, cb), Eq. (4), can be expanded in a regular perturbation series in $1/c^2$; this is a matrix analog of a theorem of Hunziker.[10] Similarly, shifting the zero of energy by writing $\epsilon = -2c^2 + \epsilon'$, "elimination of large components," gives, using the notation of (10),

$$\{H_l^{(-Z)} + \epsilon' I - [K_\kappa^{(-Z)}(-\epsilon')]^\dagger\}b = 0 . \tag{12}$$

Next, we examine the formal nonrelativistic limit $c\to\infty$. Whether we allow the parameter γ to vary or not, $K_\kappa^{(Z)}(\epsilon)\to 0$ and a, ϵ become an eigenvector and eigenvalue of $H_l^{(Z)}$. Thus, the quality of the solution depends, to start with, on the quality of this matrix representation of the Schrödinger operator, which may *underestimate* energies. Similarly $-\epsilon'$ and b are an eigenvalue and eigenvector of $H_l^{(-Z)}$, having a *repulsive* Coulomb potential in line with the identification of negative-energy states with those of positrons. The N eigenvalues ϵ_i' are therefore negative, $c\to\infty$, with the true zero of energy $\epsilon = -2c^2\to-\infty$. The N eigenvalues of (10), $\epsilon_i^{(N),+}$, approximate the Rydberg eigenvalues and should converge as predicted by the Hylleraas-Undheim theorem.[11]

To analyze the case of c finite, rewrite (10) as a Rayleigh quotient

$$\epsilon = a^\dagger[H_l^{(Z)} - K_\kappa^{(Z)}(\epsilon)]a/a^\dagger a . \tag{13}$$

Whenever ϵ *is such that* $(ZW+\epsilon I)$ is *positive definite,* $a^\dagger K_\kappa^{(Z)}(\epsilon)a$ is positive, so that relativistic corrections decrease ϵ. Suppose also that W has a finite spectral radius $\rho_W^{(N)}$; this is required by (3). Then, for all a such that $x = \Pi_\kappa^\dagger a$ is non-null

$$\begin{aligned} R(\epsilon,a) &= a^\dagger K_\kappa^{(Z)}(\epsilon)a/a^\dagger a \\ &= \frac{1}{2}\left[x^\dagger\left(I + \frac{ZW+\epsilon I}{2c^2}\right)^{-1}(ZW+\epsilon I)/2c^2 x\right]\Big/ a^\dagger a \\ &\leqq \frac{a^\dagger T_l a}{a^\dagger a}\left(1 + \frac{Z\rho_W^{(N)}+\epsilon}{2c^2}\right)^{-1}(Z\rho_W^{(N)}+\epsilon)/2c^2 . \end{aligned}$$

Provided $\sup_N \rho_W^{(N)}$ is finite (a requirement which may restrict the basis set), $R(\epsilon,a) < a^\dagger Ta/a^\dagger a$, on $0 > \epsilon > -2c^2$ and the right-hand side of (13) is

uniformly bounded below. This, in conjunction with the separation property $\epsilon_i^{(k+1)} < \epsilon_i^{(k)} < \epsilon_{i+1}^{(k+1)}$, $i = 1, 2, \ldots, k$, $k = 1, 2, \ldots$ of leading diagonal submatrices of any real symmetric matrix[11] like that of (10), suffices to explain observation B of Drake and Goldman's calculation.[3] A similar analysis shows that the remaining N eigenvalues all lie in the lower continuum.

This discussion reveals several points at which the numerical results may be sensitive to the choice of basis set. The numerical difficulties observed in many recent calculations on atoms and molecules[1-3] which are usually just attributed to the presence of negative-energy states are obviously closely connected. Here we suggest *sufficient* conditions for convergence which may be useful to test particular choices of basis set. Necessary conditions are at present out of reach. We expect the analysis given here to generalize in several directions. There should be no difficulty in extending it to include more general effective atomic central potentials of the Hartree-Fock type with Coulomb singularities at the origin. Likewise, it should not be difficult to replace the Coulomb singularity with the potential appropriate to a distributed nuclear charge. It should also be possible to generalize the argument to cope with the self-consistent-field problem for atoms and molecules, exploiting the similarities in matrix structure[2] to (7). Any progress in these directions will be reported elsewhere.

The support of Academia de la Investigación Científica de México under the exchange agreement with the Royal Society, and of Instituto de Física, UNAM, are gratefully acknowledged. I am particularly grateful to Dr. C. F. Bunge and Dr. T. A. Brody for conversations which helped considerably to clarify the arguments presented in this paper, and to Professor E. R. Davidson for some enlightening comments on basis sets.

---

*On leave of absence from Dept. of Theoretical Chemistry, 1 South Parks Road, Oxford, OX1 3TG United Kingdom.

[1]F. Mark, H. Lischka, and F. Rosicky, Chem. Phys. Lett. 71, 507 (1980).

[2]F. Mark and F. Rosicky, Chem. Phys. Lett. 74, 562 (1980).

[3]G. W. F. Drake and S. P. Goldman, Phys. Rev. A 23, 2093 (1981).

[4]R. D. Richtmyer, *Principles of Advanced Mathematical Physics* (Springer, New York, 1978), Vol. 1, Secs: 10.17, 11.6.

[5]It is often asserted that the components $g(r)$ and $f(r)$ are in different spaces, but this is only true in a very trivial sense. The theory makes it clear that each such space is a copy of $L^2(\mathfrak{R})$ and, indeed, the symmetric way in which $g(r)$ and $f(r)$ enter the equations demands they be treated on the same footing.

[6]Z/r is a multiplicative operator.

[7]It is the prescription of exponential decay at large distances and the correct power-series exponent near the origin that ensures that calculations based on numerical integration methods do not suffer from variational collapse into negative energy states.

[8]The terms positive energy and negative energy, which are used conventionally in this context, are enclosed by quotation marks as they are not strictly appropriate with our choice of energy zero.

[9]P. O. Löwdin and H. Shull, Phys. Rev. 101, 1730 (1956).

[10]W. Hunziker, Commun. Math. Phys. 40, 215 (1975).

[11]J. K. L. MacDonald, Phys. Rev. 43, 830 (1933); the theorem, in a different guise is familiar to numerical analysts, e.g., A. Ralston, *A first course in numerical analysis* (McGraw-Hill, New York, 1965), Theorem 10.15, p. 495.