SLAC–PUB–5569
May 1991
(M)

Schrodinger and Dirac Variational Representations with Sparse Matrices: Use of the Sturmian Functions*

CHARLES T. MUNGER

Stanford Linear Accelerator Center
Stanford University, Stanford, California 94309

ABSTRACT

A variational representation for the Dirac Coulomb Hamiltonian can be developed using Sturmian functions where the matrices in the generalized eigenvalue problem $Ax = \lambda Bx$ have a bandwidth of at most 5. Standard matrix algorithms will isolate N eigenvalues and inverse iteration will isolate in sucession N eigenvectors in $\sim N^2$ computer operations and using $\sim N$ locations in memory. A new method is presented of evaluating the matrix elements which connect exact eisenstates and Sturmian variational states by the emission of radiation. The theory of the use of Sturmian functions for the Schrodinger Coulomb Hamiltonian is reviewed and extended.

Submitted to *Physical Review A*

* Work supported by the Department of Energy, contract DE–AC03–76SF00515.

Introduction

The calculation of two-photon decay rates using standard perturbation theory requires one to sum over an infinite number of intermediate atomic states. For hydrogenic ions described by the Schrodinger or Dirac equation the intermediate states are known exactly, but both evaluating the needed matrix elements and doing the final sum are difficult. To calculate rates in hydrogenic ions a powerful computational method is to replace the infinite sum over exact bound and continuum states by a finite sum over a discrete set of states obtained by diagonalizing the Coulomb Hamiltonian in a basis of N functions. One selects the functions so that all necessary radial integrals can be done analytically instead of numerically, and so the matrix eigenvalue problem is the simplest possible. This paper adapts the Sturmian functions of Rotenberg[1] to make the eigenvalue problem for both the Schrodinger and Dirac Coulomb Hamiltonians banded, and develops a simple method of evaluating the radial integrals that describe the emission of radiation of arbitrary multipolarity. This latter method avoids an expansion in powers of the photon wavenumber k and so is especially suited to the study of hydrogenic ions of high nuclear charge, where the wavelength of the emitted radiation is not small compared to the size of the ion. The simplification of the matrix problem and the new method of evaluating integrals make it possible to deal with numbers of virtual intermediate states in the thousands.

Section 1 of this paper defines the Sturmian functions, derives their key properties, and shows how they render the Schrodinger matrix problem sparse. Section 2 applies symmetries of the Schrodinger Coulomb Hamiltonian to construct an infinite set of analytically equivalent representations of any sum over intermediate states. Section 3 presents a novel way of solving for the exact bound states

of the Dirac Coulomb problem. Section 4 reviews and extends Goldman's finite basis set methods as applied to the Dirac Coulomb problem and shows how the Sturmian functions render the matrix problem sparse. Section 5 reviews Grant's description of the matrix elements needed to describe the emission of radiation from Dirac states, and develops a new algorithm for their calculation. Finally Section 6 presents the results of some numerical tests.

1. The Sturmian functions and Schrodinger finite basis sets

To motivate the introduction of the Sturmian functions ξ consider first the eigenstates of the Schrodinger Coulomb Hamiltonian. After the usual separation of the wavefunction Ψ into radial and angular parts,

$$\Psi_l = \frac{1}{r}\psi_l(r)Y_l^m(\theta,\phi) \tag{1.1}$$

the problem to be solved is $H^S\psi = E\psi$ where the Coulomb Hamiltonian is

$$H^S = -\frac{1}{2}\frac{d^2}{dr^2} + \frac{l(l+1)}{2r^2} - \frac{Z}{r} \tag{1.2}$$

Atomic units with $\hbar = m = e = 1$ and $c = \alpha^{-1}$ will be used throughout this paper. The exact normalized bound state wavefunctions for angular momentum l may be written ψ_{lp}, where $p = 1, 2, \ldots$ indexes the wavefunctions in order of increasing energy, and the exact bound state eigenvalues are

$$E_{lp} = -\frac{1}{2}\frac{Z^2}{(l+p)^2}, \quad p = 0, 1, 2, \ldots \tag{1.3}$$

The eigenstate of lowest energy , ψ_{l1}, satisfies the differential equation

$$\left[-\frac{d^2}{dr^2} + \frac{l(l+1)}{r^2} - \frac{2Z}{r} + \frac{Z^2}{(l+1)^2}\right]\psi_{l1} = 0 \tag{1.4}$$

Consider the Sturmian functions $\xi(r)$ and eigenvalues ζ defined by the related

differential equation

$$\left[-\frac{d^2}{dr^2}+\frac{l'(l'+1)}{r^2}-\frac{2Z'}{r}\zeta+\frac{Z'^2}{(l'+1)^2}\right]\xi(r)=0 \tag{1.5}$$

where Z' and l' are numbers (not necessarily integers) greater than zero. The form of this eigenvalue equation differs from those which arise naturally in quantum mechanics in that the eigenvalue ζ appears not multiplying the eigenfunction ξ alone, but multiplying a function times ξ. Such an equation can be used to generate a set of functions complete on the interval $r=(0,\infty)$ by imposing suitable boundary conditions[2]. Require then that the functions ξ_n, like the functions ψ_{lp}, be square-integrable as $r\to 0$ and as $r\to\infty$. If the value of Z' is chosen equal to Z and the value of l' chosen equal to l, evidently one such solution has $\zeta=1$ and $\xi(r)$ proportional to ψ_{l1}. For all $Z'>0$ the complete set of solutions is

$$\begin{aligned}\zeta_n &= \frac{l'+n}{l'+1}\\ \xi_n &= \left[\frac{1}{2Z'}\frac{(n-1)!}{\Gamma(2l'+1+n)}\right]^{+1/2}\left(\frac{2Z'r}{l'+1}\right)^{l'+1}e^{-(Z'r/(l'+1))}L_{n-1}^{2l'+1}\left(\frac{2Z'r}{l'+1}\right)\end{aligned} \tag{1.6}$$

where the index n runs $1,2,3,\ldots$. The symbol $L_n^a(x)$ represents an associated Laguerre polynomial, which is a polynomial of degree n in x whose coefficients depend continuously on the upper index a. There survive in the mathematical literature different definitions of the associated Laguerre polynomials; the definition Goldman[3,4] and I use is[5]

$$L_n^a(x)=\sum_{m=0}^{n}(-1)^m\binom{n+a}{n-m}\frac{x^m}{m!} \tag{1.7}$$

where the symbol in parentheses denotes a binomial coefficient, defined for real x

and integer k by

$$\binom{x}{k} = \begin{cases} \dfrac{\Gamma(x+1)}{k!\Gamma(x-k+1)} & , k \geq 0 \\ 0 & , k < 0 \end{cases} \tag{1.8}$$

The functions ξ_n are normal with respect to the weight function $1/r$ and the normalization in equation 1.6 has been chosen so that

$$\langle \xi_n | 2Z'/r | \xi_{n'} \rangle = \delta_{nn'} \tag{1.9}$$

The functions not however normal in the familiar sense, $\langle \xi_n | 1 | \xi_m \rangle \neq \delta_{nm}$. The Laguerre polynomials have the orthogonality property that[5]

$$\int_0^\infty x^\alpha e^{-x} L_n^\alpha(x) L_m^\alpha(x)\, dx = \begin{cases} 0 & , m \neq n, \text{ Re } \alpha > -1 \\ \dfrac{\Gamma(\alpha+n+1)}{n!} & , m = n, \text{ Re } \alpha > 0 \end{cases} \tag{1.10}$$

which together with the recursion relation[5]

$$xL_n^{\alpha+1}(x) = (n+\alpha+1)L_n^\alpha(x) - (n+1)L_{n+1}^\alpha(x) \tag{1.11}$$

suffices to show that

$$\langle \xi_n | 1 | \xi_m \rangle \equiv T_{nm} = \frac{l'+1}{4Z'^2} \begin{cases} 2(n+l') & , n = m \\ -1[n(n+2l'+1)]^{1/2} & , |n-m| = 1 \ , \ n = \min n, m \\ 0 & , |n-m| > 1 \end{cases} \tag{1.12}$$

Thus the matrix of the overlap of the unorthogonal functions ξ_n is, fortuitously, tridiagonal and simple. Finally the differential equation (1.5) shows that the following matrix is tridiagonal:

$$\left\langle \xi_n \right| - \frac{d^2}{dr^2} + \frac{l'(l'+1)}{r^2} \left| \xi_m \right\rangle = \delta_{nm} - \frac{Z'^2}{(l'+1)^2} T_{nm} \tag{1.13}$$

Thus the matrices of $\langle 1 \rangle$, $\langle 1/r \rangle$, and $\langle -d^2/dr^2 + l'(l'+1)/r^2 \rangle$ in the set of functions $\{\xi\}$ have bandwidth at most 3. It follows immediately that the matri-

ces in the generalized matrix eigenvalue problem $\langle H^S \rangle X = E\langle 1 \rangle X$ also have a bandwidth of at most 3. In fact the matrix of the Coulomb Hamiltonian can be made diagonal. Define the constant $C(Z')$ by $C(Z') = -Z'^2/(2(l+1)^2)$ and consider the matrix eigenvector problem $H'X = \epsilon' TX$, where $H' = H - C(Z')$ and $\epsilon' = E - C(Z')$. This problem has the same eigenvectors X as does $HX = ETX$, but the matrix H' is diagonal:

$$\begin{aligned} H'_{nm} &= \langle \xi_n | H' | \xi_m \rangle \\ &= \frac{1}{2}\langle \xi_n | \left[\left(-\frac{d^2}{dr^2} + \frac{l(l+1)}{r^2} + \frac{Z'^2}{(l+1)^2} \right) - \frac{2Z}{r} \right] | \xi_m \rangle \\ &= \frac{1}{2}\langle \xi_n | \left(\zeta_m - \frac{Z}{Z'} \right) \frac{2Z'}{r} | \xi_m \rangle \\ &= \frac{1}{2}\left(\zeta_m - \frac{Z}{Z'} \right) \delta_{nm}, \qquad n, m = 1, 2, ... \end{aligned} \tag{1.14}$$

Therefore the matrix problem $H'X = \epsilon' TX$ is of the canonical form $AX = \lambda BX$ where A and B are real symmetric band matrices of bandwidth 1 and 3 respectively and B is positive definite. Crawford's implementation[6] of Cholesky factorization, applied to the banded matrix B, transforms in $\sim N^2$ computer operations this generalized matrix problem into the ordinary problem $AX = \lambda X$, where the new matrix A is symmetric and tridiagonal. This can be attacked by better known methods (for example, the QL or QR algorithm)[7,8] , to isolate the N eigenvalues in a net $\sim N^2$ computer operations. With the eigenvalues known the method of inverse iteration[19] may be applied to the original equation $H'X = \epsilon' TX$ to yield each of the N eigenvectors separately. Because the matrices H' and T are sparse the computation of each eigenvector takes $\sim N$ operations instead of the $\sim N^2$ which would be required if either matrix was full. The number of operations required to compute all N eigenvectors is therefore also $\sim N^2$.

The sparsity of H' and T also reduces the amount of computer memory required throughout. Storage of the non-zero entries in both matrices requires only $\sim N$ locations in memory. Crawford's algorithm and, for the resulting tridiagonal problem, the QR and QL algorithms, need only $\sim N$ locations in memory to determine all N eigenvalues. To solve by inverse iteration $H'X = \epsilon' TX$ for one eigenvector also requires for such sparse matrices only $\sim N$ locations in memory. Provided one needs each eigenvector only once, so one can use the same N locations in memory to store in succession the N components of each eigenvector, complete calculations can be done with only $\sim N$ locations in memory, instead of the $\sim N^2$ which would be needed merely to store any full matrix. For the calculation of two-photon decay rates in particular this economy of memory is possible.

The functions ξ_n and their properties are by no means novel. Similar functions first applied to problems in atomic physics by Rotenberg[1], who studied the scattering of positrons from hydrogen. Edmonds[9] in a study of the quadratic Zeeman effect in hydrogen recognised the functions rendered the generalized eigenvalue problem for the Coulomb Hamiltonian sparse. His work was preceeded the publication of the Crawford algorithm, but Clark and Taylor[10] recognised the algorithm's applicability and were able to extend[11] Edmonds' work with the functions. The treatment here is new only in that both Z' and and l' in equation 1.6 are allowed to be non-integers, a generalization essential if they are to be applied to the Dirac Coulomb Hamiltonian.

For special but useful values of Z' the sparse generalized eigenvalue problem $H'X = \epsilon' TX$ reduces still further to a sparse ordinary eigenvalue problem. Provided $Z' > Z$ the elements of H' are all positive and the eigenvalues ϵ' are all greater than zero. Because $H'_{nn} > 0$ one can define a diagonal matrix S, with

elements $S_{nn} = [H'_{nn}]^{-1/2}$, for which $SH'S = 1$. The generalized matrix equation $H'X = \epsilon' TX$ may then be transformed into the ordinary matrix equation $T^{(2)}Y = \mu Y$, where $T^{(2)} = STS$ is real, symmetric and tridiagonal and where the new eigenvalue μ and the new eigenvector Y are trivially related to the old by $\mu = 1/\epsilon'$ and $Y = S^{-1}X$. This transformation doesn't work when Z' exactly equals Z because H'_{11} is then zero and so $S_{11} = [H'_{11}]^{-1/2}$ is undefined. The choice of Z' equal to Z is otherwise a natural one because the ξ_n basis then includes the exact lowest energy eigenfunction ψ_{l1}, which becomes merely a linear combination of ξ_1 and ξ_2. The lowest variational energy considered as a function of Z' is then at its absolute minimum, and so the basis is automatically optimized with respect to variations in the parameter Z'. For $Z' = Z$ the transformation can be salvaged by using the slightly modified basis of functions g_n, in which ψ_{l1} has been decoupled by hand from all but one of the basis functions:

$$\begin{aligned} g_1 &= \psi_{l1} = \frac{Z\sqrt{2}}{l+1}\xi_1 \\ g_2 &= S_{22} \times [\xi_2 - \langle \psi_{l1} | \xi_2 \rangle \psi_{l1}] \\ g_n &= S_{nn} \times \xi_n, \quad n = 2, \ldots, N \end{aligned} \tag{1.15}$$

In this basis the matrices H' and T take the form

$$\begin{aligned} H'^{(g)} &= \begin{pmatrix} 0 & \vec{0}^T \\ \vec{0} & 1 \end{pmatrix} \\ T^{(g)} &= \begin{pmatrix} 1 & \vec{0}^T \\ \vec{0} & \dfrac{(l+1)^2}{2Z^2} T^{(3)} \end{pmatrix} \end{aligned} \tag{1.16}$$

where the matrix $T^{(3)}$ is defined by

$$T^{(3)}_{nm} = \begin{cases} 2l+3 & , n = m = 1 \\ 2\left(1 + \dfrac{l+1}{n}\right) & , n = m,\ n > 1 \\ \sqrt{1 + \dfrac{2(l+1)}{n}} & , |n-m| = 1 \\ 0 & , |n-m| > 1 \end{cases} \tag{1.17}$$

The function ψ_{l1}, proportional to g_1, is a normalized eigenvector with eigenvalue $\epsilon' = 0$. The remaining eigenvectors are linear combinations of g_n for $n > 1$ and the eigenvalues and eigenvectors follow from a solution of $T^{(3)}X = \lambda X$, where $\lambda = 2Z^2/(\epsilon'(l+1)^2)$. This is a numerical problem which is almost trivial. The matrix $T^{(3)}$ is real, symmetric, tridiagonal, and positive definite. Its elements are simple functions. It has convenient property of diagonal dominance, that is, $|T^{(3)}_{nn}| \geq |T^{(3)}_{n-1,n}| + |T^{(3)}_{n+1,n}|$ for all n, and the inequality is strict for at least one value of n (it happens the inequality is strict for all n). The diagonal dominence of $T^{(3)}$ allows for example a matrix equation like $T^{(3)}X = Y$ to be solved for a vector X given a vector Y by Gaussian elimination without pivoting, instead of by the more complicated algorithm of Gaussian elimination with pivoting.[12] Finally the matrix differs only slightly from a simple matrix whose eigenvalues and eigenvectors are known analytically. An $n \times n$ symmetric tridiagonal matrix whose diagonal elements are equal to a and whose off-diagonal elements are equal to b has eigenvalues λ_k given by[13]

$$\lambda_k = a + 2b\cos\left(\frac{k\pi}{n+1}\right) \tag{1.18}$$

and eigenvectors $x^{(k)}$, normalized so $\vec{x}^{(k)} \cdot \vec{x}^{(k)} = 1$, with components given by[13]

$$x^{(k)}_j = \left(\frac{2}{n+1}\right)^{1/2} \sin\left(\frac{kj\pi}{n+1}\right) \tag{1.19}$$

where $j,k = 1,2,\ldots,N$. With the choice $a = 2$ and $b = -1$ the elements of the simple matrix differ from the elements of $T^{(3)}$ by the first term on the $T^{(3)}$ diagonal, $2l+3$, and elsewhere only by terms of leading order $\sim 1/n$ for large n. The existence of this nearby matrix with known eigenvalues and eigenvectors both aids the debugging of computer code which solves for the eigenvalue and eigenvectors of $T^{(3)}$, and gives useful analytic estimates of the magnitude of both the eigenvalues and the components of the eigenvectors.

2. Symmetries of the Schrodinger Coulomb Hamiltonian

There are symmetries buried in the Schrodinger Hamiltonian, equation 1.2, which both have applications in practical problems and which which are related to the more complicated symmetries of the Dirac Coulomb Hamiltonian, which will be treated in section 3. Consider the Schrodinger Coulomb Hamiltonian in equation 1.2 as a operator which depends on l:

$$H(l) = -\frac{d^2}{dr^2} + \frac{l(l+1)}{r^2} - \frac{2Z}{r} \tag{2.1}$$

One can solve for the eigenvalues and eigenvectors of this Hamiltonian by using a system of raising and lowering operators[14], a method which will now be reviewed. Consider the operator $J(l)$ defined by

$$J(l) = -Z + \frac{l^2}{r} + l\frac{d}{dr} \quad , l > 0 \tag{2.2}$$

Simple algebra shows that

$$J(l)H(l) = \begin{cases} H(l-1)J(l) & l > 0 \\ H(l)J(l) & l = 0 \end{cases} \tag{2.3}$$

Therefore if ψ_l is an eigenstate of $H(l)$ with eigenvalue λ, then $J(l)\psi_l$ is an eigenstate of $H(l-1)$ with the same value of λ, provided $l > 0$. Thus $J(l)$ for $l > 0$

is a lowering operator with respect to l. It is convenient to rescale by defining an operator $q(l)$ by

$$q(l) = \frac{1}{l}J(l) = -\frac{Z}{l} + \frac{l}{r} + \frac{d}{dr}, \qquad l > 0 \tag{2.4}$$

Consider the hermitian conjugate of $q(l)$ defined by

$$q^\dagger(l) = -\frac{Z}{l} + \frac{l}{r} - \frac{d}{dr} \qquad , l > 0 \tag{2.5}$$

The operator $q^\dagger$ satisfies the commutation relation

$$q^\dagger(l+1)H(l) = H(l+1)q^\dagger(l), \quad l \geq 0 \tag{2.6}$$

so $q^\dagger(l)$ is a raising operator. Both $q^\dagger$ and q respectively raise and lower the angular momentum of both bound and continuum eigenstates. The Hamiltonian $H(l)$ can be rewritten in terms of these operators in the two forms

$$\begin{cases} q^\dagger(l)q(l) = H(l) + Z^2/l^2 & , l \geq 1 \\ q(l)q^\dagger(l) = H(l-1) + Z^2/l^2 & , l \geq 1 \end{cases} \tag{2.7}$$

These operators q and $q^\dagger$ form an efficient way to construct the exact solutions of the problem, since the equation $H(l)\Psi = \lambda\Psi$ may be rewritten for all l as

$$q(l+1)q^\dagger(l+1)\psi = \left(\lambda + \frac{Z^2}{(l+1)^2}\right)\psi \tag{2.8}$$

One can conclude that any function satisfying $q^\dagger(l+1)\psi = 0$ has eigenvalue $\lambda = -Z^2/(l+1)^2$. Any such function cannot be raised in l further. The normalized solutions to the simple first order differential equation $q^\dagger(l+1)\psi = 0$ are

$$\psi_l = \left[\left(\frac{l+1}{2Z}\right)^{2l+3}\Gamma(2l+3)\right]^{-1/2} r^{l+1}e^{-Zr/(l+1)} \tag{2.9}$$

These are of course the yrast $1s$, $2p$, $3d$, ... bound states. The other bound states states can be found by repeatedly lowering l, down of course to $l = 0$ which cannot

be lowered further. There are $l+1$ such states. If $H(L)\psi_L = \lambda\psi_L$ and $\langle\psi_L|\psi_L\rangle = 1$, then $H(L-1)\psi_{L-1} = \lambda\psi_{L-1}$ and $\langle\psi_{L-1}|\psi_{L-1}\rangle = 1$, where

$$\psi_{L-1} = \left[\lambda + \frac{Z^2}{L^2}\right]^{-1/2} q(L)\psi_L \tag{2.10}$$

One can of course step up in l as well as down. If $H(L)\psi_L = \lambda\psi_L$ and $\langle\psi_L|\psi_L\rangle = 1$, then $H(L+1)\psi_{L+1} = \lambda\psi_{L+1}$ and $\langle\psi_{L+1}|\psi_{L+1}\rangle = 1$, where

$$\psi_{L+1} = \left[\lambda + \frac{Z^2}{(L+1)^2}\right]^{-1/2} q^\dagger(L+1)\psi_L \tag{2.11}$$

Equations 2.10 and 2.11 apply not only to bound states, which have $\lambda < 0$ and for which the raising operation in 2.11 terminates at some maximum value of l for each bound state energy, but also to continuum states which have $\lambda > 0$ and for which l may be raised indefinitely. The extra factors involving λ in equations 2.10 and 2.11 then preserve the appropriate normalization of the continuum states, that of one electron confined to a spherical volume of large radius.

Thus the operators q and $q^\dagger$ may be used to find quite elegantly the eigenstates of the Schrodinger Coulomb Hamiltonian. They also have practical value. A sum over a set of intermediate radial wavefunctions corresponding to angular momentum l may be transformed, by repeatedly raising and lowering l with q and $q^\dagger$, into an infinity of different sums over the radial wavefunctions corresponding to different angular momenta. These new sums like the original may be evaluated using discrete states and discrete eigenenergies found for the artificial new angular momentum. A sum like

$$\sum_n f(E_{1n})\langle\psi_{01}|\hat{O}_1\psi_{1n}\rangle\langle\psi_{02}|\hat{O}_2\psi_{1n}\rangle \tag{2.12}$$

where $\hat{O}_1$ and $\hat{O}_2$ are operators connecting exact $1s$ and $2s$ states to virtual p states

with energies E_{1n}, and where f is some function of the energies of the intermediate states, has for example the analytically equivalent forms

$$\begin{aligned} &= \sum_{n\neq 1} \frac{f(E_{0n})}{E_{0n}+Z^2} \langle \psi_{0n}|q(1)\hat{O}_1\psi_{01}\rangle\langle\psi_{0n}|q(1)\hat{O}_2\psi_{02}\rangle, \\ &= f(E_{11})\langle\psi_{11}|\hat{O}_1|\psi_{01}\rangle\langle\psi_{11}|\hat{O}_2|\psi_{02}\rangle \\ &\quad + \frac{f(E_{2n})}{E_{2n}+Z^2/4}\sum_n \langle\psi_{2n}|q^\dagger(2)\hat{O}_1\psi_{01}\rangle\langle\psi_{2n}|q^\dagger(2)\hat{O}_2\psi_{02}\rangle \end{aligned} \tag{2.13}$$

When variational states with the appropriate l are substituted into two of the infinity of such sums one can hope that one will converge as $N \rightarrow \infty$ from above and one from from below, allowing one bracket the true value.

3. Symmetries of the Dirac Coulomb Hamiltonian

Consider next the extension of these ideas from the Schrodinger to the Dirac Coulomb Hamiltonian. This generalization relies heavily on the seminal work of Goldman and of Drake[15], which is now briefly reviewed. The Dirac equation for a Coulomb potential can be written as $H^D\Psi = E\Psi$, where

$$H^D = \vec{\alpha}\cdot\vec{p} + \beta - \frac{Z\alpha^2}{r} \tag{3.1}$$

where $\vec{\alpha}$ and β are the usual 4×4 Dirac matrices. The solutions may be written

$$\psi = \begin{pmatrix} i\dfrac{g(r)}{r}\Omega_{jlM} \\ -\dfrac{f(r)}{r}\Omega_{j\tilde{l}M} \end{pmatrix}, \quad \tilde{l} = 2j-1 \tag{3.2}$$

where $g(r)$ and $f(r)$ are the large and small radial wavefunctions and Ω_{jlM} is a two-component spherical spinor. The large and small components satisfy the coupled

equations

$$\begin{pmatrix} \left(\frac{1}{\alpha} - \frac{\alpha Z}{r}\right) & \left(\frac{\kappa}{r} - \frac{d}{dr}\right) \\ \left(\frac{\kappa}{r} + \frac{d}{dr}\right) & -\left(\frac{1}{\alpha} + \frac{\alpha Z}{r}\right) \end{pmatrix} \begin{pmatrix} g(r) \\ f(r) \end{pmatrix} = \alpha E \begin{pmatrix} g(r) \\ f(r) \end{pmatrix} \tag{3.3}$$

where κ is the Dirac quantum number, $\kappa = \pm(j + \frac{1}{2})$ for $l = j \pm \frac{1}{2}$. Goldman[3] introduced the following unitary transformation of the functions f and g which defines a new pair of functions ϕ and θ:

$$\begin{pmatrix} \phi(r) \\ \theta(r) \end{pmatrix} = \frac{1}{\sqrt{1 + q(\kappa)^2}} \begin{pmatrix} q(\kappa) & 1 \\ 1 & -q(\kappa) \end{pmatrix} \begin{pmatrix} g(r) \\ f(r) \end{pmatrix} \tag{3.4}$$

where the parameter $q(\kappa)$ is defined by $q(\kappa) = Z\alpha/(\kappa + \gamma)$, and where γ is defined by $\gamma = \sqrt{\kappa^2 - (Z\alpha)^2}$. The eigenvalue equation transforms to

$$\begin{pmatrix} \eta(\kappa) & \left(\frac{\alpha Z}{\kappa} + \frac{\alpha\gamma}{r} + \alpha\frac{d}{dr}\right) \\ \left(\frac{\alpha Z}{\kappa} + \frac{\alpha\gamma}{r} - \alpha\frac{d}{dr}\right) & -\eta(\kappa) - 2\frac{\alpha^2 Z}{r} \end{pmatrix} \begin{pmatrix} \phi \\ \theta \end{pmatrix} = \epsilon \begin{pmatrix} \phi \\ \theta \end{pmatrix} \tag{3.5}$$

where the constant $\eta(\kappa)$ is defined by $\eta = -\gamma/\kappa$ and the new eigenvalue ϵ is related to the old by $\epsilon = \alpha^2 E$.

One might hit upon the unitary tranformation in equation 3.4 as the only one which removes all dependence on r from the term in the upper left corner of this matrix. The use of the functions ϕ and θ instead of the standard functions g and f allows one to solve separately for the bound states with a common $|\kappa|$ which occur in energy-degenerate pairs (*e.g.*, the $2s_{1/2}$ and $2p_{1/2}$ states) and the states which are isolated in energy (*e.g.*, the $1s_{1/2}$). Furthermore one solves for one function θ which is common to each degenerate pair, instead of two. The proof runs as follows. It is impossible to find a solution with the function ϕ identically zero and

the function $\theta \neq 0$. The function θ can be zero if and only if the eigenvalue ϵ is equal to η and ϕ satisfies the first order differential equation

$$\left(\frac{Z}{\kappa} + \frac{\gamma}{r} - \frac{d}{dr}\right)\phi(r) = 0 \tag{3.6}$$

For $\kappa > 0$ this equation has no normalizable solution. For $\kappa < 0$ there is the one normalized solution

$$\phi(r) = \left[\left(\frac{2Z}{|\kappa|}\right)^{-(2\gamma+1)}\Gamma(2\gamma+1)\right]^{-1/2} r^{\gamma}e^{-Zr/|\kappa|} \tag{3.7}$$

which is the one bound state of given $|\kappa|$ has no accompanying degenerate state. These are the $1s_{1/2}$, $2p_{3/2}$, $3d_{5/2}$, ..., states. For all other states of given $|\kappa|$ one has $\epsilon \neq \eta$ and $\theta \neq 0$ so one may solve the top line of equation 3.5 for ϕ in terms of θ,

$$\phi = \frac{-1}{\eta(\kappa) - \epsilon}\left(\frac{Z}{\kappa} + \frac{\gamma}{r} + \frac{d}{dr}\right)\alpha\theta \tag{3.8}$$

and eliminate ϕ in the second line. One finds that the function θ satisfies the differential equation

$$\left[-\frac{d^2}{dr^2} + \frac{\gamma(\gamma+1)}{r^2} + \frac{1-\epsilon^2}{\alpha^2} - \frac{2Z\epsilon}{r}\right]\theta = 0 \tag{3.9}$$

This, remarkably, has no dependence on the sign of κ. For given $|\kappa|$, then, except for the one already excluded state, the solutions with opposite signs for κ must have the same energy ϵ and, except for a multiplicative constant, the same function θ for a lower component.

One can go further and solve equation 3.9 for the bound states, though it is not a standard eigenvalue equation (note the eigenvalue ϵ appears squared). In the

process one finds the Dirac and Schrodinger exact solutions are curiously linked. Consider the differential equation

$$\left[-\frac{d^2}{dr^2}+\frac{\gamma(\gamma+1)}{r^2}-2E_m-\frac{2Z'}{r}\right]\Psi_m=0 \tag{3.10}$$

The differential operator appearing in equation 3.10 looks like the Schrodinger Coulomb Hamiltonian, equation 2.1, except the number playing the role of the angular momentum, γ, is not an integer, and neither necessarily is the constant Z' playing the role of the nuclear charge. If one reviews the solving of equation 2.1 for bound states by means of the operators q and $q^\dagger$, equations 2.2 through 2.11, one finds it carries through if, instead of being only a positive integer, l is allowed to be any positive number. For bound states the eigenvalues E_m of equation 3.10 are therefore given by

$$E_m=-\frac{Z'^2}{2(\gamma+m)^2}, \qquad m=1,2,\ldots \tag{3.11}$$

There is a solution, square integrable at infinity, ϵ, θ to equation 3.9 if and only if there is a solution, square integrable at infinity, E_m,Ψ_m to equation 3.10, with a correspondence between constants given by

$$\frac{1-\epsilon^2}{\alpha^2}=-2E_m$$

$$Z\epsilon=Z' \tag{3.12}$$

Solving for ϵ and using $E=\epsilon/\alpha^2$ yields

$$\frac{E}{m_e c^2}=\left[1+\frac{(Z\alpha)^2}{\left(m+\sqrt{\kappa^2-(Z\alpha)^2}\right)^2}\right]^{-1/2}, \qquad m=1,2,\ldots \tag{3.13}$$

This is the Sommerfeld formula. The derivation given excludes the lowest energy

state for $\kappa < 0$, whose eigenvalue was shown earlier to be $\epsilon = \eta(\kappa)$. However η/α^2 happens to equal the righthand side of equation 3.13 if m is assigned the value 0. This completes the derivation of the bound state energies.

4. Dirac Finite Basis Sets

The unitary transformation 3.4 certainly eases the finding of the exact eigenstates of the Dirac Coulomb Hamiltonian. It also eases finding of variational eigenstates by rendering the matrix eigenvalue equation sparse. Equation 3.5 may be written in Hamiltonian form as $h\Phi = \epsilon\Phi$ where Φ is the vector $\big(\phi(r), \theta(r)\big)$ and

$$h = \begin{pmatrix} \eta(\kappa) & \alpha\left(\frac{Z}{\kappa} + \frac{\gamma}{r} + \frac{d}{dr}\right) \\ \alpha\left(\frac{Z}{\kappa} + \frac{\gamma}{r} - \frac{d}{dr}\right) & -\eta(\kappa) - \frac{2Z\alpha^2}{r} \end{pmatrix} \tag{4.1}$$

It is convenient to define the four operators

$$\begin{aligned} B(\kappa) &= \alpha\left(\frac{Z}{\kappa} + \frac{\gamma}{r} + \frac{d}{dr}\right) \\ B^\dagger(\kappa) &= \alpha\left(\frac{Z}{\kappa} + \frac{\gamma}{r} - \frac{d}{dr}\right) \\ v &= \frac{Z\alpha^2}{r} \\ A &= -\eta(\kappa) - 2v \end{aligned} \tag{4.2}$$

whereupon the operator $h(\kappa)$ may be simply written

$$h = \begin{pmatrix} \eta & B \\ B^\dagger & A \end{pmatrix} \tag{4.3}$$

One now has a pair of constructions introduced by Goldman[3]:

Construction 1. Choose N distinct functions $\{w\}$ to expand θ and $M \geq N+1$ distinct functions $\{u\}$ to expand ϕ, such that $\{B(\kappa)w\} \in \{u\}$ and $\{B(-\kappa)w\} \in \{u\}$, and so that if a function f is any linear combination of the functions $\{w\}$, then $B(\kappa)f \neq 0$ and $B(-\kappa)f \neq 0$. Then diagonalizing h in the basis

$$\begin{cases} \begin{pmatrix} u_j \\ 0 \end{pmatrix} & j = 1, \ldots, M \\ \begin{pmatrix} 0 \\ w_j \end{pmatrix} & j = M+1, \ldots, M+N+1 \end{cases} \tag{4.4}$$

produces $2N$ legitimate eigenvectors with distinct eigenvalues $\epsilon \neq \eta$. The eigenstates satisfy a differential equation analogous to equation 3.8,

$$\phi = \frac{-1}{\eta(\kappa) - \epsilon} B(\kappa)\theta \ ,$$

and consequently the N functions $\{Bw\} \in \{u\}$ suffice to expand the upper components ϕ. The $M-N$ extra functions in $\{u\}$ produce spurious eigenvectors with eigenvalues $\epsilon = \eta(\kappa)$ and lower components $\theta = 0$ and are discarded.[16]

Construction 2. Choose N distinct functions $\{w\}$ such that for each, $B(\kappa)w \neq 0$ and $B(-\kappa)w \neq 0$. Then diagonalizing h in the basis

$$Q_j = \begin{cases} \begin{pmatrix} B(\kappa)w_j \\ 0 \end{pmatrix} & j = 1, \ldots, N \\ \begin{pmatrix} 0 \\ w_j \end{pmatrix} & j = N+1, \ldots, 2N \end{cases} \tag{4.5}$$

produces $2N$ legitimate eigenvectors with eigenvalues $\epsilon \neq \eta(\kappa)$. These eigenvalues and eigenvectors are identical to those of Construction 1 for a common choice of the set $\{w\}$.

The constructions have the following additional properties[15,3]: (1) the legitimate eigenvectors, together if $\kappa < 0$ with the lone exact state (equation 3.7), form a suitable variational basis which is complete in the limit $N \to \infty$ if the set of functions $\{w\}$ then spans the lower functions θ of the exact solutions[17]; (2) half the legitimate eigenvalues are positive and half negative; (3) the legitimate eigenvalues for the constructions with $\kappa > 0$ and $\kappa < 0$ are degenerate, and pairs with degenerate eigenvalues have, up to a constant factor, same lower component θ:

$$\theta_{-\kappa}(r) = \left(\frac{\epsilon + \eta(\kappa)}{\epsilon - \eta(\kappa)}\right)^{1/2} \theta_{\kappa}(r) \tag{4.6}$$

The first construction offers the convenience that a proper choice of the functions u_j and w_j renders the basis 4.4 orthogonal, and so makes it possible to solve an ordinary instead of a generalized matrix eigenvalue problem. A suitable set of functions involving the Laguerre polynomials was used by Goldman[4]:

$$\begin{cases} u_j(r) = \left[\dfrac{2\beta\ (j-1)!}{\Gamma(2\gamma + j)}\right]^{1/2} (2\beta r)^{\gamma} e^{-\beta r} L_{j-1}^{2\gamma}(2\beta r) & j = 1, \ldots, N+1 \\ w_j(r) = \left[\dfrac{2\beta\ (j-1)!}{\Gamma(2\gamma + j + 2)}\right]^{1/2} (2\beta r)^{\gamma+1} e^{-\beta r} L_{j-1}^{2\gamma+2}(2\beta r) & j = 1, \ldots, N \end{cases} \tag{4.7}$$

where β is an adjustable parameter. The generalized eigenvalue problem $hX = \epsilon QX$ reduces to the ordinary one $hX = \epsilon X$. The matrix of the Hamiltonian is full, however, so the number of computer operations required to isolate the eigenvalues or eigenvectors remains $\sim N^3$. Consider however the use of the second construction, and examine the matrix of h in an (unorthogonal) basis, equation 4.5. The the $2N \times 2N$ matrix equation $hX = \epsilon QX$ may be written in simple $N \times N$

blocks:

$$\begin{pmatrix} \eta(\kappa)M & M \\ M & A \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \epsilon \begin{pmatrix} M & 0 \\ 0 & U \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \tag{4.8}$$

where

$$\begin{aligned} M_{ij} &= \langle w_i | B^\dagger B | w_j \rangle \\ A_{ij} &= \langle w_i | A | w_j \rangle \\ U_{ij} &= \langle w_i | 1 | w_j \rangle \end{aligned} \tag{4.9}$$

where I have used the notation $\langle f_1 | \hat{o} | f_2 \rangle = \int_0^\infty f_1 (\hat{o} f_2)\, dr$. The operator $B^\dagger B$ is equal to

$$B^\dagger B = \alpha^2 \left[-\frac{d^2}{dr^2} + \frac{\gamma(\gamma+1)}{r^2} + \frac{2Z}{\kappa}\frac{\gamma}{r} + \frac{Z^2}{\kappa^2} \right] \tag{4.10}$$

Consider the Sturmian problem

$$\left[-\frac{d^2}{dr^2} + \frac{l'(l'+1)}{r^2} - \frac{2Z'}{r}\zeta + \frac{Z'^2}{(l'+1)^2} \right] \xi_{l'}(r) = 0 \tag{4.11}$$

whose solutions eigenfunctions and eigenvalues are given by equation 1.6. It is an immediate consequence of the fact that the matrices of $\langle 1/r \rangle$, $\langle 1 \rangle$ and $\langle -d^2/dr^2 + l'(l'+1)/r^2 \rangle$ in the ξ basis are at most tridiagonal that the matrices A, M, and U will be at worst tridiagonal if the parameter l' is chosen equal to γ. The functions ξ and $B\xi$ depend as $r \to 0$ as $r^{\gamma+1}$ and r^γ, respectively, as do the exact eigenstates of the Dirac Coulomb Hamiltonian. By reordering the rows and columns of the matrix eigenvalue problem from the order $Bw_1, \ldots, Bw_N, w_1, \ldots, w_N$ to $Bw_1, w_1, \ldots, Bw_N, w_N$ one sees that the generalized eigenvalue problem $AX = \lambda BX$ has a bandwidth of at most 9.

For a special choice of the parameter Z' in the ξ basis one can simplify the matrix eigenvalue problem further. Choose the constants

$$l' = \gamma$$

$$Z' = \frac{1+\gamma}{|\kappa|} Z \equiv Z'_0 \tag{4.12}$$

which makes two terms in the differential operator defining the functions ξ match two terms appearing in the differential operator $B^\dagger B$. This choice of Z' incidently makes the constant β in the exponential $\sim e^{-\beta r}$ term of the basis match that of the lone state for $\kappa < 0$. One then finds that the matrix of M is diagonal:

$$\begin{aligned} M_{ij} &= \langle \xi_i | B^\dagger B | \xi_j \rangle \\ &= \langle \xi_i | \alpha^2 \left(\left[-\frac{d^2}{dr^2} + \frac{\gamma(\gamma+1)}{r^2} + \frac{Z^2}{\kappa^2} \right] + \frac{2Z}{\kappa}\frac{\gamma}{r} \right) | \xi_j \rangle \\ &= \alpha^2 \langle \xi_i | \frac{2Z'}{r}\zeta_j + \frac{2Z}{r}\frac{\gamma}{\kappa} | \xi_j \rangle \\ &= \alpha^2 \left(\zeta_j + \frac{\kappa}{|\kappa|}\frac{\gamma}{\gamma+1} \right) \langle \xi_i | \frac{2Z'}{r} | \xi_j \rangle \\ &= \alpha^2 \left(\zeta_j + \frac{\kappa}{|\kappa|}\frac{\gamma}{\gamma+1} \right) \delta_{ij} \end{aligned} \tag{4.13}$$

Note $M_{jj} > 0$ for all j and κ. Thus it is possible to define a scaled basis

$$g_n(Z'_0) = \alpha^{-1} \left[\zeta_n + \frac{\kappa}{|\kappa|}\frac{\gamma}{1+\gamma} \right]^{-1/2} \xi_n \tag{4.14}$$

so that M will reduce to the unit matrix. As for the Schrodinger problem the maximum simplification is achieved in a matrix eigenvector problem where the eigenvalues have been shifted. Solve then $h'\Phi = \epsilon'\Phi$, with $h' = h + \eta$ and $\epsilon' = \epsilon + \eta$.

The form of the operator h' is

$$h' = \begin{pmatrix} 2\eta & B \\ B^\dagger & 2v \end{pmatrix} \tag{4.15}$$

and the matrix of h' using the functions g has the block form

$$h' = \begin{pmatrix} 2\eta & 1 \\ 1 & -D(Z_0') \end{pmatrix} \tag{4.16}$$

where $D(Z_0')$ is a diagonal matrix with positive non-zero entries:

$$\begin{aligned} D_{ij}(Z_0') &= \langle g_i | 2v | g_j \rangle \\ &= \left[\zeta_i + \frac{\kappa}{|\kappa|}\frac{\gamma}{\gamma+1}\right]^{-1/2} \left[\zeta_j + \frac{\kappa}{|\kappa|}\frac{\gamma}{\gamma+1}\right]^{-1/2} \frac{Z}{Z'} \langle \xi_i | 2Z'/r | \xi_j \rangle \\ &= \left[\zeta_j + \frac{\kappa}{|\kappa|}\frac{\gamma}{\gamma+1}\right]^{-1} \left(\frac{|\kappa|}{\gamma+1}\right) \delta_{ij}, \qquad i,j = 0,1,2,\ldots \end{aligned} \tag{4.17}$$

Without the shift of eigenvalue the matrix appearing in the place of $D(Z_0')$ would be tridiagonal. The matrix of the overlap of the basis states, Q, where $Q_{ij} = \langle Q_i | Q_j \rangle$, is also simple, having the block form

$$Q = \begin{pmatrix} 1 & 0 \\ 0 & T(Z_0') \end{pmatrix} \tag{4.18}$$

where $T(Z_0')$ is the $N \times N$ symmetric, positive definite, tridiagonal matrix given by

$$T_{nm}(Z_0') = \left(\frac{\kappa}{2Z\alpha}\right)^2 \times$$

$$\begin{cases} 2(n+\gamma)\left[n+(s+1)\gamma\right]^{-1} & , n = m \\ -1\left[n(n+2\gamma+1)\right]^{1/2} \times & \\ \qquad \left[(n+(s+1)\gamma)(n+1+(s+1)\gamma)\right]^{-1/2} & , |n-m| = 1 \ , \ n = \min(n,m) \\ 0 & , |n-m| > 1 \end{cases} \tag{4.19}$$

where $s = \kappa/|\kappa|$. The band matrices appearing in the generalized eigenvalue problem now have a maximum bandwidth of 5. As for the Schrodinger problem

the eigenvalues may be found by applying in succession the Crawford, Schwartz,[18] and the QR or QL algorithm and using $\sim N$ locations in memory. The QL or QR algorithm will solve for the eigenvectors in $\sim N^3$ time and using $\sim N^2$ memory; it is better therefore to find the eigenvectors by inverse iteration[19], which requires $\sim N^2$ time and $\sim N$ locations in memory. A sparse matrix problem suitable for solution by inverse iteration may be constructed by the following argument for all Z', instead of for just the special value given by equation 4.10.

The original matrix problem, equation 4.8, for an arbitrary set of functions $\{w\}$ is equivalent to the following pair of matrix equations for the vectors x_1 and x_2:

$$\begin{aligned} \eta M x_1 + M x_2 &= \epsilon M x_1 \\ M x_1 + A x_2 &= \epsilon U x_2 \end{aligned} \tag{4.20}$$

Because $M_{ij} = \langle Bw_i | Bw_j \rangle$ is the overlap matrix of N independent functions Bw, the matrix M is positive definite. Therefore $MX \neq 0$ for any vector X. At least one of the two vectors x_1, x_2 for a solution must be non-zero. The second equation in 4.20 shows then that $x_2 \neq 0$ and the then the first that $x_1 \neq 0$ and $\epsilon \neq \eta$. From the first equation in 4.20 one then conclude that x_1 and x_2 are related by

$$x_1 = \left(\frac{-1}{\eta - \epsilon} \right) x_2 \tag{4.21}$$

Therefore one knows an entire two-component eigenvector if one knows only its eigenvalue and its (suitably normalized) lower component x_2. Solve then only for x_2 by using this equation to eliminate x_1 in equation 4.20. One finds that x_2

satisfies

$$\left[\epsilon^2 \langle 1 \rangle_w + \epsilon \left\langle \frac{2Z\alpha^2}{r} \right\rangle_w + \left\langle \alpha^2 \frac{d^2}{dr^2} - \alpha^2 \frac{\gamma(\gamma+1)}{r^2} - 1 \right\rangle_w \right] x_2 = 0 \qquad (4.22)$$

where $\langle \hat{o} \rangle_w$ has been assigned the meaning, 'the matrix of the operator $\hat{o}$ in the basis of functions w". Note $\langle 1 \rangle_w \neq 1$ if the set of functions w is not orthogonal. All dependence on the sign of κ has disppeared in equation 4.22, which is not surprising because it is merely the matrix version of our earlier differential identity, equation 3.9. One sees that the variational solutions with opposite signs of κ, like the exact solutions, are degenerate in energy and have the same functions as lower components. This conclusion was first reached by Goldman[3,20]

The $2N \times 2N$ standard matrix equation 4.8, which is linear in ϵ, has been deflated into the $N \times N$ matrix equation 4.22, which is nonlinear in ϵ. A series of tricks will, using the ξ_n basis and for a range of Z', reinflate 4.22 into a sparse standard matrix equation which will yield the desired eigenvalues ϵ and lower components x_2. Equation 4.22 has the form $[A\epsilon^2 + B\epsilon + C]x = 0$. Shift the eigenvalue to $\delta = \epsilon - \epsilon_0$, where ϵ_0 is a constant to be determined. There results the new equation

$$[A\delta^2 + (2\epsilon_0 A + B)\delta + (A\epsilon_0^2 + B\epsilon_0 + C)]X = 0 \qquad (4.23)$$

which is of the same form with the substitutions

$$\begin{aligned} \delta &\rightarrow \epsilon \\ A' &\rightarrow A \\ B' &\rightarrow 2\epsilon_0 A + B \\ C' &\rightarrow A\epsilon_2^2 + B\epsilon_0 + C \end{aligned} \qquad (4.24)$$

The form of C' is

$$C' = -\alpha^2 \left\langle -\frac{d^2}{dr^2} + \frac{\gamma(\gamma+1)}{r^2} - \frac{2Z\epsilon_0}{r} + \left(\frac{1-\epsilon_0^2}{\alpha^2}\right) \right\rangle_w \tag{4.25}$$

Comparing this to equation 1.5 one sees that the matrix C' will be diagonal in the ξ_n basis if ϵ_0 satisfies

$$\left(\frac{1-\epsilon_0^2}{\alpha^2}\right) = \frac{Z'^2}{(\gamma+1)^2} \tag{4.26}$$

Solving for ϵ_0, one will require

$$\epsilon_0(Z') = \sigma\sqrt{1-\left(\frac{Z'\alpha}{\gamma+1}\right)^2} \tag{4.27}$$

where $\sigma = \pm 1$ is an arbitrary sign. Because the shift ϵ_0 must be real C' can be made diagonal only if $Z'\alpha \leq \gamma + 1$. In the ξ basis one finds the matrices

$$\begin{aligned} C'_{ij} &= -\alpha^2 \left(\zeta_j - \frac{Z}{Z'}\epsilon_0\right)\delta_{ij} \\ B'_{ij} &= \frac{Z}{Z'}\alpha^2\delta_{ij} + 2\epsilon_0 T_{ij} \\ A'_{ij} &= T_{ij} \end{aligned} \tag{4.28}$$

Its convenient to define a diagonal matrix a with diagonal elements $a(n)$ given by

$$a(n) = a_{nn} = \zeta_n - \frac{Z}{Z'}\epsilon_0 \tag{4.29}$$

Now any matrix equation $[A\epsilon^2 + B\epsilon + C]X = 0$ may be rewritten if $\epsilon \neq 0$ as $C(1/\epsilon)^2 + B(1/\epsilon) + A]x = 0$. If det(C) $\neq 0$ a unique dummy vector y may be defined for each x by $Cy = (1/\epsilon)Cx$. Then solving $[A\epsilon^2 + B\epsilon + C]X = 0$ is

equivalent to solving the ordinary eigenvector problem

$$\begin{bmatrix} 0 & -C \\ -C & -B \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix} = \epsilon \begin{bmatrix} -C & 0 \\ 0 & +A \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix} \tag{4.30}$$

Apply this formula to the shifted version of the Dirac problem, 4.23, and its matrices A', B' and C'. The conditions so that $\delta \neq 0$ and $\det(C') \neq 0$ are discussed later. One obtains

$$\alpha^2 \begin{bmatrix} 2a\epsilon_0 & a \\ a & -Z/Z' \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix} = (\epsilon + \epsilon_0) \begin{bmatrix} a & 0 \\ 0 & T \end{bmatrix} \begin{bmatrix} y \\ x_2^{(\xi)} \end{bmatrix} \tag{4.31}$$

where the superscript on $x_2^{(\xi)}$ indicates the column vector x_2 is in the ξ basis. The matrix T (equation 1.12 with $l' \to Z'$) is positive definite. The $2N \times 2N$ matrix on the righthandside of this equation will therefore be positive definite if the matrix a is. If the $2N \times 2N$ matrix is positive definite then equation 4.31 can be solved for its eigenvalues by the applying in succession the Crawford, Schwartz, and QR or QL algorithms in $\sim N^2$ computer operations. The requirement that a be positive definite translates into the following bounds on Z':

$$\begin{cases} 0 < Z' < (1+\gamma)/\alpha & , \epsilon_0 < 0 \\ \dfrac{Z(1+\gamma)}{\sqrt{(1+\gamma)^2 + (Z\alpha)^2}} < Z' < (1+\gamma)/\alpha & , \epsilon_0 > 0 \end{cases} \tag{4.32}$$

These bounds I henceforth adopt. They are sufficient (but not perhaps necessary) for the construction 4.30 to be valid.

Because one has chosen $a(n) > 0$ one can define a rescaled basis

$$g_n(Z') = \alpha^{-1} a(n)^{-1/2} \xi_n(Z'), \qquad n = 1, 2, \ldots \tag{4.33}$$

wherein the matrix equation takes its simplest form:

$$\begin{bmatrix} +2\epsilon_0 & 1 \\ 1 & -D(Z') \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix} = (\epsilon + \epsilon_0) \begin{bmatrix} 1 & 0 \\ 0 & T(Z') \end{bmatrix} \begin{bmatrix} y \\ x_2^{(g)} \end{bmatrix} \tag{4.34}$$

where $D(Z')$ is a diagonal matrix with positive non-zero entries and and $T(Z')$ is tridiagonal:

$$\begin{aligned} D_{nm}(Z') &= \frac{Z}{Z'}[a(n)]^{-1}\delta_{nm} \\ T_{nm}(Z') &= (\gamma+1)\left(\frac{1}{2Z'\alpha}\right)^2 \times \\ &\begin{cases} 2(n+\gamma)[a(n)]^{-1} & , n = m \\ -1[n(n+2\gamma+1)]^{+1/2}[a(n)a(n+1)]^{-1/2} & , |n-m| = 1,\ n = \min(n,m) \\ 0 & , |n-m| > 1 \end{cases} \end{aligned} \tag{4.35}$$

Remarkably, when one chooses Z' equal to its earlier magic value, Z'_0, and sets the arbitrary sign σ in the definition of ϵ_0 to $-\kappa/|\kappa|$, one finds that $D(Z') \to D(Z'_0)$, $T(Z') \to T(Z'_0)$, $\epsilon_0 \to \eta$, and $g_n(Z') \to g_n$, reproducing precisely the original sparse matrix equations in 4.17 and 4.18. Solving 4.35 for eigenvectors by inverse iteration is very fast because one need solve per iteration only one $N \times N$, tridiagonal matrix equation $AX = Y$ for a vector X in terms of a vector Y.

Having by inverse iteration solved 4.34 for an eigenvalue ϵ and, up to a constant factor, a lower component vector $x_2^{(g)}$, one is left with the task of finding a normalized lower vector in the ξ basis. One chooses a sign of κ, calculates the vector $[x_2^{(\xi)}]_n = \alpha^{-1}a(n)^{-1/2}[x_2^{(g)}]_n$, and calculates both $\eta(\kappa)$ and the normalization

constant $A(\kappa) > 0$ defined by

$$A(\kappa)^2 = \left(\frac{1}{\eta - \epsilon}\right)^2 \times$$
$$[x_2^{(\xi)}]^T \left[\alpha^2 \left(\zeta_j - \eta(\kappa)\frac{Z}{Z'}\right)\delta_{ij} + \left((\eta(\kappa) - \epsilon)^2 + \frac{(Z\alpha)^2}{\kappa^2} - \frac{(Z'\alpha)^2}{(\gamma+1)^2}\right) T_{ij}\right] [x_2^{(\xi)}] \tag{4.36}$$

Note $A(\kappa)$ can be computed in $\sim N$ operations because T has bandwidth 3. The lower and upper component function θ_κ and ϕ_κ are given by

$$\begin{aligned} \theta_\kappa(r) &= \frac{1}{A(\kappa)} \sum_{j=1}^{N} (x_2)_j \cdot \xi_j(r) \\ \phi_\kappa(r) &= \left(\frac{-1}{\eta(\kappa) - \epsilon}\right) \frac{1}{A(\kappa)} \sum_{j=1}^{N} (x_2)_j \cdot B(\kappa)\xi_j(r) \end{aligned} \tag{4.37}$$

The normalized lower and upper components for the opposite sign of κ can be computed by changing the sign of κ everywhere in this formula and computing $A(-\kappa)$, or by using the degeneracy 4.6. of the states for opposite signs of κ to get

$$\begin{aligned} \theta_{-\kappa}(r) &= \left(\frac{\epsilon + \eta(\kappa)}{\epsilon - \eta(\kappa)}\right)^{1/2} \frac{1}{A(\kappa)} \sum_{j=1}^{N} (x_2)_j \cdot \xi_j(r) \\ \phi_{-\kappa}(r) &= \left(\frac{1}{\epsilon^2 - \eta^2(\kappa)}\right)^{1/2} \frac{1}{A(\kappa)} \sum_{j=0}^{N} (x_2)_j \cdot B(-\kappa)\xi_j(r) \end{aligned} \tag{4.38}$$

Any pair of functions $(\theta_\kappa(r), \phi_\kappa(r))$ may of course be multiplied by a common overall sign whose value is a matter for convention.

5. Radial Integrals

The theory presented in Sections 1-4 indicates that virtual intermediate states may be found faster, more stably, and with greater economy of memory and with less roundoff error, when expanded in terms of the ξ basis than any other analytic basis. Much of the benefit would be lost if, to calculate a matrix element connecting by the emission of radiation an exact state to the intermediate one, one were forced to expand the Laguerre polynomial $L_{n-1}^{2\gamma+1}$ in the function ξ_n into its component powers of r and to integrate term by term. Such an expansion should be avoided. One also wishes to avoid any expansion in the wavenumber $k = \omega/c$ of the emitted photon. Such an expansion involves the sum of many terms of opposite sign and will converge slowly if $1/k$ is small compared to the spacial extent of either the the exact or intermediate state. Furthermore one wishes to accomodate any choice of the parameter Z' which appears in the exponential of the ξ basis. This can all be done.

The form of the Dirac radial integrals required to calculate with manifest gauge invariance the emission of radiation was worked out by Grant[21]. He wrote his solution to the Dirac radial equation in a central field as[22]

$$\Psi_{nlm}(\vec{r}) = \frac{-i}{r} \begin{bmatrix} iP_{nk}(r)\chi_{km}(\theta,\phi) \\ -Q_{nk}(r)\chi_{-k,m}(\theta,\phi) \end{bmatrix} \tag{5.1}$$

which shows his functions P and Q are the same as the functions f and g used by Goldman (equation 3.2) except for an extra phase factor $-i$. To calculate the matrix element for the emission of a photon of multipolarity L from a pair of states

α, β he found it sufficient to calculate the set of three integrals[21]

$$I_L^{\pm}(\omega) = \int_0^{\infty} \left(P_\alpha Q_\beta \pm Q_\alpha P_\beta\right) j_L(kr)\, dr$$
$$J_L(\omega) = \int_0^{\infty} \left(P_\alpha P_\beta + Q_\alpha Q_\beta\right) j_L(kr) dr \tag{5.2}$$

where j_L is a spherical Bessel function.

Consider replacing one of these exact states (say β) by a variational state. Then P_β and Q_β are each a linear combination via 3.4 of ξ_n and $B\xi_n$. The operator B contains the derivitive d/dr but using the relation[5]

$$\frac{d}{dx} L_n^\alpha(x) = -L_{n-1}^{\alpha+1}(x) \tag{5.3}$$

both ξ_n and $B\xi_n$ may be written as the sum of a few terms each of which have the form of a leading power of r times an exponential times a Laguerre polynomial. The exact states P_α and Q_α have the form of a leading power of r times a polynomial in r times an exponential. The degree of the polynomial of the exact state will be small if the state α lies low on the ladder of bound states of given κ_α. So the contribution of the vector ξ_n to the radial integral is the sum of a small number of terms of the form

$$\int_0^{\infty} e^{-\Lambda r} r^{\gamma_\alpha} r^{n_\alpha} j_L(kr) e^{-\lambda r/2} r^{\gamma_\beta} L_n^{Q'}(\lambda r)\, dr \tag{5.4}$$

Here Λ is the constant in the exponential in the exact state, r^{γ_α} represents the lowest power of r multiplying the exact wavefunction, r^{n_α} represents an additional integer power of r arising from the expansion of the polynomial of the exact state, λ

is the constant $2Z'/(\gamma_\beta+1)$ which appears in the ξ basis, r^{γ_β} is the lowest possible leading power of r arising from the result of operating on ξ_n with B, and n' is an integer in general different than n and Q is a positive number in general different than $2\gamma_\beta+1$.

One starts by using the relation[5]

$$j_L(x) = (-x)^L \left(\frac{1}{x}\frac{d}{dx}\right)^L j_0(x) \tag{5.5}$$

and repeated integrations by parts to transform equation 5.4 into an integration over j_0 instead of j_L. If a function $f(r)$ decreases exponentially at infinity and is no more singular as $r \to 0$ than $r^{L-1+\epsilon}$, where $\epsilon > 0$, then

$$\int_0^\infty f(r) j_L(kr)\, dr = \int_0^\infty \left[\left(\frac{d}{dr}\frac{1}{r}\right)^L r^L f(r)\right] j_0(kr)\, dr \tag{5.6}$$

The integrand in equation 5.4 satisfies the needed conditions. The factor $e^{-(\Lambda+\lambda/2)r}$ guarantees the correct behavior as r goes to infinity. Conservation of angular momentum enforces $L \leq j_\alpha + j_\beta$ and because $\gamma_{\alpha,\beta} = \sqrt{(j_{\alpha,\beta}+1/2)^2 - (Z\alpha)^2}$ the factor $r^{\gamma_\alpha+\gamma_\beta}$ guarantees the correct behavior as r goes to zero. Integrating equation 5.4 by parts and again using 5.3 to eliminate the derivitives of a Laguerre polynomial which appear, one now is left with a sum of integrals of the form

$$\int_0^\infty e^{-(\Lambda+\lambda/2)r} r^b j_0(kr) L_{n''}^{Q''}(\lambda r)\, dr \tag{5.7}$$

where Q'' and n'' have been in general modified again. If $j_\alpha = 1/2$ and $j_\beta = 1/2$ and $L = 1$, as $(Z\alpha)^2 \to 1^-$ one can have $b \to -1^+$, so some of the integrals needed

may have (integrable) singularities as $r \to 0$. One now substitutes for the Laguerre polynomial a confluent hypergeometric function using the indentity[5,23]

$$L_n^{\alpha}(t) = \frac{\Gamma(\alpha+1+n)}{n!\,\Gamma(\alpha+1)} F(-n, \alpha+1, t) \tag{5.8}$$

The resulting integral one can evaluate in terms of a hypergeometric function of a complex variable using the formula

$$\int_0^{\infty} e^{-(\Lambda+\lambda/2)r} r^b j_0(kr) F(-n, c, r)\, dr = \frac{1}{k}\Gamma(b)\,\mathrm{Im}\left[\left(\frac{z}{\lambda}\right)^{-b} F\left(-n, b; c; z\right)\right] \tag{5.9}$$

where the number z is given by

$$z = \frac{\lambda}{\Lambda + \frac{\lambda}{2} - ik} \tag{5.10}$$

Formula 5.9 is valid (at least) for the range of parameters Λ , b, k, c , and λ all real with $\Lambda > 0$, $\lambda > 0$, and $b > -1$. This formula may be established by using the contour integral representations in the complex t plane[23]

$$\begin{aligned} F(-n, c, z) &= \frac{(-1)^n \Gamma(c)}{\Gamma(c+n)} \frac{n!}{2\pi i} \int_C t^{-n-1}(1-t)^{c+n-1} e^{zt}\, dt \\ F\left(-n, \beta; c; z\right) &= \frac{(-1)^n \Gamma(c)}{\Gamma(c+n)} \frac{n!}{2\pi i} \int_C t^{-n-1}(1-t)^{c+n-1}(1-zt)^{-\beta}\, dt \end{aligned} \tag{5.11}$$

which are simultaneously valid if n is an integer ≥ 0, if the common contour C is a path which circles the origin anticlockwise without including the points $t = 1$ or $t = 1/z$, if $|\arg(1-z)| < \pi$, if $(1-zt)^{-\beta} = 1$ when $z = 0$, and if $c \neq 0, -1, \ldots, -n+1$. One also needs the identity[5] $j_0(x) = \sin(x)/x$ and the integral formula[24]

$$\int_0^{\infty} e^{-pr} r^{b-1} \sin(kr) = \frac{i\Gamma(b)}{2}\left[(p+ik)^{-b} - (p-ik)^{-b}\right] \tag{5.12}$$

which is valid for $\mathrm{Re}\,(b) > 0$ and $\mathrm{Re}\,(p) > |\mathrm{Im}\,k|$.

One now considers the evaluation of the hypergeometric function appearing First, note that the function need be evaluated only in the disk in the complex plane $|z-1|<1$. Consider the scaled variables $L=\lambda/\Lambda$ and $K=k/\Lambda$. Then

$$z=x+iy=\frac{L(1+L/2)}{(1+L/2)^2+K^2}+i\frac{LK}{(1+L/2)^2+K^2} \tag{5.13}$$

whence

$$\left(x-\frac{L}{L+2}\right)^2+y^2=\left(\frac{L}{L+2}\right)^2 \tag{5.14}$$

which describes a circle radius $L/(L+2)$ centered at $(x,y)=\big(L/(L+2),0\big)$. As the scale parameter λ runs from zero to infinity the circle expands from the origin through the disk $|z-1|<1$. For fixed λ, as k runs from zero to infinity the point z travels along the circle anticlockwise from the circle's intercept with the positive real axis to the origin.

Second note the leading argument of the hypergeometric is a negative integer, $-n$. The power series expansion about $z=0$ in the complex plane terminates, so the hypergeometric reduces to a simple polynomial in z. Explicitly[23]

$$F\left(-n,b;c;z\right)=\sum_{j=0}^{n}(-1)^j\binom{n}{j}\frac{(b)_j}{(c)_j}z^j \tag{5.15}$$

where the Pochhammer symbol $(b)_j$ is defined by[23]

$$(b)_j=\begin{cases}1 & ,j=0\\ b(b+1)(b+2)\cdots(b+j-1) & ,j\geq 1\end{cases} \tag{5.16}$$

The first two polynomials have the simple forms

$$\begin{aligned}F_0&=F(0,b;c;z)=1\\ F_1&=F(-1,b;c;z)=1-\frac{b}{c}z\end{aligned} \tag{5.17}$$

where I have introduced the shorthand $F_n\equiv F\left(-n,b;c;z\right)$. Except for small n the

series expansion 5.15 for the hypergeometric is not useful for computation, because the combinatorials in the sum are large and the final small result involves a delicate cancellation between terms of great magnitude and opposite sign. Instead one may use the following Gauss recursion relation[23]

$$F_{n+1} = \left[\frac{n}{n+c} + \frac{(c-b)z}{n+c} + (1-z)\right] F_n - \left(\frac{n}{n+c}\right)(1-z)F_{n-1} \qquad (5.18)$$

to calculate F_n for large n starting from the simple seed values in equation 5.17 for $n = 0$ and $n = 1$.

A problem with all linear, three-term recursion relations such as equation 5.18 is that they have have not one but two solutions, and the existence of the second may make the recursion numerically unstable[12]. One can model this phenomenon using the simpler recursion

$$S_{k+1} = (a+b)S_k - abS_{k-1} \qquad (5.19)$$

which for constants a, b has the explicit solutions a^k and b^k. Suppose $|a| < 1$ and initial conditions for S_0 and S_1 select precisely the solution a^k. Any round-off error in the propagation of the recurrance will mix in some of the spurious solution b^k. If $|b| > 1$ the spurious solution grows so fast that both the absolute and relative error in the numbers generated by the recurrance will deteriorate exponentially. If $|a| < |b| < 1$ the absolute error remains small but the relative error still deteriorates exponentially; for $|b| < |a|$ both the absolute and the relative error remain small. If $|a| < 1$ and $|b| < 1$ and the initial conditions call for a mix of the solutions a^k and b^k, the relative error for large k depends on the degree to which the initial conditions determine numerically the coefficient of the slowest falling solution.

The more complicated recursion, 5.18, may be analysed similarly. For the hypergeometrics needed for radial integrals it is always the case that $b > 0$, $c > 0$, and $|z-1| < 1$ so I will consider only these ranges of values. For large n one finds by straight substitution that the two solutions $F(1)$ and $F(2)$ to the recurrance vary asymtotically with large n as $F(1) \sim O(n^{-b})$ and $F(2) \sim O(n^{c-b}(1-z)^n)$. Because both $F(1)$ and $F(2)$ tend to zero as $n \to \infty$ the absolute error in the numbers generated by the recurrance will always be small. To determine the relative error one needs to know how much of the slower dying $F(1)$ is in the asymtotic form of the desired solution F_n. This asymtotic form can be extracted from the generating function for the polynomials F_n. This function $G(t)$ is defined by[23]

$$G(t) = (1-t)^{b-c}\,[1-(1-z)t]^{-b} = \sum_0^\infty \frac{t^n}{n!}(c)_n\, F\,(-n, b; c; z) \tag{5.20}$$

The singularities of $G(t)$ in the complex t plane can be used, following a theorem of Darboux, to find the asymtotic form for large n of the coefficients a_n in the power series expansion expansion $G(t) = \sum a_n t^n$ and hence the asymtotic form of the F_n[25]. The nature of the singularities depends on the values of b and c, and so one has the following asymtotic forms:

$$F_n \sim \begin{cases} (nz)^{-b}\dfrac{\Gamma(c)}{\Gamma(c-b)} & , \quad c-b \neq 0, -1, -2, \ldots \\ \dfrac{\Gamma(c)}{\Gamma(b)}\left(\dfrac{z}{z-1}\right)^{b-c}(1-z)^n & , \quad c-b = 0, -1, -2, \ldots \end{cases} \tag{5.21}$$

which are valid for $b > 0$, $c > 0$, $|z-1| < 1$. Thus the hypergeometric contains a large component of $F(1)$ unless $c-b$ nearly zero or nearly a negative integer, and

the relative precision will be preserved by the recurrance. If $c - b$ is exactly zero or a negative integer the relative precision will deteriorate exponentially.

Some simple examples may illustrate this general result. For $b = c$ one has $F(1) = 1$ and $F(2) = (1 - z)^n$. Because $c - b = 0$ the function F_n should have no component of $F(1)$ as $n \to \infty$. One finds in this special case that F_n is exactly $(1-z)^n$ and the result holds trivially. For $b = 1$ and $c = 2$ one has $F(1) = 1/(k+1)$ and $F(2) = (1 - z)^k/(k + 1)$. Because $c - b$ is not a negative integer or zero F_n should contain a large component of $F(1)$. The form for F_n in this case,

$$F_n = \frac{1}{(n+1)z}\left[1 - (1 - z)^{n+1}\right] \tag{5.22}$$

shows that it does and that its asymtotic form follows the estimate in equation 5.21.

The case where $c - b$ a negative integer can occur readily for Schrodinger problems where both b and c are restricted to positive integers, and for Dirac problems where the two states in a matrix element happen to have the same value of $|\kappa|$. The hypergeometric F_n fortunately reduces when $|c - b|$ is an integer to analytically simple closed forms which may easily be evaluated to high relative precision. For $c - b$ equal to a positive integer these forms make it easy to check the correct functioning of the recurrance. First if $b = c$ one notes the special case[26]

$$F(a, b; b; z) = (1 - z)^{-a} \tag{5.23}$$

so one has a simple result if $b = c$. Then if $b > c$ one can repeatedly use the recursion relation[26]

$$F(a, b; c; z) = \frac{1}{1 - z}F(a, b - 1; c; z) + \frac{a - c}{c(1 - z)}F(a, b; c + 1; z) \tag{5.24}$$

to lower the difference between b and c until 5.23 may be applied. Should in lowering b a value $b = 0$ be encountered one has at once

$$F(a, 0; c; z) = 1 \tag{5.25}$$

Next if $c = b + 1$ an integral form for the hypergeometric[26]

$$F(a, b; c; z) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 t^{b-1}(1-t)^{c-b-1}(1-tz)^{-a}\, dt \tag{5.26}$$

which is valid for $\mathrm{Re}(c) > \mathrm{Re}(b) > 0$ and for $|\arg(1-z)| < \pi$, applies and the integral reduces to

$$F(a, b; b+1; z) = b \int_0^1 t^{b-1}(1-tz)^{-a}\, dt, \qquad b > 0,\ |\arg(1-z) < \pi| \tag{5.27}$$

which can be integrated exactly by parts. Then if $c > b + 1$ one uses repeatedly the recursion[26]

$$F(a, b; c; z) = \frac{-b}{c-b-1} F(a, b+1; c; z) + \frac{c}{c-b-1} F(a, b; c-1; z) \tag{5.28}$$

until $c = b + 1$, and then applies 5.27. One finds by these methods for example that

$$\begin{aligned}
F(-n, 2; 4; z) &= 3 \times \frac{2 - (1-z)^{n+1}(2 + 2(1+n)z + (1+n)(2+n)z^2}{(1+n)(2+n)(3+n)z^3} \\
F(-n, 3; 4; z) &= 6 \times \frac{-2 + z(n+3) + (1-z)^{n+2}(2 + z(n+1))}{(1+n)(2+n)(3+n)z^3} \\
F(-n, 4; 4; z) &= (1-z)^n \\
F(-n, 5; 4; z) &= (1-z)^{n-1}\left(1 - \left(\frac{n+c}{c}\right)\frac{1}{1-z}\right)
\end{aligned} \tag{5.29}$$

which expressions, unlike the series expansion, are readily evaluated numerically for all n if $z \neq 0$. The restriction to $z \neq 0$ is unimportant because that point is only achieved in the unphysical limit where the photon wavenumber k goes to infinity.

6. Numerical Tests and Discussion

Table 1 shows the result of solving using FORTRAN for the eigenvalues and eigenvectors of the Dirac Hamiltonian, using 4.36 and the values $Z = Z' = 92$, $\alpha = 1/137.036$, $\kappa = 1$, and $N = 200$. The eigenvalues were found in Double precision using the NAG[27] implementations of the Crawford (F01BUF and F01BVF), Schwartz (F01BWF). amd QR (F02AVF) algorithms. These eigenvalues served as the seed values for an inverse iteration routine which gave eigenvalues and normalized eigenvectors in quadruple precision. The degeneracy of the eigenvalues for different signs of ϵ_0, and the proportionality (4.6) between the lower component vectors in the ξ basis were reproduced to parts in $\sim 10^{13}$ for double precision and parts in $\sim 10^{29}$ for quadruple precision for all 400 eigenstates.

As a test of the numerical stability of the method for large N, the double precision matrix diagonalization was repeated for $Z = 92$, $Z' = Z'_0$, and $N = 2400$ (which means for a 4800×4800 matrix). The degeneracy of the eigenvalues ϵ for opposite signs of κ was maintained for all 4800 eigenvalues to parts in 10^{13}. The eigenvalues from one sign of κ could be isolated in about 10 minutes on an IBM 370.

As a test of the recursion relation for the hypergeometric, equation 5.18 was used in double precision to generate F_{5000} for $b = 1$, $c = 2$, and for a range of z in the disk $|1 - z| < 1$. The results agreed to parts in 10^{13} with the result of the exact formula 5.22 for this special case.

It is a pleasure to thank Professor S. P. Goldman for a fruitful correspondance, and to acknowledge illuminating conversations with Jacqueline Holen, Tony Y. Lin and Dale Koetke. This work was supported by the Department of Energy under contract DE-AC03-76SF00515.

Table 1				
1	9.3304196879495E-01	9.3304196879495E-01	4.3E-17	1.7E-33
2	9.7129255046708E-01	9.7129255046708E-01	2.7E-15	2.9E-33
3	9.8427712040503E-01	9.8427712040502E-01	7.6E-15	-6.2E-33
4	9.9012266260062E-01	9.9012266260062E-01	2.8E-15	-1.0E-32
5	9.9323209176943E-01	9.9323209176942E-01	6.5E-15	-1.1E-32
6	9.9507740297513E-01	9.9507740297511E-01	1.3E-14	-1.3E-32
7	9.9626039606096E-01	9.9626039606094E-01	1.6E-14	-1.4E-32
8	9.9706353738649E-01	9.9706353738647E-01	1.4E-14	-1.8E-32
9	9.9763346229836E-01	9.9763346229834E-01	1.2E-14	-2.0E-32
10	9.9805237301537E-01	9.9805237301536E-01	1.1E-14	3.3E-30
11	9.9836922962717E-01	9.9836922962716E-01	1.1E-14	1.9E-24
12	9.9861465831169E-01	9.9861465831168E-01	9.0E-15	1.1E-19
13	9.9880860859954E-01	9.9880860859953E-01	7.4E-15	9.1E-16
14	9.9896452521643E-01	9.9896452521642E-01	6.7E-15	1.6E-12
15	9.9909173633374E-01	9.9909173633374E-01	6.3E-15	6.5E-10
...	.....	.....	.....	.....
196	8.2009746568666E+00	8.2009746568667E+00	1.0E-14	
197	9.8826417422065E+00	9.8826417422065E+00	3.6E-15	
198	1.2485488417102E+01	1.2485488417102E+01	1.5E-14	
199	1.7096321858787E+01	1.7096321858788E+01	7.6E-14	
200	2.7892987982837E+01	2.7892987982842E+01	1.6E-13	

Table 1 Column 1 shows the positive energy eigenvalues from a solution in double precision of 4.36 for $Z = Z' = 92$, $\alpha = 1/137.036$, $\kappa = 1$ and $N = 200$. Column 2 shows the eigenvalues from an inverse iteration routine written in quadruple precision and are taken as a standard; the underlined digits in Column 1 are the first in error. Column 3 shows the fractional error between the double and quadruple precision results. Column 4 shows the fractional error between the quadruple precision eigenvalues for bound states and the corresponding Sommerfeld values; a positive error means the variational eigenvalue lies (correctly) above the Sommerfeld. Some values are negative because of roundoff error.

REFERENCES

1. M. Rotenberg, Annals of Physics **19**, 262, 1962.
2. E. L. Ince, *Ordinary Differential Equations*, Longmans, Green and Co. Ltd. (1927), Chapter 10.
3. S. P. Goldman, Phys. Rev. A **31**, 3541 (1985).
4. S. P. Goldman, Phys. Rev. A **40**, 1185 (1989).
5. A. M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, Dover Publications, Inc., New York(1980).
6. C. R. Crawford, Numerical Mathematics **16**, 41 (1973).
7. G. H. Golub and C. F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press (1989).
8. W. H. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vettering, *Numerical Recipes (Fortran Version)*, Cambridge University Press(1989).
9. A. R. Edmonds, J. Phys. B: Atom. Molec. Phys., Vol. **6**, 1603(1973).
10. C. W. Clark and K. T. Taylor, J. Phys. B: Atom. Molec. Phys. **13**, L737 (1980).
11. C. W. Clark and K. T. Taylor, J. Phys. B: At. Mol. Phys. **15**, 1175 (1982).
12. W. H. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vettering, *Numerical Recipes (Fortran Version)*, Cambridge University Press(1989).
13. S. Barnett, *Matrices: Methods and Applications*, Clarendon Press, Oxford, (1990).
14. P. M. Morse and H. Feshbach, *Methods of Theoretical Physics, Part I*, McGraw-Hill Book Company, Inc., (1953).
15. G. W. F. Drake and S. P. Goldman, in *Advances in Atomic and Molecular Physics* **25**, 393, Academic Press (1988).

16. If $M = N+1$, for $\kappa < 0$ the function represented by the spurious eigenvector happens to converge to the exact lowest energy state (*e.g.*, for $\kappa = -1$ the $1s_{1/2}$), while for $\kappa > 0$ the eigenvector remains spurious and must be discarded.

17. The necessary restrictions on the functions $\{w\}$ so that one has a useful variational representation are not known. Because the exact eigenstates of the Dirac Hamiltonian have series expansions of the form $\theta = r^{\gamma+1}\Sigma a_k r^k$ it is sufficient to require that a linear combination of the functions $\{w\}$ reproduce any series of this form. (Goldman in reference 3 worked with specific functions of the form $w_j = r^{\gamma+1+j}e^{-\lambda r}$, where $j = 0, \ldots N-1$, but his proofs generalize easily). A set of functions $\{w\}$ without the leading $r^{\gamma+1}$ dependence may however still serve, though there are difficulties to overcome. The convergence of the variational wavefunctions to the exact as $N \rightarrow \infty$ may be weakened from uniform to, perhaps, mean-square, and some familiar operations, such as term-by-term differentiation of the expansion, may become tricky to justify. One would also have special problems near $r = 0$.

18. H R. Schwartz, in *Linear Algebra*, J. H. Wilkinson and C. Reinsch, eds, Springer-Verlag (1971); see also Numer. Math. **12**, 231 (1968).

19. G. Peters and J. H. Wilkinson, Siam Review **21**, 339 (1979).

20. Goldman's proof of this point is subtly flawed. Referring to the equations and notations of reference 3, one can define a lower component function θ_i' for $\kappa > 0$ by 3.27a and conclude that 3.22 holds. One can define corresponding upper component function ϕ_i' by 3.18, with $\kappa > 0$, so that 3.19 holds. If 3.20 for $\kappa > 0$ held as well, then the defined function (ϕ_i', θ_i') would be a variational eigenstate with $\kappa < 0$. However, one cannot reach 3.20 from 3.22 without assuming the defined function is normalized, $\langle\phi_i'|\phi_i'\rangle + \langle\theta_i'|\theta_i'\rangle = 1$, and no justification for this assumption is given. The proof given here is offered as a replacement.

21. I. P. Grant, J. Phys. B: Atom. Molec. Phys. **7**, 1458(1974).

22. I. P. Grant, Adv. Phys. **19**, 747(1970).

23. Z. X. Wang and D. R. Guo, *Special Functions*, World Scientific Publishing Co. (1989).

24. A. Erdélyi, W. Magnus, F. Oberhettinger and F. G. Tricomi, *Tables of Integral Transforms, Volume I*, McGraw-Hill Book Co., Inc., 1954.

25. H. S. Wilf, *generatingfunctionology*, Academic Press, 1990.

26. A. Erdélyi, W. Magnus, F. Oberhettinger and F. G. Tricomi, *Higher Transcendental Functions, Volume I*, McGraw-Hill Book Co., Inc., 1953.

27. Numerical Analysis Group (U.S.A.) Inc.