

# Chapter 22

## The Dirac Equation in the Algebraic Approximation

H.M. Quiney

University of Melbourne, Melbourne, Australia

---

1 Introduction	423
2 Rayleigh–Ritz Method for Schrödinger and Dirac Operators	424
3 Basis Sets for Electronic Structure Calculations	430
4 Applications to Relativistic Atoms	438
References	443

---

### 1 INTRODUCTION

The first reported attempts to solve the Dirac equation using the Rayleigh–Ritz method, also known as the algebraic approximation, were made by Synek,<sup>(1)</sup> who presented the formalism and Kim,<sup>(2,3)</sup> who was able to perform calculations on closed-shell atoms. Kim noted that the solutions exhibited strong sensitivity to the choice of the basis-set parameters and that the variational approximations obtained in this way appeared not to satisfy the usual Undheim–Hylleraas–Macdonald interleaving theorem of quantum chemistry. The variational approximations to bound-state energies obtained in this scheme did not provide upper bounds to the exact energies, and catastrophic variational collapse could be encountered in which no reasonable physical interpretation could be placed on any state.

A series of articles appeared that extended the algebraic formulation of the relativistic Dirac–Hartree–Fock (DHF) equations to diatomic and polyatomic systems, but most

failed to address the variational instability of the calculated spectrum, which was clearly related to the detailed choice of the basis set. A step in the right direction was made by Lee and McLean,<sup>(4,5)</sup> who reported that some control of the spectrum could be regained if the small-component basis set was *augmented* by functions that are generated by the action of the two-component operator  $\sigma \cdot p$  on a large-component spin-orbital basis set selected according to non-relativistic quantum-chemical prescriptions. This prescription alone is not sufficient to provide a complete solution to the variational problems associated with the use of the Dirac operator, but their success in calculating a plausible electronic structure for AuH in the DHF approximation served to stimulate considerable further research into the finite matrix representation of the Dirac operator.

The unsatisfactory situation regarding the use of the Dirac operator in quantum-chemical basis-set calculations at that time was reviewed by Kutzelnigg,<sup>(6)</sup> who offered several prescriptions to avoid variational collapse of these basis-set parametrizations. One of these methods is now known as *kinetic balance* and it forms the basis of most successful attempts to solve matrix eigenvalue equations based on finite-dimensional representations of the Dirac operator. The method of kinetic balance is reviewed in this chapter, and has been developed as a systematic and reliable implementation of the empirical prescription advocated by Lee and McLean.<sup>(4)</sup> It has been refined by several groups to satisfy subsidiary requirements associated with the use of point-charge models of nuclear structure and to accommodate the generation of large, linearly independent basis sets. The common feature of all successful approaches, however, is the one-to-one matching of the basis sets that are used to expand the upper and lower components through a functional relationship that is linear in the operator  $\sigma \cdot p$ , and the

selection of trial functions that are appropriate to the choice of nuclear model, which is adopted.

The use of kinetic balance certainly avoids the worst aspects of variational collapse and generates a spectral representation, which reproduces the relativistic fine structure to good accuracy if a sufficiently flexible basis set is employed. One may not, however, make any unequivocal statements regarding the provision of upper bounds to the state energies when kinetic balance is adopted. The analysis of Stanton and Havriliak<sup>(7)</sup> based on Gershgorin's theorem indicates that the variational estimates of bound-state eigenvalues may fall below the exact values by an amount  $O[(Z/c)^4]$ , where  $Z$  is the nuclear charge and  $c$  is the speed of light, a phenomenon that has frequently been observed in practice. These 'small variational collapses' are most often seen, however, when a basis set is highly contracted and are indicative of a lack of completeness of the set rather than of any catastrophic failure of the underlying theory. In practice, these small bounds failures cause no problem, and are usually not observed if the basis set is large or used in the uncontracted form. The important feature of the relativistic algebraic approximation is that properly matched basis sets provide accurate parametrizations of Dirac spinors, and that these representations become increasingly accurate as the basis set is extended towards completeness. The resulting spectral representation provides a convenient basis for the implementation of relativistic self-consistent field (SCF) methods and relativistic many-body theories, including the modelling of processes that involve the creation of virtual electro-positron pairs. The algebraic approximation has become an important technique in relativistic quantum chemistry and has led to the construction of several *ab initio* computational packages, which are the subjects of ongoing development.

In this chapter, we review some of the mathematical features of the Rayleigh–Ritz method and examine the features that are peculiar to basis-set parameterizations of the Dirac equation. The kinetic balance prescription is presented as a prototype for all valid relativistic basis-set methods, which use sets that are generated in matched pairs of functions. We then present a number of choices of basis set that have been used in atomic and molecular structure theory, and illustrate the use of the complete spectrum in modelling processes that formally involve the creation of virtual electron–positron pairs.

## 2 RAYLEIGH–RITZ METHOD FOR SCHRÖDINGER AND DIRAC OPERATORS

In discussing the Rayleigh–Ritz method applied to the expansion of Dirac spinors in atomic and molecular physics,

some behaviour that differs from familiar results, which are specific to the properties of the Schrödinger equation, have sometimes been misinterpreted as fatal pathologies in the relativistic context. This is particularly true of the failure of the relativistic variational approximations to provide rigorous upper bounds to the exact values, and the consequent failure of the Undheim–Hylleraas–MacDonald interleaving theorem. In this section, we shall find that the Dirac equation in the algebraic approximation satisfies a variational principle that sets sufficiently rigorous bounds on the discrete eigenvalue spectrum to ensure that accurate representations may be generated, that they may be refined by extending the basis set towards completeness, and that they are amenable to unambiguous physical interpretation within the context of quantum electrodynamics (QED). It is clear from the presence of positive- and negative-energy branches of the Dirac spectrum and the charge-conjugation symmetries of the Dirac operator, however, that the non-relativistic interleaving theorem cannot apply to all elements of the spectrum and that it need not, therefore, apply to any single element. In practice, one finds quasi-non-relativistic behaviour for the bound-state solutions of the Dirac equation in the algebraic approximation, in which the behaviour is dominated by the non-relativistic parts of the wave function. This is sufficient to implement the algebraic approximation for the accurate determination of the relativistic electronic structures of atoms and molecules.

The known mathematical results relating to the Rayleigh–Ritz method are reviewed here to examine how they can be applied to both non-relativistic and relativistic problems. This review is a summary of the analysis of Grant and Quiney,<sup>(8)</sup> in which further details may be found.

### 2.1 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

$$0 \leq \lambda_1 \leq \lambda_2 \leq \dots$$

some of which may be degenerate, with corresponding eigenfunctions  $\psi_1, \psi_2, \dots$ , respectively. The *Rayleigh quotient*, defined (in Dirac bra-ket notation) by

$$R[\psi] := \frac{\langle \psi | T | \psi \rangle}{\langle \psi | \psi \rangle} \quad (1)$$

evidently exists for all non-trivial  $\psi \in \mathcal{D}$ . The Rayleigh–Ritz method assumes that we can approximate every such

$\psi$  by its projection,  $P_n \psi$ , on a finite-dimensional subspace  $\mathcal{W}_n := \{\chi_1, \chi_2, \dots, \chi_n\} \subset \mathcal{D}$ , so that there exists some set of coefficients (in general complex) for which

$$P_n \psi = c_1^n \chi_1 + c_2^n \chi_2 + \dots + c_n^n \chi_n \quad (2)$$

A simple calculation shows that

$$F^n[\mathbf{c}^n] := R[P_n \psi] = \frac{\mathbf{c}^{n\dagger} \mathbf{T}^n \mathbf{c}^n}{\mathbf{c}^{n\dagger} \mathbf{S}^n \mathbf{c}^n} \quad (3)$$

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

$$T_{ij}^n = \langle \chi_i | T | \chi_j \rangle, \quad S_{ij}^n = \langle \chi_i | \chi_j \rangle, \quad 1 \leq i, j \leq n \quad (4)$$

respectively. The Gram matrix  $\mathbf{S}^n$  must be positive definite if equation (3) is to have a meaning, so that the set  $\mathcal{W}_n$  must be linearly independent.

This manipulation approximates the Rayleigh quotient by an algebraic function  $F^n[\mathbf{c}^n]$ . Its stationary points are the roots of the Galerkin equation<sup>(9)</sup>

$$\mathbf{T}^n \mathbf{c}^n = \Lambda^n \mathbf{S}^n \mathbf{c}^n \quad (5)$$

where  $\Lambda^n$  is a diagonal eigenvalue matrix whose elements can be ordered so that

$$0 < \Lambda_1^n \leq \Lambda_2^n \leq \dots \leq \Lambda_n^n$$

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

The literature of quantum chemistry conventionally establishes at this point that the lowest eigenvalue  $\Lambda_1^n$  is an upper bound to the true eigenvalue  $\lambda_1^n$ , based on the fact that  $T$  is a positive definite operator. This conclusion is easily generalized to encompass all operators  $T$  that are bounded below. This needs to be considered in the more general setting furnished by Stakgold;<sup>(9)</sup> the reader should be aware that Stakgold's inequalities are reversed because eigenvalues are listed in the reverse order to the one used here.

Let  $T$  be as defined above. Then the Poincaré Theorem states that eigenvalues of the Galerkin equation (5) are upper bounds to the target eigenvalues

$$0 < \lambda_1 \leq \Lambda_1^n \leq \dots \leq \lambda_n \leq \Lambda_n^n$$

Stakgold<sup>(9)</sup> lists a number of comments regarding this theorem, in particular,

1. If  $T$  is *non-positive* 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 non-positive  $T$  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  $\mathcal{W} := \{\chi_i\}_{i=1}^\infty$  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  $P_n T \rightarrow T$  in the operator norm.

Of these, it is the third observation that is the most crucial, since we must construct representations that become more accurate as the basis set is enlarged. The second point is suggestive of the existence of an interleaving theorem of the Poincaré type when  $T$  is the Dirac operator in which the positive-energy energies converge to the exact eigenvalues from above, and the negative-energy energies from below, but the assertion that this is a relativistic extension of the Undheim–Hylleraas–MacDonald theorem requires special care in the context in which calculations are actually performed. The universal approach that is adopted is to expand the upper and lower components of Dirac spinors in finite, and usually non-orthogonal basis sets of two-spinors. This is not, in general, equivalent to expanding the spinors in a finite set of four-spinors chosen from strictly non-overlapping sequences of positive- and negative-energy basis functions. It is important in this case to consider the domain of the operator, for convenience dividing it into subdomains,  $\mathcal{D}_+$  and  $\mathcal{D}_-$ . If one employs basis sets that span these subdomains, then one may infer a generalized interleaving theorem in which the eigenvalues of the Dirac operator approach from above for the positive-energy states, and from below for the negative-energy states, but such basis sets are necessarily of four-component type. In practice, one employs two-component functions to expand separately the upper and lower components of Dirac spinors, for which one may make no simple statements regarding the energy domains spanned by the functions. It is not sufficient to equate the upper components with the positive-energy domain and the lower components with the negative-energy domain in order to establish a relativistic interleaving principle.

## 2.2 Convergence of Rayleigh–Ritz approximations

The convergence of Rayleigh–Ritz eigenvalues and eigenfunctions of non-relativistic quantum Hamiltonian operators was analyzed in Reference 10. Let  $T = \mathbf{p}^2/2$  be the usual particle kinetic energy, and consider the Sobolev space

$W_2^{(1)} \subset L^2$  equipped with the norm

$$\|\psi\|_{W_2^{(1)}}^2 = \langle u | 1 + T | u \rangle = \|(1 + T)^{1/2} u\|_{L^2}^2 \quad (6)$$

Klahn and Bingel<sup>(10)</sup> establish that the approximate Rayleigh–Ritz eigenvalues converge to the eigenvalues of the target Hamiltonian (which they term *E-convergence*) if the set  $\mathcal{W}$  is complete in  $W_2^{(1)}$ . The approximation  $\psi_i^n := P_n \psi_i$  to the  $i$ th bound state converges in the mean to the eigenfunction  $\psi_i$  whenever the corresponding eigenvalue,  $E_i^n$ , converges to  $E_i$  as  $n$  increases. It suffices 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<sup>(11)</sup> to the convergence of expectation values and transition matrix elements. If  $A$  is assumed to be a strictly positive operator, self-adjoint on a domain  $\mathcal{D}(A)$ , a new  $A$ -norm may be defined by

$$\|\psi\|_A^2 = \langle \psi | A | \psi \rangle = \|A^{1/2} \psi\|_{L^2}^2 \quad (7)$$

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  $A = 1 + T$ . It may also be established that the sequence

$$\langle A \rangle^n := \langle \psi^n | A | \psi^n \rangle$$

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

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

and

$$\|\psi^n - \psi\|_A^2 \leq |\langle A \rangle^n - \langle A \rangle| + 2\|A\psi\| \cdot \|\psi^n - \psi\|$$

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  $T$ , meaning that there exists a pair of non-negative numbers  $a, b$  such that

$$|\langle \psi | A | \psi \rangle| < a \langle \psi | \psi \rangle + b \langle \psi | T | \psi \rangle, \quad \forall \psi \in \mathcal{D}(T) \quad (8)$$

This includes a wide range of operators, including *bounded* operators, for which we can set  $b = 0$ , Coulomb potentials,  $T$  itself (with  $a = 0$  and  $b = 1$ ), components of

the momentum operator  $\mathbf{p}$ , and non-relativistic atomic and molecular Hamiltonians,  $H_S$ . 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 equation (8), we choose  $k > 0$  so that  $T + k$  is strictly positive, and the sequence  $\psi^n$  is  $E$ -convergent to  $\psi$  in the  $T + k$  norm, then

$$\|\psi^n - \psi\|_A^2 \leq \max\left(\frac{a}{k}, b\right) \|\psi^n - \psi\|_{T+k}^2$$

so that  $\{\psi^n\}$  is also  $A$ -convergent to  $\psi$ . This means that it is sufficient for  $\mathcal{W}$  to be complete in the Sobolev space  $W_2^{(1)}$ . The transition matrix elements of the form  $\langle \psi_i^n | A | \psi_j^n \rangle$  also converge to the desired limit  $\langle \psi_i | A | \psi_j \rangle$  as  $n \rightarrow \infty$  provided the sequences  $\{\psi_i^n\}$  and  $\{\psi_j^n\}$  are also  $A$ -convergent.

## 2.3 Extension to Dirac operators

### 2.3.1 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 books of Kato (see Chapter 5 of Reference 12), Reed and Simon (see Chapter 10 of Reference 13) and Thaller<sup>(14)</sup> adopt 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 case of the Schrödinger operator, Reed and Simon<sup>(13)</sup> demonstrate 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<sup>(13)</sup> extends this to the full non-relativistic Hamiltonian for atoms and molecules.

The situation regarding the relativistic Dirac Hamiltonian is complicated by the need to consider a range of cases depending on the strength of the external field the nuclear model involved in its construction. For the free-electron Dirac Hamiltonian (in Hartree atomic units),

$$H_0 := c \alpha \cdot \mathbf{p} + \beta c^2 \quad (9)$$

$H_0$  is essentially self-adjoint on  $C_0^\infty(\mathbf{R}^3 \setminus 0)^4$  and self-adjoint on the Sobolev space  $\mathcal{D}(H_0) = W_1(\mathbf{R}^3)^4 \subset L^2(\mathbf{R}^3)^4$  (see Theorem 1.1 of Reference 14). 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  $C_0^\infty(\mathbf{R}^3 \setminus 0)^4$  and self-adjoint on  $\mathcal{D}(H_0)$  only if  $Z/c < \sqrt{3}/2$  (see Theorem 4.4 of Reference 14). Since  $c \approx 137$  in atomic units,

this restricts the potentials to  $Z < 118$ , which satisfactorily covers all elements in the periodic table. Grant<sup>(15)</sup> and Grant and Quiney<sup>(8)</sup> explain how this restriction may be relaxed to include all cases point-nuclear models for  $Z < 137$ .

The idealized point-charge nucleus is not an adequate 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 that is infinitely differentiable. In this case,  $H$  is again essentially self-adjoint on  $\mathcal{D}(H_0)$  (see Theorem 4.3 of Reference 14). The extension to many-electron atoms or molecules is not considered in Thaller.<sup>(14)</sup> The inclusion of a mean-field potential in the DHF equations does not, however, change the essential character of the one-body Hamiltonian, which is determined by the nature of its singular points.

### 2.3.2 Rayleigh–Ritz method for the Dirac operator

For simplicity, consider the Dirac–Coulomb atomic Hamiltonian

$$H(v) := c \boldsymbol{\alpha} \cdot \mathbf{p} + mc^2 \beta - v \frac{Z}{r} \quad (10)$$

with  $Z/c < \sqrt{3}/2$  on  $\mathcal{D}(H_0) = W_1(\mathbf{R}^3)^4$ , 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  $v$  defines a family of Dirac Hamiltonians interpolating smoothly between  $H(0) = H_0$  and  $H(1)$ , which incorporates the full strength Coulomb field. A trial wave function of the general form

$$\phi := \begin{bmatrix} \sum_{n=1}^N c_n^L M[L, n\kappa m, \mathbf{r}] \\ i \sum_{n=1}^N c_n^S M[S, n\kappa m, \mathbf{r}] \end{bmatrix} \quad (11)$$

is assumed. Later we shall identify the expansion functions  $M[T, n\kappa m, \mathbf{r}]$ ,  $T = L, S$  with one of several choices of spinor basis set; but at this stage, we merely assume that the set is complete in  $\mathcal{D}(H_0)$ . The Galerkin equations take the form

$$\begin{bmatrix} c^2 \mathbf{S}^{LL} + v \mathbf{V}^{LL} & c \boldsymbol{\Pi}^{LS} \\ c \boldsymbol{\Pi}^{SL} & -c^2 \mathbf{S}^{SS} + v \mathbf{V}^{SS} \end{bmatrix} \begin{bmatrix} \mathbf{c}^L \\ \mathbf{c}^S \end{bmatrix} = E \begin{bmatrix} \mathbf{S}^{LL} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}^{SS} \end{bmatrix} \begin{bmatrix} \mathbf{c}^L \\ \mathbf{c}^S \end{bmatrix} \quad (12)$$

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  $\boldsymbol{\Pi}^{T\bar{T}}$  (where  $\bar{T} = S$  when  $T = L$  and vice versa) are given by

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

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

The system is Hermitian and consequently 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  $\phi$  is an arbitrary trial function in  $\mathcal{D}(H_0)$ , and let

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

where  $V_\phi = \langle \phi | V | \phi \rangle$ . Suppose that  $\phi$  has been chosen so that  $E_\phi(0) > mc^2$ . Since  $V$  has, by hypothesis, a strictly negative expectation,  $V_\phi < 0$ , then  $E_\phi(v) < E_\phi(0)$  decreases monotonically as  $v$  increases from 0 to 1, and we shall have  $mc^2 > E_\phi(v) > -mc^2$  if  $E_\phi(0)$  is not too large. Examples of this behaviour may be found, for example, in Reference 16.

As only point eigenvalues of bound states can lie in  $(-mc^2, mc^2)$ ,  $E_\phi(v)$  will approximate a bound-state eigenvalue for values of  $v$  in some interval  $0 < v_0 \leq v \leq 1$ . A sufficient condition that  $E_\phi(1) > -mc^2$  is that  $0 > V_\phi \geq V_{\min} > -2mc^2$ , so that no state of the positive-energy spectrum can ever enter the ‘negative-energy sea’ as  $v$  approaches unity. Of course, a state with  $E_\phi(0) < -mc^2$  will have  $E_\phi(v) < -mc^2$  for all values  $0 \leq v \leq 1$ . 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  $-mc^2$  au, in the manner suggested by Stakgold.<sup>(9)</sup>

This behaviour applies also to the eigensolutions of the Galerkin equations (12). We shall see that with admissible sets of expansion functions discussed in this chapter, characterized by matched sets of  $N$  functions for both large and small components, a discrete pseudo-spectrum is obtained that generates  $N$  eigensolutions in the interval  $-\infty < E_i^N < -mc^2$ , ( $i = 1, 2, \dots, N$ ) and  $N$  in  $-mc^2 < mc^2 + V_{\min} < E_i^N < \infty$ , ( $i = N + 1, N + 2, \dots, 2N$ ). This behaviour is consistent with the discussion following the Poincaré theorem, in particular, with respect to the convergence of the spectral representations that span the domain of the Dirac operator. It guarantees that one may construct spectral representations of dimension  $2N$  in which precisely  $N$  solutions may be classified as ‘negative-energy states’

and  $N$  solutions are ‘positive-energy states’, with effective upper and lower bounds that separate the branches. It is not sufficient, however, to ensure that the eigenvalues are bounded from above. There is ample evidence to confirm that the observed behaviour of the positive-energy branch of the eigenvalue spectrum is pseudo-non-relativistic if the basis set is large and flexible and that the observed small failure to provide a rigorous upper bound is of little practical consequence. The difficulty in establishing strict upper bounds on relativistic finite basis-set approximations, and a feature that confounds efforts to assert a proof of the existence of such bounds, is that one may not generally assume that a state whose eigenvalue is less than  $-mc^2$  au in a given representation has a zero projection onto the space of states that are classified as being of positive energy in a *finite* representation, which is constructed by enhancing the basis set by adding one large-component function, and one small-component function. Similarly, a positive-energy bound state in a given finite-dimensional representation may possess a non-vanishing projection onto the negative-energy space of an enhanced finite-dimensional representation, possibly lowering its energy below the exact value. The enlargement of a finite-dimensional representation constructed from large- and small-component amplitudes does not, in general, act separately to lower the positive-energy eigenvalues and to raise the negative-energy eigenvalues. Instead, augmentation of the basis by adding to the large- and small-component sets causes a rotation between all states whose direction may not readily be established, but which is observed to vanish as the basis set tends towards completeness.

Grant<sup>(17)</sup> discusses the specification of a Dirac-norm, and the convergence characteristics of algebraic approximations in this norm. He concludes that smooth convergence of these finite-dimensional representations may be expected for the bound states of Dirac Hamiltonians corresponding to mean-field atomic or molecular potentials. This favourable behaviour also extends to the convergence of algebraic approximations to the matrix elements of operators, such as position, momentum and powers of the radial coordinate, which typically arise in computational atomic and molecular physics.

## 2.4 Kinetic balance

The necessity of a one-to-one mapping between the elements of the large- and small-component basis sets has been discussed by Dyllal *et al.*,<sup>(18,19)</sup> Stanton and Havriliak<sup>(7)</sup> and Grant and Quiney.<sup>(20)</sup> Here, we examine a generalization of the work of Dyllal *et al.*<sup>(21)</sup> to examine the consequences of kinetic balance on the finite-dimensional representation of the complete Dirac spectrum, rather than just its

positive-energy solutions. It is instructive to consider the free-particle Dirac equation

$$\begin{bmatrix} c^2 - \varepsilon & c\boldsymbol{\sigma}\cdot\mathbf{p} \\ c\boldsymbol{\sigma}\cdot\mathbf{p} & -c^2 - \varepsilon \end{bmatrix} \begin{bmatrix} \psi_\varepsilon^L \\ \psi_\varepsilon^S \end{bmatrix} = 0$$

Straightforward algebra yields

$$(\boldsymbol{\sigma}\cdot\mathbf{p})\psi_\varepsilon^T = \eta_T \left( \frac{c^2 + \eta_T \varepsilon}{c} \right) \psi_\varepsilon^{\bar{T}}$$

where  $\eta_T = 1$  if  $T = L$ ,  $\eta_T = -1$  if  $T = S$  and  $\bar{T} \neq T$ . Further use of the operator identity

$$(\boldsymbol{\sigma}\cdot\mathbf{p})(\boldsymbol{\sigma}\cdot\mathbf{p}) = \hat{p}^2 \mathbf{I} \quad (13)$$

and the relativistic energy-momentum relation

$$\varepsilon^2 = c^2 p^2 + c^4$$

leads to separate equations for  $\psi_\varepsilon^T$ ,

$$\hat{p}^2 \mathbf{I} \psi_\varepsilon^T = p^2 \psi_\varepsilon^T \quad (14)$$

If we now consider the *matrix* representation of the free-particle Dirac equation

$$\begin{bmatrix} (c^2 - \varepsilon) \mathbf{S}_{LL} & c \mathbf{\Pi}_{LS} \\ c \mathbf{\Pi}_{SL} & -(c^2 + \varepsilon) \mathbf{S}_{SS} \end{bmatrix} \begin{bmatrix} \mathbf{c}_L \\ \mathbf{c}_S \end{bmatrix} = 0$$

the vectors of expansion coefficients,  $\mathbf{c}_L$  and  $\mathbf{c}_S$ , form the solutions of a pair of simultaneous matrix equations

$$\mathbf{\Pi}_{\bar{T}T} \mathbf{c}_T = \eta_T \left( \frac{c^2 + \eta_T \varepsilon}{c} \right) \mathbf{S}_{\bar{T}\bar{T}} \mathbf{c}_{\bar{T}} \quad (15)$$

Equations involving either  $\mathbf{c}_L$  or  $\mathbf{c}_S$  may be obtained by elimination, yielding

$$\mathbf{\Pi}_{T\bar{T}} \mathbf{S}_{\bar{T}\bar{T}}^{-1} \mathbf{\Pi}_{\bar{T}T} \mathbf{c}_T^\dagger = p^2 \mathbf{S}_{TT} \mathbf{c}_T \quad (16)$$

Decoupling of the matrix representation of the Dirac equation (16) into valid representations of the Dirac free-particle operator equation (14) is possible only if the matrix identities

$$\mathbf{\Pi}_{T\bar{T}} \mathbf{S}_{\bar{T}\bar{T}}^{-1} \mathbf{\Pi}_{\bar{T}T} = \mathbf{p}_{T\bar{T}}^2 \quad (17)$$

are satisfied, where  $\mathbf{p}_{T\bar{T}}^2$  is the matrix representation of the operator  $(\boldsymbol{\sigma}\cdot\mathbf{p})(\boldsymbol{\sigma}\cdot\mathbf{p}) = \hat{p}^2 \mathbf{I}$  in the basis set  $\{\varphi_k^T\}$ . The matching of basis functions according to

$$\{\varphi_k^S\} = \{N_k^S \boldsymbol{\sigma}\cdot\mathbf{p} \varphi_k^L\} \quad (18)$$

generates representations that satisfy the equation (17), because

$$\mathbf{\Pi}_{LS} = \mathbf{p}_{LL}^2 \quad (19)$$

$$\mathbf{\Pi}_{SL} = \mathbf{S}_{SS} \quad (20)$$

$$\mathbf{\Pi}_{LS} = \mathbf{\Pi}_{SL}^\dagger \quad (21)$$

Without loss of generality, we have chosen the arbitrary constants  $\{N_k^S\}$  to be unity, but we need make no further assumptions about the functional form or normalization of the sets  $\{\varphi_k^L\}$  and  $\{\varphi_k^S\}$ . Substituting equations (19) and (20) into equation (16) yields two matrix representations of the free-particle Schrödinger equation for the large- and small-component functions

$$\mathbf{p}_{LL}^2 \mathbf{c}_T = \mathbf{p}_{LL}^2 \mathbf{S}_{LL} \mathbf{c}_T \quad \text{for } T = L, S \quad (22)$$

The large- and small-component expansion vectors,  $\mathbf{c}_L$  and  $\mathbf{c}_S$ , are solutions of *identical* generalized matrix eigenvalue equations which, following these manipulations, can be written entirely in terms of  $\mathbf{p}_{LL}^2$  and  $\mathbf{S}_{LL}$ , even though the large- and small-component basis functions are different. As a consequence of the identical values of  $p^2$ , which arise in the partitioned equations for  $\mathbf{c}_T$ , there is a one-to-one mapping of the positive- and negative-energy eigenvalues of the free-particle four-component equations of the form  $E \rightarrow -E$  if strict kinetic balance is enforced. Substitution of equation (20) into equation (15) determines the simple relationship

$$\mathbf{c}_S = \frac{c}{c^2 + \varepsilon} \mathbf{c}_L$$

between the large- and small-component vectors. The origin of the notation  $L$  and  $S$  for components is obvious if one considers the relative magnitudes of  $\mathbf{c}_L$ , for a *positive-energy* eigenvalue,  $\varepsilon > c^2$ . For *negative-energy* solutions, however,  $\varepsilon < -c^2$  and the conventional notation is misleading. A  $2M$ -dimensional space of four-component solutions of the Dirac equation,  $\psi_\varepsilon$ , is constructed for the eigenvalue spectrum  $\varepsilon = \pm c^2 \sqrt{1 + (p/c)^2}$  from the  $M$  linearly independent solutions of the two-component equation (22), each with eigenvalue  $p^2$ . It is straightforward to demonstrate that  $\langle \psi_\varepsilon | \psi_{-\varepsilon} \rangle = 0$ , establishing the linear independence of the spinor space.

In practice, one may choose a non-orthogonal set of functions to solve the large-component free-particle equation (22), and use these two-component solutions to construct an orthonormal set of discrete four-component solutions of the Dirac equation. These four-component solutions span the same linear space as the union of the separate two-component kinetically balanced large- and small-component spaces, provided that *all* positive- and

negative-energy solutions are included. From another point of view, we may use the prescriptions of this section to construct a  $2N$ -dimensional basis set of free-electron Dirac spinors starting with a two-component basis formed by solution of the free-electron Schrödinger equation, and the relationship between the expansion coefficients of a restricted kinetic balance (RKB) set, equation (15). Introduction of an external field causes a rotation amongst the elements of the discrete free-electron spinor set.

## 2.5 Non-relativistic limit

The four-component spinor solutions of the Dirac equation may be re-interpreted as two-component spin orbitals in the formal limit  $c \rightarrow \infty$ . From a computational point of view, however, fixing  $c$  to be more than two or three orders of magnitude larger than its natural value introduces numerical instability in the diagonalization of the basis-set representation of the Dirac equation; the practical upper limit is approximately  $c \rightarrow 20\,000$  au. This method may not be used for high-precision determinations of non-relativistic wave functions, since this value of  $c$  is unable to extinguish completely the fine structure in the core electronic levels of heavy elements.

In order to make direct comparisons between the relativistic formulation and the non-relativistic limit of quantum mechanics, we make use of the identity

$$(\boldsymbol{\sigma} \cdot \mathbf{p})(\boldsymbol{\sigma} \cdot \mathbf{p}) = p^2 \mathbf{I}$$

to write a two-component Schrödinger equation

$$\left(\frac{1}{2}(\boldsymbol{\sigma} \cdot \mathbf{p})(\boldsymbol{\sigma} \cdot \mathbf{p}) + V(\mathbf{r})\right) \psi^L(\mathbf{r}) = E \psi^L(\mathbf{r})$$

involving only the large-component two-spinor,  $\psi^L(\mathbf{r})$ , of a Dirac four-spinor,  $\psi(\mathbf{r})$ . If one adopts a fixed set of Gaussian orbital exponents, this formulation may be used to generate degenerate pairs of non-relativistic spin orbitals in the relativistic large-component basis set, with eigenvalues identical to those that would be generated in the usual scalar orbital formulation of quantum chemistry. These two-component solutions will, in general, involve unitary rotations amongst states of  $\alpha$  and  $\beta$  spin, but can readily be transformed into states of pure spin if required.

All that is required to implement this non-relativistic procedure in addition to the elements of our relativistic electronic structure formalism are the one-electron integrals over the two-component form of the kinetic energy operator,  $(\boldsymbol{\sigma} \cdot \mathbf{p})(\boldsymbol{\sigma} \cdot \mathbf{p})$ . Since the operator  $\boldsymbol{\sigma} \cdot \mathbf{p}$  is translationally invariant, we may use the result from atomic structure

theory

$$\boldsymbol{\sigma} \cdot \mathbf{p} \frac{f(r)}{r} \chi_{\kappa, m}(\vartheta, \varphi) = \frac{i}{r} \left( \frac{d}{dr} + \frac{\kappa}{r} \right) f(r) \chi_{-\kappa, m}(\vartheta, \varphi)$$

where  $f(r)$  is an arbitrary function of  $r$ . Repeated application of this identity to a two-component  $G$ -spinor yields

$$\begin{aligned} (\boldsymbol{\sigma} \cdot \mathbf{p}) (\boldsymbol{\sigma} \cdot \mathbf{p}) \frac{f(r)}{r} \chi_{\kappa, m}(\vartheta, \varphi) &= \frac{1}{r} \left( -\frac{d^2}{dr^2} + \frac{\kappa(\kappa + 1)}{r^2} \right) \\ &\times f(r) \chi_{\kappa, m}(\vartheta, \varphi) \end{aligned}$$

demonstrating that the operator  $p^2$  returns the quantum numbers of the two-component spin-angular functions to their original values, so that the procedure introduces an elementary radial differential operator of the same general form as in non-relativistic atomic theory. In order to make the comparison complete, we need only note that for a given value of orbital angular momentum,  $\ell > 0$ , two fine-structure components may be formed,  $\kappa = \ell$ , and  $\kappa = -\ell - 1$ . In either case,

$$\kappa(\kappa + 1) = \ell(\ell + 1)$$

The special case in which fine structure is absent, corresponding to  $\ell = 0$ , also satisfies the equivalence between scalar and two-spinor non-relativistic formulations, since  $\kappa = -1$  and  $\kappa(\kappa + 1) = 0$ .

The repeated use of the operator  $\boldsymbol{\sigma} \cdot \mathbf{p}$  in spherical polar form serves as a demonstration that it couples only two linearly independent function spaces, labelled by  $\pm\kappa$ . In a Cartesian representation, each application of  $\boldsymbol{\sigma} \cdot \mathbf{p}$  generates spin-orbital functions whose effective angular momentum increases without limit, which might suggest that the kinetic balance prescription is a recursive definition of a basis that can never be closed. In practice, a kinetically balanced Cartesian basis set exhibits a very high degree of linear dependence, which can only be controlled by identifying and retaining only the linearly dependent components. In contrast, the restriction to basis functions of common total angular momentum and opposite parity in the spherical polar representation indicates that this linear dependence is confined to the radial components of the functions. Only two function spaces are involved in the reciprocal relationship between components introduced by  $\boldsymbol{\sigma} \cdot \mathbf{p}$ , so that the general characteristics of basis-set construction in non-relativistic quantum chemistry are retained. In particular, atom-centred large- and small-component spinor basis sets may be augmented by including functions of increasing angular momentum, following a controlled, stepwise procedure, since the inclusion of a large-component function of symmetry-type  $\kappa$  necessitates the addition of a single,

matching basis function of symmetry-type  $-\kappa$  to the small-component basis.

### 3 BASIS SETS FOR ELECTRONIC STRUCTURE CALCULATIONS

#### 3.1 Coulomb Sturmian functions

##### 3.1.1 Definition and properties

The non-relativistic Sturmian functions are defined by Rotenberg<sup>(22,23)</sup> as the normalized solutions of the differential equation

$$\left[ -\frac{d^2}{dr^2} + \frac{l(l+1)}{r^2} - 2E_0 + 2\alpha_{nl}V(r) \right] S_{nl}(r) = 0, \quad 0 < r < \infty \quad (23)$$

vanishing at the endpoints  $r = 0$  and  $r = \infty$ . The integers  $n = 1, 2, 3, \dots$ , and  $l = 0, 1, \dots, n-1$  correspond to the usual non-relativistic quantum numbers and  $E_0$  is a fixed, negative number. The parameter  $\alpha_{nl}$  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

$$\int_0^\infty S_{nl}(r) S_{n'l}(r) V(r) dr = -\delta_{nn'} \quad (24)$$

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 equation (23) in terms of the independent variable  $x = 2\lambda r$ , so that

$$\left[ -\frac{d^2}{dx^2} + \frac{l(l+1)}{x^2} + \frac{1}{4} - \frac{\alpha_{nl}Z}{\lambda x} \right] S_{nl}(x) = 0 \quad (25)$$

having the solutions

$$\begin{aligned} S_{nl}(x) &:= \mathcal{N}_{nl} s_{nl}(x), \quad s_{nl}(x) = e^{-x/2} x^{l+1} L_{n-l-1}^{2l+1}(x), \\ &n = l+1, l+2, \dots \end{aligned} \quad (26)$$

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

$$n = \frac{\alpha_{nl}Z}{\lambda}$$

The  $L_k^\alpha(x)$  are Laguerre polynomials,<sup>(24)</sup> and  $\mathcal{N}_{nl}$  is a normalization constant. We recover the standard solutions

for the radial hydrogenic eigenfunctions when  $\alpha_{nl} = 1$ , giving  $E_0 = -Z^2/2n^2$ . The value of  $\lambda$  (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  $L_k^\alpha(x)$  are listed in, for example, Reference 24. When  $\alpha \geq 1$ , they are orthogonal on  $(0, \infty)$  with weight function  $w(x) = e^{-x}x^\alpha$ , such that

$$\int_0^\infty e^{-x}x^\alpha L_k^\alpha(x)L_{k'}^\alpha(x) dx = \frac{\Gamma(\alpha+k+1)}{k!} \delta_{k,k'} \quad (27)$$

The generating function

$$\Phi^{(\alpha)}(t, s) := L_k^\alpha(s)t^k = (1-t)^{-\alpha-1} \exp\left(\frac{ts}{t-1}\right), \quad |t| < 1 \quad (28)$$

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

$$\begin{aligned} I_l^{(p)}(t, u) &= \int_0^\infty e^{-x}x^{2l+2+p} \Phi^{(2l+1)}(t, x)\Phi^{(2l+1)}(u, x) dx \\ &= \frac{(2l+2+p)!}{(1-tu)^{2l+3+p}} [(1-t)(1-u)]^{1+p} \end{aligned} \quad (29)$$

Two cases have immediate application:

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

$$\langle kl|x^{-1}|k'l \rangle = \frac{(2l+k+1)!}{k!} \delta_{k,k'} \quad (30)$$

which agrees with equation (27), if we put  $\alpha = 2l+1$  and  $k = n-l-1$ .

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 non-vanishing matrix elements for  $k' = k \pm 1$  as well as for  $k' = k$ . We shall normalize

the Sturmians so that

$$\langle kl|k'l \rangle = \delta_{k,k'} \quad (31)$$

so that, remembering  $k = n-l-1$ , equation (30) gives

$$\mathcal{N}_{nl} = \left[ \frac{(n-l-1)!}{2n(n+l)!} \right]^{1/2} \quad (32)$$

The non-vanishing elements of the Gram matrix are thus

$$\begin{aligned} g_{nn}^l &= 1, \quad g_{n,n+1}^l = g_{n+1,n}^l = -\frac{1}{2} \sqrt{1 - \frac{l(l+1)}{n(n+1)}}, \\ &n = l+1, l+2, \dots \end{aligned} \quad (33)$$

The completeness and linear dependence of the Coulomb Sturmian basis set is discussed in detail in Reference 8.

### 3.2 $L$ -spinors

The Dirac  $L$ -spinors are derived from the relativistic analogues of the Coulomb Sturmians. We envisage representing Dirac four-component wave functions as linear combinations

$$\begin{bmatrix} \frac{1}{r} \sum_{n_r} c_{n_r}^L f_{n_r, \kappa}^L(r) \chi_{\kappa m}(\theta, \varphi) \\ i \frac{1}{r} \sum_{n_r} c_{n_r}^S f_{n_r, \kappa}^S(r) \chi_{-\kappa m}(\theta, \varphi) \end{bmatrix} \quad (34)$$

where  $f_{n_r, \kappa}^L(r)$  is a large-component radial function,  $f_{n_r, \kappa}^S(r)$  is a matching small-component radial and  $\chi_{\pm \kappa m}(\theta, \varphi)$  are the usual spin-angular functions. We will consider the  $L$ -spinor basis set in some detail and regard it as a prototype for discussing other classes of spinor basis function.

In terms of the independent variable  $x = 2\lambda r$ ,  $\lambda > 0$  constant, the  $L$ -spinor amplitudes are given by

$$\begin{aligned} f_{n_r, \kappa}^L(x) &= \mathcal{N}_{n_r, \kappa} x^\gamma e^{-x/2} \\ &\times \left\{ -(1 - \delta_{n_r, 0}) L_{n_r-1}^{2\gamma}(x) + \frac{N_{n_r, \kappa} - \kappa}{n_r + 2\gamma} L_{n_r}^{2\gamma}(x) \right\} \end{aligned} \quad (35)$$

$$\begin{aligned} f_{n_r, \kappa}^S(x) &= \mathcal{N}_{n_r, \kappa} x^\gamma e^{-x/2} \\ &\times \left\{ -(1 - \delta_{n_r, 0}) L_{n_r-1}^{2\gamma}(x) - \frac{N_{n_r, \kappa} - \kappa}{n_r + 2\gamma} L_{n_r}^{2\gamma}(x) \right\} \end{aligned} \quad (36)$$

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

$$\gamma = +\sqrt{\kappa^2 - \frac{Z^2}{c^2}}, \quad N_{n_r, \kappa} = +\sqrt{n_r^2 + 2n_r\gamma + \kappa^2} \quad (37)$$

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

$$\begin{bmatrix} \frac{1}{2} - \frac{\alpha_{n_r, \kappa} Z \mu^2}{c} \frac{1}{x} & -\frac{d}{dx} + \frac{\kappa}{x} \\ \frac{d}{dx} + \frac{\kappa}{x} & -\frac{1}{2} - \frac{Z}{\alpha_{n_r, \kappa} \mu^2 c} \frac{1}{x} \end{bmatrix} \times \begin{bmatrix} \mu^{-1} f_{n_r, \kappa}^L(x) \\ \mu f_{n_r, \kappa}^S(x) \end{bmatrix} = 0 \quad (38)$$

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

$$\mu^4 - \frac{2c}{\lambda} \mu^2 + 1 = 0 \quad (39)$$

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  $f_{n_r, \kappa}^L(x) \rightarrow \text{const } S_{n_r, l}(x)$  in the non-relativistic limit  $c \rightarrow \infty$ , which we study below.

The analogue of the non-relativistic energy parameter  $E_0 = -\lambda^2/2$  is given by

$$E_0^R = c^2 \sqrt{1 - \frac{\lambda^2}{c^2}} = c^2 + E_0 + O\left(\frac{1}{c^2}\right) \quad (40)$$

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

$$\alpha_{n_r, \kappa} = \frac{N_{n_r, \kappa} \lambda}{Z}$$

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  $\alpha_{n_r, \kappa} = 1$ .

Note that the solutions of the Sturm–Liouville equation are four-component objects, while the  $L$ -spinors are two-component objects, which define a large- and small-component basis set. These sets are used to determine solutions of a matrix eigenvalue equation by numerical diagonalization of a finite-dimensional representation. If

one employs the solutions of the Sturm–Liouville equation as a basis set, the expansion of a Dirac four-spinor must include all of the solutions, and not just the subset, which are transformed into Coulomb Sturmian functions in the non-relativistic limit.

### 3.2.1 $L$ -spinors in the non-relativistic limit

The large-component  $L$ -spinor amplitudes must coincide with the non-relativistic Coulomb Sturmians in the limit  $c \rightarrow \infty$ , corresponding to instantaneous propagation of electromagnetic disturbances, which may be established directly from the definitions, equations (35 and 36). It is easy to see that  $\gamma \rightarrow |\kappa|$ ,  $N_{n_r, \kappa} \rightarrow n$ , so that for negative values of  $\kappa = -l - 1$ , we have

$$\begin{aligned} f^L(x) &\rightarrow \text{const } x^{l+1} e^{-x/2} \left\{ -(1 - \delta_{n_r, 0}) L_{n_r-1}^{2l+2}(x) \right. \\ &\quad \left. + L_{n_r}^{2l+2}(x) \right\} \\ &= \text{const } x^{l+1} e^{-x/2} L_{n_r}^{2l+1}(x) \end{aligned} \quad (41)$$

using equation (30) of Reference 24 in the second line. Similarly, for positive  $\kappa = l$ , remembering that  $n_r \geq 1$  in this case, we have

$$\begin{aligned} f^L(x) &\rightarrow \text{const } x^l e^{-x/2} \left\{ -(1 - \delta_{n_r, 0}) L_{n_r-1}^{2l}(x) \right. \\ &\quad \left. + \frac{n_r}{n_r + 2l} L_{n_r}^{2l}(x) \right\} \\ &= \text{const } x^l e^{-x/2} \left\{ -(n_r + 2l) L_{n_r-1}^{2l}(x) + n_r L_{n_r}^{2l}(x) \right\} \\ &= \text{const } x^{l+1} e^{-x/2} L_{n_r}^{2l+1}(x) \end{aligned} \quad (42)$$

using equation (31) of Reference 24 in the third line. Since  $n_r = n - l - 1$ , it follows that

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

for both signs of  $\kappa$  in the non-relativistic limit.

A similar analysis shows that, for both signs of  $\kappa$ ,

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

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

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

The kinetic balance criterion guarantees consistency with non-relativistic equations in the limit  $c \rightarrow \infty$ .<sup>(25)</sup>

### 3.2.2 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 equation (38). Multiplying from the left by the adjoint vector  $[\mu^{-1} f_{n'_r, \kappa}^L, \mu f_{n'_r, \kappa}^S]$  and subtracting the result from the corresponding equation with  $n_r$  and  $n'_r$  interchanged gives

$$(\alpha_{n_r, \kappa} - \alpha_{n'_r, \kappa}) \int_0^\infty \left\{ f_{n'_r, \kappa}^L(x) f_{n_r, \kappa}^L(x) - (\alpha_{n_r, \kappa} \alpha_{n'_r, \kappa})^{-1} \times f_{n'_r, \kappa}^S(x) f_{n_r, \kappa}^S(x) \right\} \frac{dx}{x} = 0$$

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 (27) in the non-relativistic 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  $f_{n_r, \kappa}^L(x)$  and  $f_{n_r, \kappa}^S(x)$ , and is given by the equation

$$\begin{aligned} 1 = g_{n_r, n_r}^{(\kappa)} &= \mathcal{N}_{n_r, \kappa}^2 \left[ (1 - \delta_{n_r, 0}) \frac{\Gamma(2\gamma + n_r)}{(n_r - 1)!} \right. \\ &\quad \left. + \left( \frac{N_{n_r, \kappa} - \kappa}{2\gamma + n_r} \right)^2 \frac{\Gamma(2\gamma + n_r + 1)}{(n_r)!} \right] \\ &= \mathcal{N}_{n_r, \kappa}^2 2N_{n_r, \kappa} (N_{n_r, \kappa} - \kappa) \frac{\Gamma(2\gamma + n_r)}{n_r! (2\gamma + n_r)} \end{aligned}$$

so that

$$\mathcal{N}_{n_r, \kappa} = \left[ \frac{n_r! (2\gamma + n_r)}{2N_{n_r, \kappa} (N_{n_r, \kappa} - \kappa) \Gamma(2\gamma + n_r)} \right]^{1/2} \quad (43)$$

In a similar fashion, we can easily show that

$$g_{n_r, n_r}^\kappa = 1 \quad (44)$$

$$g_{n_r, (n_r+1)}^{(\kappa)} = g_{(n_r+1), n_r}^{(\kappa)} = \frac{\eta^T}{2} \times \left[ \frac{(n_r + 1)(2\gamma + n_r + 1)(N_{n_r, \kappa} - \kappa)}{N_{n_r, \kappa} N_{(n_r+1), \kappa} (N_{(n_r+1), \kappa} - \kappa)} \right]^{1/2}, \quad T = L, S$$

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 non-relativistic 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

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

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

$$\rho_N = 1 - \frac{C}{N^2} + O(N^{-3}) \quad (45)$$

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 behaviour is very similar to the Coulomb Sturmians.

### 3.2.3 Completeness of $L$ -spinors

The completeness of  $L$ -spinors may be established in a variety of Hilbert spaces by exploiting a result due to Klahn and Bingel.<sup>(10)</sup>

Following the presentation of Grant and Quiney,<sup>(8)</sup> this Lemma states that if  $\{\varphi_n\}_{n=1}^\infty$  is a complete system in a Hilbert space  $H$ , and  $a_{n\mu}$ , ( $1 \leq \mu \leq n$ ) be a set of arbitrary complex numbers with  $a_{nn} \neq 0$ , 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 result in the context of the  $L$ -spinor set, it is sufficient to note that equations (35 and 36) can be written as

$$\begin{aligned} f_{n_r, \kappa}^T(x) &= a_{n_r, n_r-1} x^\gamma e^{-x/2} L_{n_r-1}^{2\gamma}(x) \\ &\quad + \eta^T a_{n_r, n_r} x^\gamma e^{-x/2} L_{n_r}^{2\gamma}(x), \quad T = L, S \end{aligned}$$

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  $\kappa$ . Also, since  $N_{0, \kappa} = |\kappa|$ , the first non-vanishing  $L$ -spinor for  $\kappa > 0$  has  $n_r = 1$ . We infer that the radial amplitudes appearing in equation (34) have formal  $L$ -spinor expansions and that the  $L$ -spinors are both complete and minimal on the Sobolev spaces  $[W_2^{(p)}(\mathbf{R}^3)]^2$  for  $p = 1, 2$ . This is exactly what is required for the construction of trial wave functions for Rayleigh–Ritz approximations of Dirac four-component spinors.

### 3.2.4 Charge conjugation and $L$ -spinors

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

$$\psi \rightarrow \psi_c = C \bar{\psi}^t \quad (46)$$

where the superscript  $t$  denotes transposition and  $\bar{\psi} = \psi^* \gamma^0$  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

$$Z \leftrightarrow -Z, \quad f_{n,r,\kappa}^L(x) \leftrightarrow f_{n,r,\kappa}^S(x), \quad \kappa \leftrightarrow -\kappa, \quad \mu \leftrightarrow \mu^{-1}$$

in equation (38), the  $L$ -spinors retain the charge-conjugation symmetries of the Dirac eigenfunctions on which they are modelled. Since the mapping  $\mu \leftrightarrow \mu^{-1}$  is equivalent to changing the sign of the energy parameter  $E_0^R = +\sqrt{[1 - (\lambda^2/c^2)]}$ , equation (40), we infer that  $L$ -spinor expansions will be able correctly to represent positron (negative-energy electron) states as well as bound states. This assertion is supported by numerical applications of the  $L$ -spinor basis set presented by Grant and Quiney.<sup>(8)</sup>

## 3.3 $S$ -spinors

$S$ -spinors are the relativistic analogue of Slater functions [Slater-type orbitals (STOs)] and have the form of the most nodeless  $L$ -spinors; they are derived from the radial  $L$ -spinors by taking minimal value of  $n$  for a given value of

$\kappa$ . The explicit form of the radial  $S$ -spinor basis is

$$f_{\kappa,m,i}^T(r) = \begin{cases} N_{\kappa,i}^T r^{\gamma_\kappa} \exp(-\lambda_{\kappa,i} r) & \text{for } \kappa < 0 \\ N_{\kappa,i}^T (A_\kappa^T + \lambda_{\kappa,i} r) r^{\gamma_\kappa} \exp(-\lambda_{\kappa,i} r) & \text{for } \kappa > 0 \end{cases} \quad (47)$$

where  $T = L$  or  $T = S$ , and

$$\begin{aligned} A^L &= \frac{(\kappa + 1 - N)(2\gamma_\kappa + 1)}{2(N - \kappa)} \\ A^S &= -\frac{(\kappa - 1 - N)(2\gamma_\kappa + 1)}{2(N - \kappa)} \\ \gamma_\kappa &= \sqrt{\kappa^2 - (Z/c)^2} \\ N^2 &= \kappa^2 + 2\gamma_\kappa + 1 \end{aligned}$$

for  $\kappa > 0$ . If the  $S$ -spinor basis set contains the exponent corresponding to the ground state of the Dirac–Coulomb problem of a given symmetry, the exact ground state Bohr–Sommerfeld eigenvalue of that symmetry appears as one of the eigenvalues of the matrix eigenvalue problem. In common with the Slater basis set, care must be exercised in the selection of the  $M_\kappa$ -dimensional exponent set,  $\{\lambda_{\kappa,i}; i = 1, 2, \dots, M_\kappa\}$  of each symmetry-type,  $\kappa$ , because of the tendency of the  $S$ -spinor set to suffer from computational linear dependence.

In the  $S$ -spinor scheme, a set of exponents  $\{\lambda_{\kappa,i}\}$ ,  $i = 1, 2, \dots, M_\kappa$  is chosen, where  $M_\kappa$  is the dimension of the radial basis, and generates a matched set of large- and small-component radial functions according to equation (47). If the basis set contains the exponent  $\lambda_m = Z/N_n$ , the corresponding  $S$ -spinor is an exact solution to the hydrogenic problem.

The  $S$ -spinors are a good choice of basis functions for the calculation of many-electron atoms. Efficient methods for evaluating one- and two-electron matrix elements in this basis have been developed, and a large number of calculations have been carried out using  $S$ -spinors.

Of course, the non-analyticity of the  $S$ -spinor functions at  $r = 0$  means that multi-centre integrals over these functions are even more difficult to evaluate than the corresponding quantities involving non-relativistic Slater functions with integer values of the leading exponents. Details of the evaluation of the required one- and two-electron  $S$ -spinor integrals for use in relativistic SCF theory and relativistic many-body theory are presented by Quiney.<sup>(26)</sup>

## 3.4 $G$ -spinors

The  $S$ -spinor functions are well suited for point-nuclear atomic systems, but are practically useless for calculations of molecules, because of the unfeasibility of evaluating

multi-centre integrals. In reality, heavy nuclei are poorly described as point charges, and it has long been common practice in relativistic atomic structure calculations to employ finite nuclear charge distributions. As we have already seen, solutions of the Dirac equation for finite nuclear charge distributions have radial solutions near the nucleus with the same shape as Gaussian functions. We also know from non-relativistic quantum chemistry that multi-centre integrals over Gaussians are straightforward to evaluate, and this type of basis functions thus presents itself as a very attractive choice for relativistic molecular calculations. If we choose  $f_\mu^L(r_{A_\mu})$  to be the radial part of a spherical harmonic Gaussian-type function, which is defined with respect to a coordinate system centred at  $A_\mu$ , then

$$f_\mu^L(r_{A_\mu}) = N_\mu^L r_{A_\mu}^{l_\mu+1} \exp(-\lambda_\mu r_{A_\mu}^2)$$

and the restricted kinetically matched small component is

$$f_\mu^S(r_{A_\mu}) = N_\mu^S [(\kappa_\mu + l_\mu + 1) - 2\lambda_\mu r_{A_\mu}^2] r_{A_\mu}^{l_\mu} \exp(-\lambda_\mu r_{A_\mu}^2)$$

The exponent set  $\{\lambda_\mu; \mu = 1, \dots, N\}$  is chosen in order to provide a compact representation of the spectrum, while providing as complete a representation as possible. This is the usual ‘basis-set problem’ of quantum chemistry, and one may employ either energy-optimized exponents or even-tempered sequences of exponents.

Using the above-mentioned definition of the radial part of the  $G$ -spinor, we can now define our two-component  $G$ -spinors, which we will use in later work

$$M[L, \mu, \mathbf{r}_{A_\mu}] = \frac{N_\mu^L}{r_{A_\mu}} f_\mu^L(r_{A_\mu}) \chi_{\kappa_\mu, m_\mu}(\theta_{A_\mu}, \varphi_{A_\mu})$$

$$M[S, \mu, \mathbf{r}_{A_\mu}] = i \frac{N_\mu^S}{r_{A_\mu}} f_\mu^S(r_{A_\mu}) \chi_{-\kappa_\mu, m_\mu}(\theta_{A_\mu}, \varphi_{A_\mu})$$

The quantum numbers not shown explicitly in the  $G$ -spinor are collected in the multi-index,  $\mu$ , where

$$(\kappa, m_j, \lambda, \mathbf{A}) \mapsto \mu$$

and the normalization condition

$$\int M^\dagger[T, \mu, \mathbf{r}] M[T, \mu, \mathbf{r}] d\mathbf{r} = 1$$

for  $T = L, S$ . Using the standard integral

$$\int_0^\infty x^{2n} \exp(-\lambda x^2) dx = \frac{\Gamma(n+1/2)}{2\lambda^{n+1/2}}$$

$\lambda > 0 \quad n = 0, 1, \dots$

and the orthonormality of the spin-angular functions  $\chi_{\kappa, j, m}$ , one obtains the normalization constants

$$N_\mu^L = \left[ \frac{2(2\lambda_\mu)^{l_\mu+3/2}}{\Gamma(l_\mu+3/2)} \right]^{1/2} \quad N_\mu^S = \left[ \frac{2(2\lambda_\mu)^{l_\mu+1/2}}{\Gamma(l_\mu+5/2)} \right]^{1/2} \quad (48)$$

An interesting feature of this choice of matched basis and normalization is that  $N_\mu^S$  depends only on  $l_\mu$ , independent of the sign of  $\kappa_\mu$ , despite the considerable differences in functional form. It is convenient to define the normalization constant of a product density by the notation

$$N_{\mu\nu}^{TT'} = N_\mu^T N_\nu^{T'}$$

Details of the evaluation of the required one- and two-electron  $G$ -spinor integrals for use in relativistic SCF theory and relativistic many-body theory are presented by Quiney.<sup>(26)</sup>

The  $G$ -spinor set is widely used in four-component relativistic molecular structure calculations. In some formulations, a kinetically balanced set of Gaussian spin orbitals is used, mainly in order to exploit existing quantum-chemical technology based on scalar functions. In this case, the relativistic properties are transferred to the construction of the matrix representations of the Dirac operator, rather than to the construction of a basis set with the double-group transformational properties characteristic of a spinor basis. Ultimately, however, these two approaches are equivalent, and are connected through a well-defined sequence of linear transformations.

### 3.5 Drake and Goldman basis

In a series of articles, Drake and Goldman<sup>(27-31)</sup> introduced a generalization of the Slater-type function for use in relativistic basis-set expansions. Their original choice of normalized radial basis set of dimension  $M$  was<sup>(27)</sup>

$$f_i^T(r) = \left[ \frac{(2\lambda)^{2\gamma+2i+1}}{\Gamma(2\gamma+2i+1)} \right]^{1/2} r^{\gamma+i} \exp(-\lambda r) \quad (49)$$

where  $T = L$  or  $T = S$ ,  $i = 0, 1, \dots, (M-1)$  and  $\gamma = \sqrt{[\kappa^2 - (Z/c)^2]}$ . Unlike the  $S$ -spinor basis sets, the exponential parameter  $\lambda$  is fixed and the set is enlarged by introducing higher powers of  $r$ .

When used in this form, the matrix representations of the Dirac operator exhibit a problem with spurious roots for  $\kappa > 0$ . This is readily seen in the case of a minimal basis-set representation, in which the lowest positive-energy solution for  $\kappa = +1$  is degenerate with the lowest positive-energy solution for  $\kappa = -1$ . The proposed solution to this problem

introduces new radial basis functions that satisfy energy-independent boundary conditions at  $r = 0$ . These functions are defined by

$$f_i^L(r) = \left[ \frac{2\gamma + i}{r} + \frac{Z}{\kappa} - \lambda \right] r^{\gamma+i} \exp(-\lambda r) \quad (50)$$

$$f_i^S(r) = r^{\gamma+i} \exp(-\lambda r) \quad (51)$$

for  $i = 1, 2, \dots, M$ . In the case  $\kappa < 0$ , an extra-large-component radial function is added of the form

$$f_0^L(r) = r^\gamma \exp(-\lambda r) \quad (52)$$

Use of this basis set avoids the appearance of the spurious eigenvalue for  $\kappa > 0$ , and it is observed that the eigenvalues appear to satisfy a generalized Undheim–Hylleraas–Macdonald theorem, in the sense that the positive-energy eigenvalues approach the exact values from above.<sup>(30)</sup>

The use of basis functions of this general type, in which the members of the set are generated by including higher powers of  $r$ , introduce some quite serious practical problems. The most significant of these is the onset of severe computational linear dependence, which limits the dimension of the basis to about  $M = 14$ , if double-precision arithmetic is used.

In order to circumvent the problem with computational linear dependence, Goldman later introduced the generalized Laguerre basis set,<sup>(28,32,33)</sup> whose radial components are defined by

$$f_i^T(r) = \frac{(2\lambda)^{1/2}}{P_i(\gamma)} (2\lambda r)^\gamma L_i^{2\gamma}(2\lambda r) \exp(-\lambda r) \quad (53)$$

where  $i = 0, 1, 2, \dots, (M - 1)$ ,  $L_i^{2\gamma}(z)$  is a generalized Laguerre polynomial, and

$$P_i(\gamma) = \left[ \frac{\Gamma(2\gamma + i + 1)}{i!} \right]^{1/2} \quad (54)$$

The clear advantage of this basis set is that it spans the same linear function space as the original basis set, but the radial basis functions have been formed into orthonormal linear combinations, eliminating the problem of linear dependence. In order to derive computational benefit from this basis set, however, it is necessary to evaluate matrix elements using the analytical properties of the Laguerre polynomials, rather than through term-by-term expansions into products. Essentially the same techniques are used to evaluate these matrix elements as have already been described for the  $L$ -spinor set. While this generalized Laguerre basis set exhibits no variational problems for  $\kappa < 0$ , the appearance of spurious solutions for  $\kappa > 0$  must be eliminated

by forming linear combinations of Laguerre polynomial functions that satisfy energy-independent boundary conditions or by the introduction of additional functions. The approach in which the Laguerre polynomials are combined together to eliminate variational problems leads directly to the construction of the  $L$ -spinor basis set, which exhibits, in addition to satisfactory variational properties, the correct transformational properties under the charge-conjugation symmetries  $\kappa \rightarrow -\kappa$  and the Sturmian basis set as the non-relativistic limit of the large-component set, irrespective of the sign of  $\kappa$ .

### 3.6 B-splines

The effective range over which atomic phenomena take place is finite, and so discretization of the spectrum may be achieved by solving the Dirac equation in a cavity, subject to satisfactory boundary conditions. Johnson and co-workers<sup>(34,35)</sup> have chosen a basis of piecewise polynomial basis functions and confined the atomic pseudo-spectrum using the *MIT bag model*. Typically, the radius of the cavity is taken to be  $R = 40/Z_{\text{ion}}$  au where  $Z_{\text{ion}} = Z - N + 1$ ,  $Z$  is the nuclear charge and  $N$  is the number of electrons. The MIT bag model imposes the condition  $P(r) = Q(r)$  at  $r = R$ , which is equivalent to the requirement that there is no net flow of current across the boundary of the cavity. Strictly, it is not permissible to impose the simpler condition  $P(r) = Q(r) = 0$  since a current of positrons will be reflected from the boundary because of the infinitely large effective potential (Klein's paradox). However, the error caused by this is apparently negligible, as we shall see when we consider finite-difference discretization.

A basis of  $n$  B-splines of order  $k$  is generated in the range  $0 < r < R$ . The range is divided into segments, the endpoints of which define a knot sequence,  $\{t_i\}$ ,  $i = 1, 2, n + k$ . The B-splines of order  $k$  on this known sequence are defined recursively by the relations

$$B_{i,1}(r) = \begin{cases} 1 & t_i \leq r \leq t_{i+1} \\ 0 & \text{otherwise} \end{cases}$$

and

$$B_{i,k}(r) = \frac{r - t_i}{t_{i+k-1} - t_i} B_{i,k-1}(r) + \frac{t_{i+k} - r}{t_{i+k} - t_{i+1}} B_{i+1,k-1}(r)$$

The functions  $B_{i,k}(r)$  are piecewise polynomials of degree  $k - 1$  inside the interval  $t_i \leq r \leq t_{i+k}$  and vanish outside this interval. The knots defining the endpoints have multiplicity  $k$  at the endpoints  $r = 0$  and  $r = R$ , in the sense that  $t_1 = t_2 = \dots = t_k = 0$  and  $t_{n+1} = t_{n+2} = \dots = t_{n+k} = R$ . For atomic problems, the knots  $t_{k+1}, t_{k+2}, \dots, t_n$  are distributed on an exponential grid so that the basis function

density is greater near the nucleus than it is near the boundary.

Since the B-spline functions vanish outside a finite interval, the overlap matrix is banded. The choice of an exponential distribution of knots ensures that the basis is free of computational linear dependence, so that one can keep adding basis functions to the set in order to achieve the required level of accuracy.

The basis functions consist only of local polynomials, so that differentiation and integration of these functions may be performed exactly on a radial grid for most properties of interest. In practice, the large number of basis functions and the banded structure makes this approach rather complicated. A satisfactory scheme may be developed in which the B-splines mainly perform the role of an interpolation scheme, which facilitates the use of conventional numerical quadrature.

### 3.7 Finite-difference discretization

A method has been developed by Salomonson and Öster,<sup>(36)</sup> which discretizes the Dirac spectrum within a cavity of finite extent. One again encounters a variational collapse problem, the solution of which involves the matching of large- and small-component basis functions.

It is easy to see how the problem arises if we consider a one-dimensional free-particle Dirac equation

$$\begin{bmatrix} mc^2 & -c \frac{d}{dr} \\ c \frac{d}{dr} & -mc^2 \end{bmatrix} \begin{bmatrix} P(r) \\ Q(r) \end{bmatrix} = \varepsilon \begin{bmatrix} P(r) \\ Q(r) \end{bmatrix}$$

If we use the central difference formula for the first derivative

$$P(x_i) = \frac{1}{2h} [P(x_{i+1}) - P(x_{i-1})]$$

we may eliminate  $Q(x_i)$  to find that

$$\begin{aligned} & -\frac{1}{2m} \left( \frac{P(x_{i+2}) - 2P(x_i) + P(x_{i-1}))}{4h^2} \right) \\ & = \varepsilon \left( 1 + \frac{\varepsilon}{2mc^2} \right) P(x_i) \end{aligned}$$

Although this appears to be the standard form for a finite-difference representation of the second derivative, there is a problem – the finite-difference equations for odd  $i$  and even  $i$  are completely decoupled.

In order to couple the finite-difference equations for the large and small components, Salomonson and Öster interleaved the large- and small-component grids. In the

non-relativistic or free-particle limits, this ensures that one correctly solves two independent finite-difference representations of second-order differential equations, and that for finite  $c$  and for external field problems, the finite-difference representations of the components have the proper nodal structures, and that the spectrum is free of spurious solutions.

On a grid of  $2N$  points,  $P(r)$  and  $Q(r)$  are discretized on alternating lattice sites  $\{x_k\}$  such that we solve for

$$[P(x_1), Q(x_2), P(x_3), \dots, P(x_{2N-1}), Q(x_{2N})]$$

In this scheme,

$$\begin{aligned} \left. \frac{dP(x)}{dx} \right|_{x=x_i} &= \frac{1}{h} [P(x_{i+1}) - P(x_{i-1})] \\ & \quad i = 2, 4, 6, \dots, 2N \\ \left. \frac{dQ(x)}{dx} \right|_{x=x_j} &= \frac{1}{h} [Q(x_{j+1}) - Q(x_{j-1})] \\ & \quad j = 1, 3, 5, \dots, 2N \end{aligned}$$

In this case,  $h = x_{i+1} - x_{i-1}$  for all  $i$  in the grid.

Following practice familiar from atomic structure, the grid and the solutions are transformed according to

$$\begin{aligned} r &= e^x \\ P(r) &= \frac{1}{\sqrt{r}} P(x) \quad Q(r) = \frac{1}{\sqrt{r}} Q(x) \end{aligned}$$

which ensures that there is a high density of points near the nucleus and that the finite-difference representation is symmetric. A typical grid is defined on  $-8 \leq x \leq 3$ .

The transformed radial equations are

$$\begin{bmatrix} V(r) & c \left( -\frac{1}{\sqrt{r}} \frac{d}{dx} \frac{1}{\sqrt{r}} + \frac{\kappa}{\sqrt{r}} \frac{1}{\sqrt{r}} \right) \\ c \left( \frac{1}{\sqrt{r}} \frac{d}{dx} \frac{1}{\sqrt{r}} + \frac{\kappa}{\sqrt{r}} \frac{1}{\sqrt{r}} \right) & V(r) - 2mc^2 \end{bmatrix} \times \begin{bmatrix} P(x) \\ Q(x) \end{bmatrix} = \varepsilon \begin{bmatrix} P(x) \\ Q(x) \end{bmatrix}$$

The boundary condition near  $r = 0$  is determined by the power series expansions

$$\begin{aligned} P(x) &\simeq r^{\gamma+1/2} \\ & \quad + \frac{2[\gamma + \kappa - (Z\alpha)^2]}{(Z\alpha)^2(2\gamma + 1)} r^{\gamma+3/2} + \dots \\ Q(x) &\simeq \frac{\gamma + \kappa}{Z\alpha} r^{\gamma+1/2} \\ & \quad - \frac{2[\gamma + \kappa]}{\alpha(2\gamma + 1)} r^{\gamma+3/2} + \dots \end{aligned}$$

This fixes the relationship between  $P(r)$  and  $Q(r)$ . At the boundary, one introduces some ‘ghost’ points outside the boundary, at which the amplitudes are zero. This again determines the relationship between  $P(r)$  and  $Q(r)$  at the points just inside the boundary, which is sufficient to complete the discretization.

In practice, Salomonson and Öster implemented a scheme that employed finite-difference interpolation and differentiation formulae correct to  $O(h^6)$ .

$$\mathbf{D} = \frac{c}{1920h} \begin{bmatrix} -\frac{2250}{\sqrt{r_2r_1}} & \frac{2250}{\sqrt{r_2r_3}} & -\frac{125}{\sqrt{r_2r_5}} & \frac{9}{\sqrt{r_2r_7}} & 0 & \dots \\ \frac{125}{\sqrt{r_4r_1}} & -\frac{2250}{\sqrt{r_4r_3}} & \frac{2250}{\sqrt{r_4r_5}} & -\frac{125}{\sqrt{r_4r_7}} & -\frac{9}{\sqrt{r_4r_9}} & 0 & \dots \\ -\frac{9}{\sqrt{r_6r_1}} & \frac{125}{\sqrt{r_6r_3}} & -\frac{2250}{\sqrt{r_6r_5}} & \frac{2250}{\sqrt{r_6r_7}} & -\frac{125}{\sqrt{r_6r_9}} & \frac{9}{\sqrt{r_6r_{11}}} & 0 & \dots \\ 0 & -\frac{9}{\sqrt{r_8r_3}} & \frac{125}{\sqrt{r_8r_5}} & -\frac{2250}{\sqrt{r_8r_7}} & \frac{2250}{\sqrt{r_8r_9}} & -\frac{125}{\sqrt{r_8r_{11}}} & \frac{9}{\sqrt{r_8r_{13}}} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots \end{bmatrix}$$

$$\mathbf{K} = \frac{\kappa}{256} \begin{bmatrix} \frac{150}{\sqrt{r_2r_1}} & \frac{150}{\sqrt{r_2r_3}} & -\frac{25}{\sqrt{r_2r_5}} & \frac{3}{\sqrt{r_2r_7}} & 0 & \dots \\ -\frac{25}{\sqrt{r_4r_1}} & \frac{150}{\sqrt{r_4r_3}} & \frac{150}{\sqrt{r_4r_5}} & -\frac{25}{\sqrt{r_4r_7}} & \frac{3}{\sqrt{r_4r_9}} & 0 & \dots \\ \frac{3}{\sqrt{r_6r_1}} & -\frac{25}{\sqrt{r_6r_3}} & \frac{150}{\sqrt{r_6r_5}} & \frac{150}{\sqrt{r_6r_7}} & -\frac{25}{\sqrt{r_6r_9}} & \frac{3}{\sqrt{r_6r_{11}}} & 0 & \dots \\ 0 & \frac{3}{\sqrt{r_8r_3}} & -\frac{25}{\sqrt{r_8r_5}} & \frac{150}{\sqrt{r_8r_7}} & \frac{150}{\sqrt{r_8r_9}} & -\frac{25}{\sqrt{r_8r_{11}}} & \frac{3}{\sqrt{r_8r_{13}}} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots \end{bmatrix}$$

$$\begin{aligned}
 y(x) &= \frac{1}{256} \left[ 3y\left(x - \frac{5}{2}h\right) - 25y\left(x - \frac{3}{2}h\right) \right. \\
 &\quad + 150y\left(x - \frac{1}{2}h\right) + 150y\left(x + \frac{1}{2}h\right) \\
 &\quad \left. - 25y\left(x + \frac{3}{2}h\right) 3y\left(x + \frac{5}{2}h\right) \right] + O(h^6) \\
 y'(x) &= \frac{1}{1920h} \left[ -9y\left(x - \frac{5}{2}h\right) + 125y\left(x - \frac{3}{2}h\right) \right. \\
 &\quad - 2250y\left(x - \frac{1}{2}h\right) + 2250y\left(x + \frac{1}{2}h\right) \\
 &\quad \left. - 125y\left(x + \frac{3}{2}h\right) + 9y\left(x + \frac{5}{2}h\right) \right] + O(h^6)
 \end{aligned}$$

The finite-difference equations then have the form

$$\begin{bmatrix} \mathbf{A} & \mathbf{D}^T + \mathbf{K}^T \\ \mathbf{D} + \mathbf{K} & \mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{P} \\ \mathbf{Q} \end{bmatrix} = \varepsilon \begin{bmatrix} \mathbf{P} \\ \mathbf{Q} \end{bmatrix}$$

where

$$A_{ii} = -\frac{Z}{r_i} \quad i = 2n - 1$$

$$B_{jj} = -\frac{Z}{r_j} - 2mc^2 \quad j = 2n$$

and  $n = 1, 2, \dots, N$ . In this way, the interleaving of the components is implicit in the numerical algorithms, but the discretization points for the large and small components appear together in the matrices.

The coupling between components is achieved in the matrices  $\mathbf{D}$  and  $\mathbf{K}$ , which for their chosen interpolation and differentiation formulae take the explicit forms

Since the matrices  $\mathbf{D}$  and  $\mathbf{K}$  involve the coupling between  $P(r)$  and  $Q(r)$ , then they are constructed from product involving one odd- and one even-numbered point. This should be compared with the diagonal blocks, which involve pairs of odd- or even-numbered points. These off-diagonal blocks are not symmetric, but since they appear as  $\mathbf{D}^T$  and  $\mathbf{K}^T$  elsewhere, the Dirac matrix is symmetric, and yields real eigenvalues.

#### 4 APPLICATIONS TO RELATIVISTIC ATOMS

This section reviews some applications of the preceding theory of Rayleigh–Ritz methods to a range of physical problems. With regard to the essential features of the method, it is sufficient to demonstrate the behaviour of the scheme using simple atomic models. No special features arise in complex electronic configurations or in the calculation of molecular electronic structures other than the technical complexity of evaluating matrix elements, and the basis-set superposition problem.

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, in which the use of either  $S$ -spinor or  $G$ -spinor sets is adopted. For these nearly nodeless radial basis functions, commonly occurring matrix elements may be evaluated using conventional methods borrowed from quantum chemistry, and the reader is referred to the article by Saunders in the handbook for a comprehensive review of these techniques (**Chapter 29, this Volume**).

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

$$(a)_0 = 1, \quad (a)_k = (a+k-1)(a)_{k-1} \\ = a(a+1)\cdots(a+k-1), \quad k \geq 1$$

We also use the notations

$$G_k(a) = \frac{(a+1)_k}{k!}, \quad H_{mn}(a) = \sum_{k=0}^{\min m,n} G_k(a-1) \quad (55)$$

Then,

$$V_{mn}^T = -Z\lambda \left[ \frac{m!n!(a+m)(a+n)}{N_{m\kappa}N_{n\kappa}(N_{m\kappa}-\kappa)(N_{n\kappa}-\kappa)(a)_m(a)_n} \right]^{1/2} \\ \times \left\{ H_{m-1,n-1}(a) + \eta^T \frac{N_{n\kappa}-\kappa}{n+a} H_{m-1,n}(a) \right. \\ \left. + \eta^T \frac{N_{m\kappa}-\kappa}{m+a} H_{m,n-1}(a) \right. \\ \left. + \frac{N_{m\kappa}-\kappa}{m+a} \frac{N_{n\kappa}-\kappa}{n+a} H_{mn}(a) \right\} \quad (56)$$

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

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

$$\left. \begin{aligned} & - G_{m-1,n-1}(a) - \frac{N_{m\kappa}-\kappa}{m+a} G_{m,n-1}(a) \\ & + \frac{N_{n\kappa}-\kappa}{n+a} G_{m-1,n}(a) + \frac{N_{m\kappa}-\kappa}{m+a} \frac{N_{n\kappa}-\kappa}{n+a} G_{mn}(a) \end{aligned} \right\} \quad (57)$$

whilst the symmetric tridiagonal Gram matrices  $\mathbf{S}^{TT}$  are related to the expressions (44) by

$$S_{ij}^{TT} = \frac{g_{ij}^\kappa}{2\lambda}$$

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

## 4.1 Hydrogenic atoms

The rate of convergence of the bound-state eigenvalues of hydrogenic ions with the  $S$ -spinor or  $L$ -spinor basis sets is comparable with that observed in corresponding non-relativistic calculations using Slater or Sturmian basis functions, respectively. This is because both of these relativistic basis sets correctly reproduce the weak singularity of the radial spinors in the neighbour of the nucleus by the explicit inclusion of the leading-order radial factor  $r^\gamma$  in both large- and small-component sets. Once this non-analytic behaviour is extracted, the exponential and polynomial parts of the basis are able to reproduce the spinors to high accuracy. If the exponential parameter is chosen so that a large- and small-component pair of  $L$ -spinors coincides exactly with a bound-state function, or a large- or small-component pair of  $S$ -spinors coincides exactly with the ground state of that symmetry, the corresponding exact eigenvalue of that state will appear in the eigenvalue spectrum of the matrix representation.

If a point nucleus model is adopted, the efficiency of the  $G$ -spinor basis set deteriorates rapidly as the nuclear charge is increased, because the basis set is able to model the features of the spinors in the neighbour of the nucleus only if a large number of basis functions with large exponents is included. This is a problem that is distinct from the difficulty with which the spherical Gaussian basis set is able to reproduce the Coulomb cusp in non-relativistic orbital functions. In the relativistic case, the linear expansion method is attempting to describe a non-analytic feature of the spinors, leading to very slow convergence. An effective solution to this problem is an essential ingredient in the construction of models of relativistic ab initio quantum chemistry because Gaussian basis function amplitudes are

required in order to evaluate the multi-centre interaction matrix elements. Finite nuclear models are now widely adopted whenever the  $G$ -spinor basis set is used, because the non-analytic behaviour of the spinors is modified at the nucleus. This imparts the remarkable property that the solutions in this region are well approximated by a single Gaussian function and that the basis set no longer requires the open-ended addition of functions with huge exponent values in order to represent the Coulomb cusp, which is merely an artefact of an unphysical model. The use of a  $G$ -spinor basis set is now almost always associated with the use of a finite nuclear model, while the  $L$ -spinor and  $S$ -spinor basis functions, or the Drake and Goldman family of basis sets, are used exclusively with point-nuclear models.

## 4.2 Coulomb perturbation model 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 non-relativistically;<sup>(37)</sup> 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 non-relativistic 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  $k \geq 3$ , 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. Quiney *et al.*<sup>(38)</sup> demonstrated that all these difficulties could be avoided if we solved the Schrödinger equation for the non-relativistic hydrogenic atom using simple finite basis sets of either exponential or Gaussian form. The more difficult Dirac case was studied in Reference 39; the higher-order diagrams of order  $k \geq 3$  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  $Z'$ .

The energy of the perturbed system at second order in Rayleigh–Schrödinger energy component is given by

$$\begin{aligned} \mathcal{E} = & \langle \Phi | \mathcal{H}_0 | \Phi \rangle + \langle \Phi | \mathcal{H}_1 | \Phi \rangle \\ & + \sum_{m \neq 0} \frac{\langle \Phi_0 | H_1 | \Phi_m \rangle \langle \Phi_m | H_1 | \Phi_0 \rangle}{E_0^{(0)} - E_m^{(0)}} \end{aligned} \quad (58)$$

and the exact energy in second-order Brillouin–Wigner perturbation theory is given by

$$\mathcal{E} = \langle \Phi | \mathcal{H}_0 | \Phi \rangle + \langle \Phi | \mathcal{H}_1 | \Phi \rangle + \sum_{k \neq 0} \frac{\langle \Phi | \mathcal{H}_1 | \Phi_k \rangle \langle \Phi_k | \mathcal{V} | \Phi \rangle}{\mathcal{E} - E_k} \quad (59)$$

where

$$\langle \Phi_k | \mathcal{V} | \Phi \rangle = \langle \Phi_k | \mathcal{H}_1 | \Phi \rangle + \sum_{p \neq 0} \frac{\langle \Phi_k | \mathcal{H}_1 | \Phi_p \rangle \langle \Phi_p | \mathcal{V} | \Phi \rangle}{\mathcal{E} - E_p}, \quad \forall k \quad (60)$$

The reaction operator is denoted  $\mathcal{V}$  and equation (60) is the Lippmann–Schwinger (or Dyson) equation. The higher-order components of the Brillouin–Wigner energy expansion,  $E_0^{(k)}(\mathcal{E})$ , may be evaluated using repeated matrix multiplication. If we define the matrix elements

$$\begin{aligned} W_{mn}(\mathcal{E}) &= \frac{\langle \Phi_m | H_1 | \Phi_n \rangle}{\mathcal{E} - E_n^0} \\ V_m &= \langle \Phi_0 | H_1 | \Phi_m \rangle \\ U_m &= W_{0m}^* \end{aligned}$$

then for  $k \geq 2$ , we have

$$E_0^{(k)}(\mathcal{E}) = \mathbf{V}^\dagger \{\mathbf{W}(\mathcal{E})\}^{k-2} \mathbf{U}$$

The value of  $\mathcal{E}$  appearing in the denominator of  $W_{mn}(\mathcal{E})$  may be the exact ‘dressed value’,  $\mathcal{E}_0$ , or it may be some intermediate approximation to  $\mathcal{E}_0$  obtained at finite order in perturbation theory. For the relativistic and non-relativistic model problems, we know the value  $\mathcal{E}_0$  exactly, but in general, we must determine its value self-consistently. In Rayleigh–Schrödinger perturbation theory, we may determine the energy components by evaluating the derivatives of the energy with respect to the perturbation, but we may not so readily determine the energy components at each order in the relativistic theory.

Grant and Quiney<sup>(8)</sup> evaluated the energy components of the Rayleigh–Schrödinger perturbation series for the Coulomb perturbation  $H_1 = -1/r$ , using an  $L$ -spinor set, computed with block dimension  $N = 100$  taking as fixed the exponential parameter as  $\lambda = Z$ . They demonstrated that the negative-energy state contribution in the second-order energy component 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. This study also provides a convincing demonstration that one may perform the sum-over-states required in relativistic perturbation theories by summing over the discrete representation provided by the algebraic

approximation and that this representation tends towards completeness in a controllable fashion as the basis set is enlarged. Attempts to solve many-electron problems using methods that attempt to project out negative-energy contributions, as noted by Kutzelnigg,<sup>(6)</sup> always introduce unquantifiable errors as well as formidable technical complications, which are absent in a more direct approach. A study of the energy components of the same model problem in Brillouin–Wigner perturbation theory has also recently been presented by Quiney *et al.*<sup>(40)</sup> in both relativistic and non-relativistic approximations. They found that one may sum the energy through sufficiently high order in Brillouin–Wigner perturbation theory using Sturmian or  $L$ -spinor sets to reproduce the total dressed energy,  $\mathcal{E}$ , to machine accuracy. It was also found that the energy components were very well approximated by a geometric progression, so that reliable estimates of  $\mathcal{E}$  could be obtained by extrapolations based on a few terms in the series.

### 4.3 Relativistic sum rules

Goldman and Drake<sup>(29)</sup> have derived relativistic generalizations of the one-electron electric dipole oscillator-strength sum rules

$$S_k = \sum_n (E_n - E_0)^k |\langle \psi_0 | \mathbf{r} | \psi_n \rangle|^2 \quad (61)$$

where the sum over  $n$  implies a summation over discrete states and an integration over both positive- and negative-energy continuum states. The evaluation of these using the algebraic approximation provides a useful test of the accuracy and efficiency of the method using a given choice functional form for the basis set, or using particular values of the auxiliary parameters, such as the exponent values and basis-set dimension, which define the set.

In the non-relativistic limit of one-electron point-nuclear systems, these sum rules take the simple forms

$$S_0^{NR} = \langle \psi_0 | r^2 | \psi_n \rangle \quad (62)$$

$$S_1^{NR} = \frac{3}{2} \quad (63)$$

$$S_2^{NR} = \langle \psi_0 | \nabla^2 | \psi_n \rangle \quad (64)$$

$$S_3^{NR} = \langle \psi_0 | \delta(\mathbf{r}) | \psi_n \rangle \quad (65)$$

where  $Z$  is the nuclear charge. The result  $S_1^{NR}$  is the Thomas–Reiche–Kuhn sum rule, while  $S_0^{NR}$  follows immediately from the completeness of the intermediate states,  $\{\psi_n\}$ . For  $k = -1$ , the sum rule yields the dipole polarizability,  $\alpha_d$ .

In the relativistic case, the derivation of analytical expressions of the sum rules is somewhat involved and may be found in Reference 29. The results are

$$S_0^D = \langle \psi_0 | r^2 | \psi_n \rangle \quad (66)$$

$$S_1^D = 0 \quad (67)$$

$$S_2^D = 3c^2 \quad (68)$$

$$S_3^D = 6c^2 \langle \psi_0 | V - E_0 | \psi_n \rangle + 2c^2 \langle \psi_0 | r \frac{dV}{dr} | \psi_n \rangle \quad (69)$$

Only the non-relativistic sum rule  $S_0^R$  retains the form of its non-relativistic analogue,  $S_0^R$ , which occurs because of closure on the complete set of spinors,  $\psi_n$ . For  $k > 0$ , the sum rules do not resemble the non-relativistic results because of the strong enhancement of the matrix elements by the energy factor,  $(E_0 - E_n)^k$ , which emphasizes the role of the negative-energy states. As a test of basis-set quality, the evaluation of these sum rules provides a useful test as to whether the complete spectrum has been partitioned correctly into its positive- and negative-energy components. For the sum rule  $S_k^D$ , however, the contribution from the negative-energy states are diminished by a scale factor of  $(Z/c)^4$ , one factor of  $(Z/c)^2$  coming from the matrix elements and the other from the energy denominator.

The second-order static dipole polarizability  $\alpha_{zz}$  for the state  $|\psi_0\rangle$ , energy  $\epsilon_0$ , of a hydrogenic atom is given by the formula

$$\alpha_{zz} = 2 \sum_{n \neq 0} \frac{\langle \psi_0 | z | \psi_n \rangle \langle \psi_n | z | \psi_0 \rangle}{\epsilon_n - \epsilon_0} \quad (70)$$

where the restriction  $n \neq 0$  excludes  $\psi_0$  from the sum over states. In the Dirac case, the sum includes both positive- and negative-energy eigenstates. For brevity, we treat only the case in which  $\psi_0$  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  $m_j = +1/2$ . 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

$$\alpha_d = \frac{2}{9} (\Delta_{+1} + 2\Delta_{-2}) \quad (71)$$

Szmytkowski<sup>(41)</sup> has presented analytical expressions for the quantities  $\Delta_{+1}$  and  $\Delta_{-2}$

$$\Delta_{+1} = \frac{\gamma_1(\gamma_1 + 1)(2\gamma_1 + 1)(4\gamma_1 + 5)}{8Z^4}$$

$$\Delta_{-2} = \left[ \frac{\Gamma(\gamma_2 + \gamma_1 + 2)}{\Gamma(\gamma_2 - \gamma_1 - 1)} \right]^2 \frac{1}{8Z^4\Gamma(2\gamma_1 + 1)} \\ \times \sum_{n=0}^{\infty} \frac{[\Gamma(\gamma_2\gamma_1 + n - 2)]^2}{n!\Gamma(2\gamma_2 + n + 1)} \frac{f_n(\gamma_1, \gamma_2)}{(\gamma_2 - \gamma_1 + n)}$$

where

$$f_n(\gamma_1, \gamma_2) = 3(1 - \gamma_1^2)(n + \gamma_2)^2 + (4\gamma_1^3 + 8\gamma_1^2 + \gamma_1 - 12) \\ \times (n + \gamma_2^2) - (\gamma_1^4 + 8\gamma_1^3 + \gamma_1^2 - 12)$$

These serve as standards against which we are able to compare numerical results obtained using relativistic finite basis-set techniques. Grant and Quiney<sup>(8)</sup> conducted a numerical investigation of the evaluation of the dipole polarizability by computing the sums

$$\Delta_{\kappa} = \sum_{n \neq 0}^+ \frac{(0|r|n^+\kappa)(n^+\kappa|r|0)}{\epsilon_n^+ - \epsilon_0^+} + \sum_n^- \frac{(0|r|n^-\kappa)(n^-\kappa|r|0)}{\epsilon_n^- - \epsilon_0^+} \quad (72)$$

where superscripts  $+/-$  designate the two branches of the pseudo-spectrum 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,  $\psi_0$ , is represented exactly by the  $L$ -spinors with  $n_r = 0$ ,  $\kappa = -1$ . The basis-set dimensions,  $N_{\kappa}$ , were adjusted in this study so that the computed sums matched the exact values to within an accuracy  $\eta$ , defined by

$$Z^4 |\Delta_{\kappa} - \Delta_{\kappa}^{\text{analytic}}| \leq \eta \quad (73)$$

The value  $\eta = 10^{-6}$  ensures agreement of the numerical values with the analytic values to six significant figures. It was found that any desired agreement up to 12 figures could be obtained by increasing the basis-set dimension. The results obtained by this spectral expansion were relatively insensitive to the choice of  $\lambda$  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  $\lambda$  in each application.

#### 4.4 Gauge invariance of matrix methods

Gauge invariance of the Dirac equation implies that under the transformation

$$\mathbf{A} \rightarrow \mathbf{A}' = \mathbf{A} + \nabla \Lambda(\mathbf{r}) \quad (74)$$

where  $\Lambda(\mathbf{r})$  is an arbitrary differentiable function the Dirac Hamiltonian transforms according to

$$\hat{h}_D \rightarrow \hat{h}'_D = \exp[-i\Lambda(\mathbf{r})] \hat{h}_D \exp[i\Lambda(\mathbf{r})]$$

which transforms the solution of the Dirac equation according to

$$\psi'(\mathbf{r}) = \exp[-i\Lambda(\mathbf{r})]\psi(\mathbf{r})$$

In order to illustrate the problems that may arise in the relativistic treatment of magnetic properties using finite basis-set methods, one may consider the simplest non-trivial example,  $\Lambda(\mathbf{r}) = z$ , which introduces an interaction Hamiltonian  $c\alpha_z$ . The untransformed and transformed Hamiltonians,  $\hat{h}_D$  and  $\hat{h}'_D$ , respectively, are given by

$$\hat{h}_D = c\boldsymbol{\alpha} \cdot \mathbf{p} + \beta mc^2 - \frac{Z}{r} \\ \hat{h}'_D = c\boldsymbol{\alpha} \cdot \mathbf{p} + \beta mc^2 - \frac{Z}{r} + c\alpha_z$$

and since they are related by an elementary gauge transformation, they generate identical eigenvalue spectra, with solution sets related by a complex phase factor.

Numerical results for this simple test case were presented by Quiney.<sup>(42)</sup> While it was confirmed that the results for  $\hat{h}'_D$  and  $\hat{h}_D$  converge as the basis set is enlarged, it was found necessary to saturate the space of  $G$ -spinors with  $\ell = 4$  functions in order to achieve the level of agreement between the two representations considered here. There are other less obvious problems associated with the use of the gauge function  $\Lambda(\mathbf{r})$ , the most important of which is that the transformed Dirac Hamiltonian can no longer be block-diagonalized in representations labelled by  $(\kappa, m_j)$ . For a finite representation, this splits the  $p_{3/2}$  states into two manifolds, characterized by  $m_j = \pm 1/2$  and  $m_j = \pm 3/2$ . The  $2p_{3/2, \pm 3/2}$  eigenvalue persists unchanged throughout all sets of calculations, because the only functions with  $m_j = \pm 3/2$  included in the basis are of  $p$ -type, and consequently there are no non-zero off-diagonal matrix elements of  $c\alpha_z$  involving these functions. Moreover, the basis-set representation of  $\hat{h}'_D$  requires the use of complex arithmetic, while the representation of  $\hat{h}_D$  may be constructed using real arithmetic, if the convention is adopted that the small-component amplitudes are chosen to be purely imaginary.

The reason for the complex nature of the representation of  $\hat{h}'_D$  is easy to find; the inclusion of basis functions of higher angular momentum serves to construct the phase factor  $\exp[-i\Lambda(\mathbf{r})]$  as a multipole expansion, which in this case takes the form

$$\exp[-i\Lambda(\mathbf{r})] = 1 - iz - \frac{z^2}{2!} + i\frac{z^2}{3!} + \dots$$

We may accommodate this gauge transformation exactly by defining  $B$ -spinors,  $B[T, \mu, \mathbf{r}_{A_\mu}]$ , the two-component equivalents of London orbitals,

$$B[T, \mu, \mathbf{r}_{A_\mu}] = \exp[-i\Lambda(\mathbf{r})]M[T, \mu, \mathbf{r}_{A_\mu}]$$

where  $M[T, \mu, \mathbf{r}_{A_\mu}]$  is a  $G$ -spinor. The phase factor  $\exp[-i\Lambda(\mathbf{r})]$  has the effect of including basis functions with large angular momentum into the basis set.

## REFERENCES

- Synek, M. (1964) *Phys. Rev. A*, **136**, 1552.
- Kim, Y.-K. (1967) *Phys. Rev.*, **159**, 190; errata.
- Kim, Y.-K. (1967) *Phys. Rev.*, **154**, 17.
- Lee, Y.S. and McLean, A.D. (1982) *J. Chem. Phys.*, **76**, 735.
- McLean, A.D. and Lee, Y.S. (1982) Dirac–Hartree–Fock molecular calculations relativistic effects on  $R_e$  and  $D_e$  in AgH and AuH, in *Current Aspects of Quantum Chemistry 1981*, ed. R. Carbo, Elsevier, Amsterdam, 219–238.
- Kutzelnigg, W. (1984) *Int. J. Quantum Chem.*, **25**, 107.
- Stanton, R.E. and Havriliak, S. (1984) *J. Chem. Phys.*, **81**, 1910.
- Grant, I.P. and Quiney, H.M. (2000a) *Phys. Rev. A*, **62**, 022508, 1–14.
- Stakgold, I. (1979) *Green's Functions and Boundary Value Problems*, Wiley & Sons, New York.
- Klahn, B. and Bingel, W.A. (1977) *Theor. Chim. Acta (Berl.)*, **44**, 9.
- Klahn, B. and Morgan, J.D. (1984) *J. Chem. Phys.*, **81**, 410.
- Kato, T. (1976) *Perturbation Theory for Linear Operators*, 2nd edition, Springer-Verlag, Berlin.
- Reed, M. and Simon, B. (1975) *Methods of Modern Mathematical Physics, II: Fourier Analysis and Self-Adjointness*, Academic Press, New York.
- Thaller, B. (1992) *The Dirac Equation*, Springer-Verlag, Berlin.
- Grant, I.P. (1996) in *Atomic, Molecular and Optical Physics Reference Book*, Chap. 22, ed. G.W.F. Drake, American Institute of Physics, New York.
- Greiner, W. (1985) *Relativistic Quantum Mechanics*, 2nd edition, Springer-Verlag, Berlin.
- Grant, I.P. (1989) in *Relativistic, Quantum Electrodynamics and Weak Interaction Effects in Atoms*, Vol. 189, eds P.J. Mohr, W.R. Johnson, and J. Sucher, American Institute of Physics, New York, 235–253.
- Dyall, K.G., Grant, I.P., and Wilson, S. (1984a) *J. Phys. B*, **17**, L45.
- Dyall, K.G., Grant, I.P., and Wilson, S. (1984c) *J. Phys. B*, **17**, 493.
- Grant, I.P. and Quiney, H.M. (1988) *Adv. Atom. Mol. Phys.*, **23**, 37.
- Dyall, K.G., Grant, I.P., and Wilson, S. (1984b) *J. Phys. B*, **17**, 1201.
- Rotenberg, M. (1962) *Ann. Phys. (N.Y.)*, **19**, 262.
- Rotenberg, M. (1970) *Adv. Atom. Mol. Phys.*, **6**, 233.
- Abramowitz, A. and Stegun, I.A. (1970) *Handbook of Mathematical Functions*, Dover Publications, New York.
- Grant, I.P. (1986) *J. Phys. B: At. Mol. Opt. Phys.*, **19**, 3187.
- Quiney, H.M. (2002) Integrals over spinors, in *Handbook of Molecular Physics and Quantum Chemistry*, Vol. 2, ed. S. Wilson, Wiley & Sons, Chichester.
- Drake, G.W.F. and Goldman, S.P. (1981) *Phys. Rev. A*, **23**, 2093.
- Drake, G.W.F. and Goldman, S.P. (1988) *Adv. Atom. Mol. Phys.*, **25**, 393.
- Goldman, S.P. and Drake, G.W.F. (1982) *Phys. Rev. A*, **25**, 2877.
- Goldman, S.P. (1985) *Phys. Rev. A*, **31**, 3541.
- Goldman, S.P. and Dalgarno, A. (1986) *Phys. Rev. A*, **57**, 408.
- Goldman, S.P. (1989) *Phys. Rev. A*, **39**, 976.
- Goldman, S.P. (2000) *Phys. Rev. A*, **61**, 052513.
- Johnson, W.R., Blundell, S.A., and Sapirstein, J. (1988) *Phys. Rev. A*, **37**, 307.
- Sapirstein, J. and Johnson, W.R. (1996) *J. Phys. B*, **29**, 5213.
- Salomonson, S. and Öster, P. (1989) *Phys. Rev. A*, **40**, 5548.
- Rosicky, P. and Karplus, M. (1977) *J. Chem. Phys.*, **67**, 5419.
- Quiney, H.M., Grant, I.P., and Wilson, S. (1985) *J. Phys. B*, **18**, 577.
- Quiney, H.M., Grant, I.P., and Wilson, S. (1985) *J. Phys. B*, **18**, 2805.
- Quiney, H.M., Hubac, I., and Wilson, S. (2001) *J. Phys. B: At. Mol. Opt. Phys.*, **34**, 4323.
- Szmytkowski, R. (1997) *J. Phys. B*, **30**, 825.
- Quiney, H.M. (2001) Relativistic quantum mechanics of atoms and molecules, in *New Trends in Quantum Systems in Chemistry and Physics. I. Basic Problems and Model Systems*, Vol. 6, eds J. Maruani, C. Minot, R. McWeeny, Y.G. Smeyers, and S. Wilson, Kluwer Academic Publishers, Dordrecht, Progr. Theor. Chem. and Phys., 135–173.