Minimization methods for the one-particle Dirac equation

Jean Dolbeault, Maria J. Esteban, Eric Séré and Michel Vanbreugel
CEREMADE (UMR C.N.R.S. 7534)
Université Paris IX-Dauphine
Place du Maréchal de Lattre de Tassigny
75775 Paris Cedex 16 - France
E-mail: dolbeaul, esteban or sere@ceremade.dauphine.fr, root@ceremade.dauphine.fr
(September 17, 2000)

Abstract. *Taking into account relativistic effects in quantum chemistry is crucial for accurate computations involving heavy atoms. Standard numerical methods can deal with the problem of* variational collapse *and the appearance of* spurious roots *only in special cases. The goal of this letter is to provide a general and robust method to compute particle bound states of the Dirac equation.*

PACS numbers : 31.30.Jv, 31.15.Pf, 03.65.Pm, 02.60.Cb

The free Dirac operator has a negative continuum in its spectrum. This is the cause for a variational instability which makes the numerical computations of one-particle bound states of Dirac equations difficult. Various approaches based for instance on squared Dirac operators [1,2], min-max formulations [3,4], special basis [5,6] or even more elaborated methods have been proposed [7,8], as well as perturbative expansions of non-relativistic models and derivation of effective Hamiltonians. None of these remedies provides a complete and satisfactory answer. From a numerical viewpoint, the *variational collapse* and the existence of *spurious states* [9,10] are serious problems which have been solved in special cases by taking appropriate projections or imposing additional conditions, for instance on the boundary [5].

Here we propose exact and stable variational methods. First we present an abstract min-max approach and show how to reduce it to a constrained minimization problem. The method requires no specific precaution and allows for instance the presence of several nuclei. It is then illustrated by three computations of the *ground state* for a particular class of spherical potentials: a **shooting method** (specific to the case of spherical potentials), which is compared with a **minimization method** under constraint (see Table I), and a **direct minimization method** based on a decomposition on a general basis (see Table II) – which is equivalent to the minimization method – but has a form which is closer to the min-max formulation and is easier to generalize to non central potentials.

To compute the bound states of the one-particle Dirac Hamiltonian

$$H = \alpha\cdot(-i\nabla) + mc^2\beta + V\ , \tag{1}$$

where α and β are the usual Dirac-Pauli matrices, it is natural to characterize the energy levels of H by min-max methods applied to the Rayleigh quotients $(\varphi, H\varphi)/(\varphi,\varphi)$. Indeed, the spectrum of the Dirac Hamiltonian H being unbounded from below, it is hopeless to just minimize the Rayleigh quotient without any further precaution, since this would take us to $-\infty$. We refer to [11–15] for a mathematical study of min-max formulations. In [15] it was proved for a large class of potentials that if $F_+\oplus F_-$ is an orthogonal decomposition of a well-chosen space of smooth square integrable functions, the sequence of min-max levels

$$\lambda_k = \inf_{\substack{G \text{ subspace of } F_+ \\ \dim\, G=k}} \ \sup_{\substack{\varphi\in(G\oplus F_-)\\ \varphi\neq 0}} \frac{(\varphi, H\varphi)}{||\varphi||^2} \tag{2}$$

is equal to the sequence of the energy levels of H (counted with multiplicity) in the interval $(-mc^2, +mc^2)$.

This expression has however not much pratical interest and we will now reformulate it as a minimization problem (see [13] for mathematical justifications). Composing H with its associated positive energy projector Λ^+ would allow to minimize Rayleigh quotients but this idea is formal since the above projector is *a priori* unknown. Instead we will introduce a (nonlinear) constraint which implicitely defines the projector. The Dirac equation $H\psi = \lambda\psi$, originally written for 4-spinors, is first reduced to an equivalent equation for 2-spinors: the lower 2-spinor, or *small component* is written in terms of the *large component* (see [16,6]). More precisely, for any ψ with values in $\mathbb{C}^4$, let us write $\psi = \binom{\varphi}{\chi}$, with φ, χ taking values in $\mathbb{C}^2$. Then the equation $H\psi = \lambda\psi$ is equivalent to the system

$$\left\{\begin{array}{rcl} L\chi &=& (\lambda - mc^2 - V)\ \varphi \\ L\varphi &=& (\lambda + mc^2 - V)\ \chi \end{array}\right. \tag{3}$$

with $L = ic(\vec\sigma.\vec\nabla) = \sum_{k=1}^3 ic\sigma_k\partial/\partial x_k$, σ_k being the Pauli matrices ($k = 1,2,3$). As long as $\lambda + mc^2 - V \neq 0$, the system (3) can be written as

$$L\Big(\frac{L\varphi}{g_{_E}}\Big) + V\varphi \ = \ E\varphi\ ,\quad \chi = \frac{L\varphi}{g_{_E}} \tag{4}$$

where $g_{_E} = E + 2mc^2 - V$, $E = \lambda - mc^2$.

Then, system (3) written for $\phi = \varphi/\sqrt{g_{_E}}$ reads

$$H_E\phi = (E-V)(E+2mc^2-V)\phi \tag{5}$$

where the operator $H_E\phi := \sqrt{g_E}\ L(\frac{1}{g_E}L(\sqrt{g_E}\ \phi))$ is symmetric (and self-adjoint for an appropriate domain). Thus E is a solution of $(\phi,\phi)\,E^2 + 2(\phi,(mc^2-V)\phi)E - (\phi,(2mc^2-V)V\phi) - (\phi,H_E\phi) = 0$,

$$E = J^{\pm}(E,\phi) := \frac{1}{(\phi,\phi)}\left(\pm\sqrt{\Delta(E,\phi)} - (\phi,(mc^2-V)\phi)\right) \qquad (6)$$

where $\Delta(E,\phi) := |(\phi,V\phi)|^2 + (\phi,\phi)[m^2c^4(\phi,\phi) + (\phi,H_E\phi) - (\phi,V^2\phi)]$. Heuristically (6) is one of the analogues in quantum mechanics of Einstein's relation: $E = \pm(p^2c^2 + m^2c^4)^{1/2} - mc^2 + V$.

It turns out that the critical points of $J^{\pm}(E,\phi)$ under the constraint $E = J^{\pm}(E,\phi)$ are exactly (see [13]) the bound states of the Dirac operator H corresponding to energies in the gap (and those corresponding to J^+ converge as $c \to +\infty$ to the nonrelativistic energy levels of the Schrödinger operator). Note that if V is nonpositive, the range of J^- is contained in $(-\infty, -mc^2]$ which corresponds to the negative part of the continuum of the spectrum of the Dirac operator. We may in that case define the *ground-state* as the lowest energy level in the gap, which turns out to be the minimum of $J^+(E,\phi)$ under the constraint $E = J^+(E,\phi)$.

The functional J^+ is well defined only if further conditions are assumed on the potential. If for instance $V = \frac{\zeta}{|x|}$ is the Coulomb potential, $\zeta = \frac{Ze^2}{\hbar c}$ has to be less than 1 (which means $Z < 1/\alpha = 137.036...$), otherwise H is not well defined as a self-adjoint operator: see [17,13,15] for more details. The fact that the whole spectrum in the gap is actually characterized by $J^{\pm}$ comes from the equivalence of the method with the min-max characterization of the energy levels in the gap [15].

If V is spherically symmetric (see for instance [17,18]) the bound states can be expressed in terms of the spherical harmonics using partial wave Dirac operators acting on the space $(L^2(0,+\infty))^2$ of the square integrable real functions on $(0,+\infty)$, which have the form

$$h = \begin{pmatrix} mc^2 + V & -c\frac{d}{dr} + \frac{c\kappa}{r} \\ c\frac{d}{dr} + \frac{c\kappa}{r} & -mc^2 + V \end{pmatrix} \qquad (\kappa = \pm 1,\ \pm 2, ...) \qquad (7)$$

From now on we choose a system of units in which $m = 1$ and $c = 1$. Problem (3) takes the form

$$\begin{cases} u' = (1+\lambda)v - (Vv + \frac{\kappa}{r}u) \\ v' = (1-\lambda)u + Vu + \frac{\kappa}{r}v) \end{cases} \qquad (8)$$

The solutions of this system are characterized by two parameters, λ and $\delta = v(1)/u(1)$ for instance, and we shall denote by X the set of the solutions of (8) such that $u(1) = 1$ when λ and δ vary. However, the condition that u and v are in $L^2(0,+\infty)$ determines λ and δ. One can show that this condition is equivalent to assuming that

$$\begin{aligned} &\lim_{r\to 0_+} r(|u(r)|^2 + |v(r)|^2) = 0\ , \\ &\lim_{r\to +\infty}(|u(r)|^2 + |v(r)|^2) = 0\ , \end{aligned} \qquad (9)$$

thus providing us with a first numerical **shooting method** to determine λ and δ (this is of course valid only for spherically symmetric potentials). We shall refer to this method by the letter **s** and use it as a comparison test for the numerical results obtained by the minimization approach given below. The approximated energy levels computed with this method will therefore be denoted by λ^s in Table I.

Let us describe now the **minimization method** based on J^+. Similarly to (4), v can be eliminated in terms of u:

$$\frac{v}{r^\kappa} = (r^{2\kappa}(1+\lambda-V))^{-1}\frac{d}{dr}(r^\kappa u)\ . \qquad (10)$$

We are looking for the ground-state so we may choose $\kappa = -1$ and problem (8) is now equivalent to solving

$$h_\lambda\phi = (1+\lambda-V)(1-\lambda+V)\phi \qquad (11)$$

where h_λ is formally a self-adjoint operator:

$$h_\lambda\phi = \sqrt{1+\lambda-V}\,\frac{d}{dr}\left[\frac{r^2}{1+\lambda-V}\frac{d}{dr}(\sqrt{1+\lambda-V}\,\phi)\right] \qquad (12)$$

and $\phi(r) = r^{-1}u(r)/\sqrt{1+\lambda-V}$ is now a function defined on $(0,+\infty)$. Equation (5) is then equivalent to

$$(\phi,\phi)\lambda^2 - 2(\phi,V\phi)\lambda + (\phi,V^2\phi) - [\|\phi\|^2 + (\phi,h_\lambda\phi)] = 0 \qquad (13)$$

where $(.,.)$ is the usual scalar product in $L^2(0,+\infty)$ and $\|.\|$ the corresponding norm. The problem is then reduced to finding a critical point of $J^+(\lambda-1,.)$ with $\lambda - 1 = J^+(\lambda-1,\phi)$ and

$$J^+(\lambda-1,\phi) + 1 = \frac{\sqrt{\Delta(\lambda-1,\phi)} + (\phi,V\phi)}{\|\phi\|^2}\ , \qquad (14)$$

$$\Delta(\lambda-1,\phi) = (\phi,V\phi)^2 + \|\phi\|^2[(\phi,h_\lambda\phi) + \|\phi\|^2 - (\phi,V^2\phi)]\ . \qquad (15)$$

To solve this constrained problem numerically, the natural idea is to introduce a penalization method and to minimize $J^+(\lambda-1,\phi) + A|(\lambda-1) - J^+(\lambda-1,\phi)|^2$ for A large enough. Actually if we assume that ϕ is given by (12) with (u,v) in X, the condition that u and v are in $L^2(0,+\infty)$ is equivalent to assuming that the integrals involved in (14) are finite. Of course these integrals are numerically computed on an interval (ϵ,R) and the approximate value $J^+_{\epsilon,R}$ of J^+ is finite even if the constraint is not satisfied, but we observe that $\lim_{(\epsilon,R)\to(0,+\infty)} J^+_{\epsilon,R}(\lambda-1,\phi) = +\infty$ unless $\lambda - 1 = J^+(\lambda-1,\phi)$. A minimization of J^+ (numerically, of $J^+_{\epsilon,R}$) on the set X without constraint is therefore equivalent to a constrained minimization of J^+. This method will be referred by the letter **m** in Table I, which contains the results of our computations using both the shooting and the minimization methods described above. The set of functions over which the approximated energy levels are computed consists of all the solutions of (8), with δ and λ to be determined. Once this is done, as a byproduct of the above minimization method, the wave functions associated with the computed energy levels are given by (8). This is also true for the shooting method.

The main advantage of the minimizing setting described above is that it can be extended to nonsymmetric situations (non central potentials), but of course for a minimizing set which is larger than X. Indeed, above we have minimized J^+ only among the set of functions which are already solutions of a radially symmetric system (8). In the rest of this letter, for convenience, we still assume that the potential is radial, but consider a general basis of $L^2(0,+\infty)$ and introduce a third formulation, which is intermediate between the abstract minmax theory and the minimization of J^+ described above. Its main advantage is that the constraint $E = J^+(E,\phi)$ will be automatically satisfied. We will therefore call this method the **direct minimization method**.

As in Equation (4) we may rewrite (3) as

$$\chi = (\lambda + mc^2 - V)^{-1} L\varphi \ , \tag{16}$$

$$L\left(\frac{L\varphi}{\lambda + mc^2 - V}\right) = (\lambda - mc^2 - V)\varphi \ , \tag{17}$$

at least for any $\lambda \in]-mc^2, +mc^2[$ if V takes negative values. Multiplying Equation (17) by φ and integrating with respect to $x \in \mathbb{R}^3$, we get :

$$\begin{aligned} f_\varphi(\lambda) &:= \int \frac{|L\varphi|^2}{\lambda + mc^2 - V}\, dx \\ &= (\lambda - mc^2)\|\varphi\|^2_{L^2} - \int V|\varphi|^2\, dx =: g_\varphi(\lambda) \ . \end{aligned} \tag{18}$$

Since for a given φ, $f_\varphi(\lambda)$ is decreasing and $g_\varphi(\lambda)$ is increasing, if there exits a $\lambda = \lambda[\varphi]$ such that (18) is satisfied, then it is unique (the existence of such a λ for all u depends on the properties of the potential V). According to [15], for those V's, the ground state is such that

$$\lambda_1 = \min_{\varphi} \lambda[\varphi] \ . \tag{19}$$

For a radial potential we may use the radial Dirac equation and consider (8) instead of (3). For $m = 1$ and $c = 1$, $\lambda = \lambda_r[u]$ is then the unique solution of

$$\begin{aligned} f(\lambda) &= \int_0^{+\infty} \frac{|(r^\kappa u)'|^2}{r^{2\kappa}(1 + \lambda - V(r))}\, dr \\ &= (\lambda - 1)\int_0^{+\infty} |u(r)|^2\, dr - \int_0^{+\infty} V(r)|u(r)|^2\, dr \ . \end{aligned} \tag{20}$$

To solve it numerically, it is more convenient to rewrite $f(\lambda)$ as

$$f(\lambda) = \sum_{k=0}^{+\infty} \left[(-1)^k \int_0^{+\infty} \frac{r^{-2\kappa}|(r^\kappa u)'|^2}{(1 - V(r))^{k+1}}\, dr\right] \lambda^k \ . \tag{21}$$

The solution (with $\kappa = -1$) is then approximated on a finite basis $(u_i)_{i=1,2,\dots n}$: $u = \sum_{i=1}^n x_i u_i$. If

$$f_{ijk} = (-1)^{k-1} \int_0^{+\infty} \frac{r^2 (u_i/r)'(u_j/r)'}{(1 - V(r))^k}\, dr \tag{22}$$

and

$$V_{ij} = \int_0^{+\infty} u_i(r) u_j(r) V(r)\, dr \ , \tag{23}$$

TABLE I. Comparison of the shooting (s) and the minimization (m) methods for $\kappa = -1$, $m = c = 1$, $V(r) = -\zeta r^{-\beta}$, $\zeta = 0.5$ and $\beta \in (0,1)$. The system (8) is numerically solved with a stepsize adaptative Runge-Kutta method on the interval $(\epsilon = 10^{-4}, R = 15)$. For the shooting method we minimize the quantity $\epsilon(|u(\epsilon)|^2 + |v(\epsilon)|^2) + \theta(|u(R)|^2 + |v(R)|^2) = \Delta_{\mathrm{s}}$ for some scale parameter $\theta > 0$ (which is chosen to balance both boundary terms), while for the minimization method, the quantity $J^+(\lambda - 1, \phi)$ is directly minimized, the quantity $|J^+(\lambda - 1, \phi) - (\lambda - 1)|^2$ being computed *a posteriori* at $\lambda = \lambda^m$. For $\beta = 1$, the result is known explicitely: $\lambda_1 = [1 - \zeta^2]^{1/2} = 0.866025...$, $\delta_1 = -[(1-\lambda)/(1+\lambda)]^{1/2} = -0.267949...$

β	δ^{s}	δ^{m}	λ^{s}	λ^{m}	J^+	$\lvert(\lambda^m - 1) - J^+(\lambda^m - 1, \phi)\rvert^2$	Δ^{s}
1	-0.267954	-0.267943	0.866034	0.866013	0.866014	$1.8\ 10^{-12}$	0.00029
0.9	-0.235187	-0.235174	0.856725	0.856698	0.856698	$2.1\ 10^{-14}$	0.00053
0.8	-0.207802	-0.207788	0.843181	0.843146	0.843146	$5.2\ 10^{-14}$	0.00063
0.7	-0.183397	-0.183379	0.825877	0.825832	0.825831	$4.3\ 10^{-13}$	0.00076
0.6	-0.160651	-0.160627	0.804699	0.804639	0.804639	$4.1\ 10^{-13}$	0.00094
0.5	-0.138654	-0.138619	0.779161	0.779071	0.779070	$3.4\ 10^{-13}$	0.0012
0.4	-0.116645	-0.116584	0.748381	0.748221	0.748220	$3.8\ 10^{-13}$	0.0018
0.3	-0.0938375	-0.0937016	0.710904	0.710537	0.710536	$3.5\ 10^{-13}$	0.0049
0.2	-0.069224	-0.068798	0.664252	0.663067	0.663067	$2.4\ 10^{-13}$	0.0097
0.1	-0.0412322	-0.0392963	0.60391	0.59833	0.59833	$1.4\ 10^{-13}$	0.018

the approximating equation for λ corresponding to (18) is

$$\sum_{i,j=1}^{n}\left((\sum_{k=1}^{m} f_{ijk}\lambda^{k-1})+V_{ij}\right)x_i x_j+(1-\lambda)\sum_{i=1}^{n}x_i^2=0\ ,$$

where the series in λ has been truncated at order m. It is actually more convenient to define

$$A^{n,m}(\lambda)=\left((\sum_{k=1}^{m} f_{ijk}\lambda^{k-1})+(1-\lambda)\delta_{ij}+V_{ij}\right)_{i,j=1,2,\dots n}$$

and to approximate λ_1 by $\lambda_1^{n,m}$ defined as the first positive root of $\lambda\mapsto\mu(A^{n,m}(\lambda))$ where $\mu(A)$ denotes the first eigenvalue of the matrix A (see [15] for more details). Note that $(\lambda_1^{n,m})_{m\geq 1}$ is an alternating sequence (which essentially converges at a geometric rate): two consecutive eigenvalues determine an interval containing $\lim_{m\to+\infty}\lambda_1^{n,m}$. Numerical results for a special basis are given in Table II, with two types of approximations: a finite basis with n elements is used and the series in powers of λ is truncated at a finite order m.

Regarding excited states, they can be obtained by the first and second methods (shooting and minimization), with κ fixed to values different from -1 and for appropriate values for the quantum numbers corresponding to the decomposition into spherical harmonics. In the case of the direct minimization method, the computation of excited states is more straightforward. It is indeed reduced to the computation of the ith root of $\lambda\mapsto\mu_i(A^{n,m}(\lambda))$, where $\mu_i(A)$ denotes the ith eigenvalue of A.

In this letter we have proved that computations of one-particle bound states based on a variational formulation of the Dirac equation are not subjected to the numerical difficulties due to the negative continuum of the spectrum, thus showing the possibility of a mathematical foundation for such numerical methods. These computations are general, robust and not restricted to central potentials. Optimal numerical accuracy has not been our primary concern, since we were rather interested in showing the feasability of an approach based on the variational structure of the equation. Hopefully, this new approach will contribute to the elaboration of more efficient and consistent numerical methods.

[1] H. Wallmeier, W. Kutzelnigg, Chem. Phys. Lett. **78**, 341 (1981).
[2] W.E. Baylis, S.J. Peel, Phys. Rev. A **28**, 2552 (1983).
[3] J.D. Talman, Phys. Rev. Lett. **57**, 1091 (1986).
[4] S.N. Datta, G. Deviah, Pramana, **30**, 387 (1988).
[5] G.W.F. Drake and S.P. Goldman, Phys. Rev. A **23**, 2093 (1981).
[6] W. Kutzelnigg, Chemical Physics **225**, 203 (1997).
[7] R.N. Hill, C. Krauthauser, Phys. Rev. Lett. **72**, 2151 (1994).
[8] G.W.F. Drake and S.P. Goldman, Adv. Atomic Molecular Phys. **25**, 393 (1988).
[9] W. Kutzelnigg, International J. Quantum Chem. **XXV**, 107 (1984).
[10] I.P. Grant, A.I.P. Conf. Proc. **189**, 235 (1989).
[11] M.J. Esteban, E. Séré, *Partial Differential Equations and Their Applications,* CRM Proceedings and Lecture Notes, **12**, Eds. P.C. Greiner, V. Ivrii, L.A. Seco and C. Sulem. AMS (1997).
[12] M. Griesemer, H. Siedentop, Preprint mp-arc 97-492, to appear in J. London Math. Soc.
[13] J. Dolbeault, M.J. Esteban, E. Séré, Calc. Var. **10**, 321 (2000).
[14] M. Griesemer, R.T. Lewis, H. Siedentop, Preprint 1999.
[15] J. Dolbeault, M.J. Esteban, E. Séré, J. Funct. Anal. **174**, 208 (2000).
[16] R. van Leeuwen, E. van Lenthe, E.J. Baerends, J.G. Snijders, J. Chem. Phys. **101**, 1272 (1994).
[17] B. Thaller. *The Dirac Equation.* Texts and Monographs in Physics. Springer-Verlag, Berlin, 1st edition (1992).
[18] J.D. Bjorken, S.D. Drell, Relativistic Quantum Mechanics, McGraw-Hill (1964).

TABLE II. Direct minimisation method for $\kappa=-1$, $m=c=1$, $V(r)=-\zeta r^{-\beta}$, $\zeta=0.5$ and β close to 1. The approximating space is of dimension $n=10$ (we consider the orthonormal basis generated by the ground state of the hydrogen atom and $n-1$ Hermite functions, which is probably not very well adapted) and the series are truncated at $m=14$ or $m=15$ (the corresponding values $\lambda_1^{10,14}$ and $\lambda_1^{10,15}$ are respectively a lower and an upper bound of $\lim_{m\to+\infty}\lambda_1^{10,m}$). This last approximation is certainly rather crude as shown by the case of the Coulomb potential. As in Table I, J^+ is obtained through a minimization on the set X, and $\Delta^{\rm m}:=(1-\sum_{i=1}^{10}(u^m,u_i)^2)^{1/2}$ measures the error (in the L^2-norm) when the corresponding solution is approximated on the finite basis.

β	0.90	0.93	0.95	0.97	0.99	1.00
$\lambda_1^{10,14}$	0.855681	0.858516	0.860228	0.861792	0.863200	0.863843
$\lambda_1^{10,15}$	0.858012	0.861112	0.863004	0.864749	0.866338	0.867071
J^+	0.856698	0.859984	0.861954	0.863735	0.865310	0.866014
$\Delta^{\rm m}$	0.0082	0.0058	0.0046	0.0033	0.0020	0.0022