Maria J. Esteban

Mathematical questions about the computation of eigenvalues of Dirac operators with critical potentials in atomic and molecular physics

<https://doi.org/10.5802/crphys.16>

Part of the Thematic Issue: Prizes of the French Academy of Sciences 2019

Some rights reserved.

This article is licensed under the Creative Commons Attribution 4.0 International License.
http://creativecommons.org/licenses/by/4.0/

Les Comptes Rendus. Physique sont membres du Centre Mersenne pour l’édition scientifique ouverte
www.centre-mersenne.org
Mathematical questions about the computation of eigenvalues of Dirac operators with critical potentials in atomic and molecular physics

Abstract. This Note describes various analytical and computational results concerning the calculation of Dirac eigenvalues, or more generally, of operators with gaps. An algorithm based on an abstract theorem characterizing the eigenvalues in gaps was found years ago, but it is only recently that a delicate analysis to identify and study the domains of those operators has allowed to put that algorithm on a firm basis concerning the choice of approximation basis sets, and this both for light and for heavy atoms. The works described here concern joint papers with several collaborators: J. Dolbeault, M. Lewin, M. Loss, E. Séré and M. Vanbreugel.

Résumé. Cette Note présente divers résultats analytiques et numériques concernant le calcul des valeurs propres des opérateurs de Dirac avec des potentiels critiques en physique atomique et moléculaire.

Keywords. Dirac operator, Eigenvalue, Operator domain, Eigenvalue computation, Self-adjoint operator.

2020 Mathematics Subject Classification. Primary 49S05; Secondary 35Q75, 47-08, 65D07, 70-08, 81-08, 81Q05, 81Q10, 81V45, 81V55.

Après avoir analysé le problème de manière abstraite, nous avons obtenu des applications particulièrement intéressantes pour l’opérateur de Dirac perturbé par un potentiel extérieur électromagnétique. Utilisant une décomposition de l’espace des spineurs proposée par Talman [16], l’application du théorème abstrait nous a permis de trouver un algorithme qui est très performant pour le calcul des valeurs propres d’opérateurs de Dirac, et ceci sans avoir à faire face à l’apparition de valeurs spurieuses. De nombreux résultats dans la littérature proposent des choix de bases finies particulières pour éviter la pollution spectrale. Dans un des derniers travaux présentés dans cette Note, nous montrons que dans notre approche il y a une énorme liberté dans le choix des bases finies, et que dans tous les cas, il n’y aura pas de valeur spurieuse, et que l’algorithme convergera vers les valeurs propres. Des résultats récents s’intéressant à l’étude détaillée des domaines des opérateurs nous ont donc permis de donner une base solide à cet algorithme pas seulement pour les calculs impliquant des atomes légers, mais aussi, et surtout, dans le cas des atomes lourds, qui sont les plus intéressants dans une théorie relativiste.

1. Introduction

In 1928 [17] Paul Dirac derived an operator for quantum electrodynamics, starting from the usual classical expression of the energy of a free relativistic particle of momentum \(p \in \mathbb{R}^3\) and mass \(m\),

\[
E^2 = c^2 |p|^2 + m^2 c^4.
\]  

(1)

His aim was to propose a local differential operator of first order with respect to \(p = -i\hbar \nabla\):

\[
D_{m,c,h} = -ich \alpha \cdot \nabla + mc^2 \beta = -ich \sum_{k=1}^{3} \alpha_k \partial_k + mc^2 \beta,
\]  

(2)

where \(\alpha_1, \alpha_2, \alpha_3\) and \(\beta\) are Hermitian matrices which have to satisfy the following anticommutation relations:

\[
\begin{align*}
\alpha_k \alpha_\ell + \alpha_\ell \alpha_k &= 2 \delta_{k\ell} I_4, \\
\alpha_k \beta + \beta \alpha_k &= 0, \\
\beta^2 &= I_4.
\end{align*}
\]  

(3)

It can be proved [18] that the smallest dimension in which (3) can take place is 4 (i.e. \(\alpha_1, \alpha_2, \alpha_3\) and \(\beta\) should be \(4 \times 4\) Hermitian matrices), meaning that \(D_c\) has to act on \(L^2(\mathbb{R}^3, \mathbb{C}^4)\). The usual representation in \(2 \times 2\) blocks is given by

\[
\beta = \begin{pmatrix} I_2 & 0 \\ 0 & -I_2 \end{pmatrix}, \quad \alpha_k = \begin{pmatrix} 0 & \sigma_k \\ \sigma_k & 0 \end{pmatrix} \quad (k = 1, 2, 3),
\]  

where the Pauli matrices are defined as

\[
\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
\]
In which follows we will always work in the framework of atomic units, where we can assume that the constants $\hbar, c$ and $m$ are equal to 1. For this particular choice of units, the unperturbed Dirac operator will be denoted by $H_0$:

$$H_0 := -i\alpha \cdot \nabla + \beta.$$ 

It is well known that the Dirac operator is essentially self-adjoint and that its spectrum is all continuous spectrum and equal to

$$(-\infty, -1] \cup [1, +\infty).$$

Of course the main interesting situation for atomic or molecular physics is to consider the Dirac operator perturbed by an external electromagnetic potential, corresponding to the interaction of the electron(s) with other charged particles and its evolution under the action of an external magnetic field. Without magnetic field, one is thus interested in looking at operators of the form $H_0 + V$, and with magnetic field, one has to replace the usual gradient operator $\nabla$ by the magnetic gradient $\nabla_A = \nabla + iA$, where $A$ is a potential related to the magnetic field $B$ by $B = \text{curl } A$. In this case, the magnetic Dirac operator is denoted by $H_A$:

$$H_A := -i\alpha \cdot \nabla_A + \beta.$$ 

If now we consider the perturbed operator $H_A + V$ and the potential $V$ is for instance Coulomb-like, eigenvalues of $H_A + V$ appear in the gap of the essential spectrum and those eigenvalues correspond to discrete electronic states in the atom or the molecule. Computing eigenvalues of operators in gaps is notoriously difficult. For instance, the ground state, or better said, the minimal electronic eigenvalue, cannot be found by a simple minimization procedure.

Another serious issue is that depending on the choice of the approximating basis sets and on the computing algorithm, spurious eigenvalues not converging towards exact eigenvalues can appear. For examples and discussions on this issue see for instance [1–7] in the Physics and Chemistry literature. Interesting analysis and propositions to deal with this issue can be found in [8–15].

In [19–22], under adequate assumptions, variational min–max formulas were provided for the eigenvalues in gaps of self-adjoint operators $A$. These formulas are based on a decomposition $\mathcal{H} = \mathcal{H}^+ \oplus \mathcal{H}^-$ given by two orthogonal projectors $\Lambda^\pm$ of the ambient Hilbert space $\mathcal{H}$, and take the general form

$$
\lambda^{(k)} = \inf_{\dim(W) = k} \sup_{W \subset \mathcal{H}^+, \psi \in W \oplus \mathcal{H}^-} \frac{\langle \psi, A\psi \rangle}{\|\psi\|^2}.
$$

Here, $\mathcal{H}^\pm = \Lambda^\pm \mathcal{H}$, with $F$ a dense subspace of the domain of $\mathcal{H}$, such that the quadratic form $\langle \psi, A\psi \rangle$ is well-defined on $\mathcal{H}^+ \oplus \mathcal{H}^-$. See also the recent articles [23–25]. Based on a simple and very useful orthogonal decomposition proposed by Talman [16], it was proved by Dolbeault, Esteban and Séré in [22] for the case without magnetic field, and later in [26, 27] by Dolbeault, Esteban and Loss for the case of an external constant magnetic field, that the above abstract result implies that for electrostatic potentials having at worst singularities of the Coulomb type, $-\nu/|x|$, with $0 < \nu \leq 1$, the eigenvalues of the operator $H_A + V$ can be found by the following simple and computable procedure: for functions $\varphi \in L^2(\mathbb{R}^3, \mathbb{C}^2)$, consider the quadratic form

$$Q_{A,V,\lambda}(\varphi) := \int_{\mathbb{R}^3} \left( \frac{|\varphi \cdot \nabla_A \varphi|^2}{1 + \lambda - V} + (1 - \lambda + V)|\varphi|^2 \right) dx$$

which is decreasing in $\lambda$. If the electrostatic potential is not very large, and in particular, if the singularities of $V$ are not worse than $-\nu/|x|$, with $0 < \nu \leq 1$, condition which is necessary to prove that the operator $H_A + V$ can be defined as a self-adjoint operator in a physically meaningful manner (see Esteban–Loss [26, 28]), then, the quadratic form $Q_{A,V,\lambda}$ is positive for $\lambda$ in some interval $[-\infty, b)$, $b \in (-1, 1)$. Moreover the first (smallest) eigenvalue of $H_A + V$ will be the smallest
λ for which there is a function ϕ satisfying QA,ν,λ(ϕ) = 0. More concretely, and equivalently, let T(λ) be the operator defined via the quadratic form which acts on 2-spinors:

$$(ϕ, T(λ)ϕ) := \int_{\mathbb{R}^3} \left( \frac{|σ \cdot ∇Aϕ|^2}{λ + 1 - V} + (1 - λ + V)|ϕ|^2 \right) dx$$

and consider its first eigenvalue, μ1(λ). Since T(λ) is monotone decreasing with respect to λ, there exists a unique λ1 such that μ1(λ1) = 0. Then λ1 is the smallest eigenvalue of HA + V in the gap (−1, 1). Furthermore, for every positive integer k, if λk is the unique root of the equation μk(λ) = 0, then λk is the k-th eigenvalue of HA + V in the gap (−1, 1), counted with multiplicity. All these results were proved in [22] in all cases where the operator HA + V is self-adjoint and when A corresponds to a constant magnetic field and when V is not too large (the concrete conditions are stated in the theorems contained in [22]).

It is quite simple to propose now an algorithm to compute the eigenvalues of HA + V. For that purpose let us choose an n-dimensional space of functions from $\mathbb{R}^3$ to $\mathbb{C}^2$ and generated by \{ϕ1, ϕ2, ..., ϕn\}. Let $T_n(λ)$ be the $n \times n$ matrix whose elements are given by

$$T_n^{i,j}(λ) = \int_{\mathbb{R}^3} \left( \frac{(σ \cdot ∇Aϕ_i, σ \cdot ∇Aϕ_j)}{λ + 1 - V} + (1 - λ + V)(ϕ_i, ϕ_j) \right) dx.$$

Let μn(λ) the smallest eigenvalue of $A_n(λ)$. Then, the unique zero of the map $λ \mapsto μ_n(λ)$, λn, is an approximation of the first eigenvalue of HA + V in the gap (−1, 1) if the set {ϕ1, ϕ2, ..., ϕn, ...} generates a space F like the one present in the above abstract theorem about eigenvalues in gaps. Or more generally, if the set {ϕ1, ϕ2, ..., ϕn, ...} generates the domain of the operator HA + V. Using this algorithm, in [29], in collaboration with Dolbeault, Séré and Vanreugel, we computed the ground state and the ground state energy for an electron in the electrostatic field created by light and heavy nuclei (H, He+, Ge23+ and Th89+). The algorithm converged beautifully, without the presence of spurious eigenvalues that are often present in Dirac eigenvalues computations. In this case we chose the functions ϕn as Hermite polynomials. In [30], with Dolbeault and Séré we performed the same kind of computations for a diatomic configuration, with both light and heavy atoms in two separate locations (H2+ and Th279+). These computations were done choosing B-spline functions for the functions ϕn and in cylindrical coordinates. Again the computations ran perfectly, and the values obtained in these computations fitted perfectly experimental data and data obtained by using other algorithms. In [27], with Dolbeault and Loss we also made computations for the magnetic case, again for light and heavy nuclei, and once again the results fitted extremely well existing results obtained by other means. In particular we were able to produce new eigenvalue approximations for heavy atoms in cases that had not been dealt with before.

But, even if the above computations were excellent and the algorithm was robust and very efficient, there was a deep problem behind them. Indeed, for electrostatic potentials involving Coulomb-like singularities −ν/|x|, with 0 < ν < $\sqrt{3}/2$, the domain of the Dirac–Coulomb operator $H_0 − ν/|x|$ is equal to the Sobolev space $H^{1/2}(\mathbb{R}^3, \mathbb{C}^4)$ defined as the 4-spinors which are square integrable and such that all their first derivatives are also square integrable. But when $ν > \sqrt{3}/2$ it is known that the domain is contained in $H^{1/2}(\mathbb{R}^3, \mathbb{C}^4)$ and contains $H^1(\mathbb{R}^3, \mathbb{C}^4)$, but it is not equal to any of these two spaces. In the concrete case of Coulomb potentials, the domain could be computed explicitly, but not for other potentials having the same singularities, but not being exactly Coulomb-like. It was therefore delicate to use the above abstract theorem to derive the algorithm and use it with basis sets that maybe were not generating the domain of the operator. This is why recently in [31–33]), with Lewin and Séré we have addressed this issue, describing in full detail the domains of the Dirac–Coulomb-like operators for large $ν$, that is, for $\sqrt{3}/2 ≤ ν ≤ 1$ and trying to see which properties are necessary for the basis sets for the above algorithm to converge to the eigenvalues, and not to some numbers above them.
In [31] we described the domains of Dirac–Coulomb-like operators for $0 < \nu \leq 1$. In various papers written in the 70’s and 80’s we can find proposals of physically meaningful definitions of Dirac–Coulomb operators as self-adjoint operators for $0 < \nu < 1$ and in all cases the domain was shown to be a strict subspace of $H^{1/2}(\mathbb{R}^3, \mathbb{C}^4)$. In the end, all those proposals were shown to be equivalent, of course. For full details on all those developments see [18, 34–42]. The limit case $\nu = 1$ is harder, and this was dealt with by Esteban and Loss in [28] by using a novel method to prove self-adjointness for operators with gaps. This was later extended by Arrizabalaga et al. in [43, 44]. In the case $\nu = 1$ the domain is not a subspace of $H^{1/2}(\mathbb{R}^3, \mathbb{C}^4)$ anymore. This full description of the domain of Dirac–Coulomb-like operators was the first step towards the full understanding of how to interpret the results of the numerical computations done by the above algorithm. In [31] we found an additional result proving that the space $C^\infty_c(\mathbb{R}^3 \setminus\{0\}, \mathbb{C}^2)$ (compactly supported functions which are infinitely derivable in all points except at the origin) is dense in the space of the upper components (2-spinors) of the elements of the domain. This density argument is key in the proof of our main result which states that in the min–max characterization of the eigenvalues (4) of the operator $H_0 - \nu/|x|$, $0 < \nu \leq 1$, the space $F$ can be any among those satisfying

\[
C^\infty_c(\mathbb{R}^3 \setminus\{0\}, \mathbb{C}^4) \subset F \subset H^{1/2}(\mathbb{R}^3, \mathbb{C}^4)
\]

The immediate consequence of this result is that when running the above algorithm we can take basis sets that span any space between $C^\infty_c(\mathbb{R}^3 \setminus\{0\}, \mathbb{C}^4)$ and $H^{1/2}(\mathbb{R}^3, \mathbb{C}^4)$! This creates an impressive flexibility in the choice of the basis sets, ensuring that whichever basis set we take in this class, will lead towards the eigenvalues of $H_0 - \nu/|x|$.

Of course, the above results would not be very useful if they were only applicable to Dirac–Coulomb operators. In [31] we describe the class of electrostatic potentials $V$ to which the above results also apply, the main condition being that they are not too positive, so that they are mainly attractive, and also that if they are singular at some point, the singularity cannot exceed $-\nu/|x|$ for $0 < \nu \leq 1$.

The above recent results settle the question of when and how we can run the above algorithm with guarantee that what we will find in the end will be good approximations of the eigenvalues of perturbed Dirac operators.

More recently, with Lewin and Séré we have gone further and consider the case of several singularities, case that is important in molecular computations. In [32, 33], we have extended the results of [31] to this case. And not only to this case, but to the much more general case of Dirac operators perturbed by general singular measures. In this case, which goes beyond Coulomb singularities, more technically refined arguments are needed to describe the domains and prove the validity of the min–max arguments used to find the corresponding eigenvalues. In the above papers, we have also considered other interesting questions that we have not been able to answer completely concerning, for instance, the dependence of the eigenvalues on the geometric shape of the nuclear distribution. A very simple question is: if we consider two identical nuclei of charge $\nu \in (0, 1/2)$, at a distance $R > 0$, would the energy of an electron be above the energy of an electron in the electrostatic field created by a single nucleus of charge $2\nu$? In the non-relativistic case the answer is yes. And not only that, actually the ground state energy grows with the distance $R$. For Dirac operators such results do not exist. Only numerical results have been obtained for instance by Artemyev et al. [45] showing that the dependence of the ground state energy with respect to the distance between the two nuclei follows the same pattern in the relativistic and in the non-relativistic cases.

Acknowledgement

This Note has been prepared while the author was being partially supported by the project molQED, Project-ANR-17-CE29-0004 of the French National Research Agency (ANR).
References


