

# VARIATIONAL METHODS IN RELATIVISTIC QUANTUM MECHANICS

MARIA J. ESTEBAN, MATHIEU LEWIN, AND ERIC SÉRÉ

ABSTRACT. This review is devoted to the study of stationary solutions of linear and nonlinear equations from relativistic quantum mechanics, involving the Dirac operator. The solutions are found as critical points of an energy functional. Contrary to the Laplacian appearing in the equations of nonrelativistic quantum mechanics, the Dirac operator has a negative continuous spectrum which is not bounded from below. This has two main consequences. First, the energy functional is strongly indefinite. Second, the Euler-Lagrange equations are linear or nonlinear eigenvalue problems with eigenvalues lying in a spectral gap (between the negative and positive continuous spectra). Moreover, since we work in the space domain  $\mathbb{R}^3$ , the Palais-Smale condition is not satisfied. For these reasons, the problems discussed in this review pose a challenge in the Calculus of Variations. The existence proofs involve sophisticated tools from nonlinear analysis and have required new variational methods which are now applied to other problems.

In the first part, we consider the *fixed eigenvalue problem* for models of a free self-interacting relativistic particle. They allow to describe the localized state of a spin-1/2 particle (a fermion) which propagates without changing its shape. This includes the Soler models, and the Maxwell-Dirac or Klein-Gordon-Dirac equations.

The second part is devoted to the presentation of min-max principles allowing to characterize and compute the eigenvalues of linear Dirac operators with an external potential, in the gap of their essential spectrum. Many consequences of these min-max characterizations are presented, among them a new kind of Hardy-like inequalities and a stable algorithm to compute the eigenvalues.

In the third part we look for *normalized solutions* of nonlinear eigenvalue problems. The eigenvalues are *Lagrange multipliers*, lying in a spectral gap. We review the results that have been obtained on the Dirac-Fock model which is a nonlinear theory describing the behavior of  $N$  interacting electrons in an external electrostatic field. In particular we focus on the problematic definition of the ground state and its nonrelativistic limit.

In the last part, we present a more involved relativistic model from Quantum Electrodynamics in which the behavior of the vacuum is taken into account, it being coupled to the real particles. The main interesting feature of this model is that the energy functional is now bounded from below, providing us with a good definition of a ground state.

---

*Date:* January 8, 2008. Final version to appear in *Bull. Amer. Math. Soc.*

1991 *Mathematics Subject Classification.* 49S05, 35J60, 35P30, 35Q75, 81Q05, 81V70, 81V45, 81V55.

*Key words and phrases.* Relativistic quantum mechanics, Dirac operator, variational methods, critical points, strongly indefinite functionals, nonlinear eigenvalue problems, ground state, non-relativistic limit, Quantum Chemistry, mean-field approximation, Dirac-Fock equations, Hartree-Fock equations, Bogoliubov-Dirac-Fock method, Quantum Electrodynamics.

## CONTENTS

Introduction	2
Notations and basic properties of the Dirac operator	7
1. Nonlinear Dirac equations for a free particle	8
1.1. Soler models: existence by O.D.E. techniques	9
1.2. Soler models: existence by variational techniques	10
1.3. Existence of solutions by perturbation theory	11
1.4. Nonlinear Dirac equations in the Schwarzschild metric	12
1.5. Solutions of the Maxwell-Dirac equations	12
1.6. Nonlinear Dirac evolution problems	17
2. Linear Dirac equations for an electron in an external field	18
2.1. A variational characterization of the eigenvalues of $D_c + V$	20
2.2. Numerical method based on the min-max formula	25
2.3. New Hardy-like inequalities	27
2.4. The nonrelativistic limit	27
2.5. Introduction of a constant external magnetic field	28
3. The Dirac-Fock equations for atoms and molecules	29
3.1. The (non-relativistic) Hartree-Fock equations	30
3.2. Existence of solutions to the Dirac-Fock equations	31
3.3. Nonrelativistic limit and definition of the Dirac-Fock “ground state”	34
4. The mean-field approximation in Quantum Electrodynamics	36
4.1. Definition of the free vacuum	39
4.2. The Bogoliubov-Dirac-Fock model	42
4.3. Global minimization of $\mathcal{E}_{\text{BDF}}^\nu$ : the polarized vacuum.	45
4.4. Minimization of $\mathcal{E}_{\text{BDF}}^\nu$ in charge sectors	47
4.5. Neglecting Vacuum Polarization: Mittleman’s conjecture	50
References	53

## INTRODUCTION

In this paper, we present various recent results concerning some linear and non-linear variational problems in relativistic quantum mechanics, involving the Dirac operator.

Dirac derived his operator in 1928 [44], starting from the usual classical expression of the energy of a free relativistic particle of momentum  $p \in \mathbb{R}^3$  and mass  $m$

$$(1) \quad E^2 = c^2|p|^2 + m^2c^4$$

( $c$  is the speed of light), and imposing the necessary relativistic invariances. By means of the usual identification

$$p \longleftrightarrow -i\hbar\nabla$$

where  $\hbar$  is Planck’s constant, he found that an adequate observable for describing the energy of the free particle should therefore be a self-adjoint operator  $D_c$  satisfying the equation

$$(2) \quad (D_c)^2 = -c^2\hbar^2\Delta + m^2c^4.$$

Taking the locality principle into account, Dirac proposed to look for a local operator which is first order with respect to  $p = -i\hbar\nabla$ :

$$(3) \quad D_c = -i\hbar \boldsymbol{\alpha} \cdot \nabla + mc^2\beta = -i\hbar \sum_{k=1}^3 \alpha_k \partial_k + mc^2\beta,$$

where  $\alpha_1, \alpha_2, \alpha_3$  and  $\beta$  are hermitian matrices which have to satisfy the following anticommutation relations:

$$(4) \quad \begin{cases} \alpha_k \alpha_\ell + \alpha_\ell \alpha_k &= 2 \delta_{k\ell} \mathbf{1}, \\ \alpha_k \beta + \beta \alpha_k &= 0, \\ \beta^2 &= \mathbf{1}. \end{cases}$$

It can be proved [159] that the smallest dimension in which (4) 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

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

By (1), the time-dependent Dirac equation describing the evolution of a free particle is

$$(5) \quad i\hbar \frac{\partial}{\partial t} \Psi = D_c \Psi.$$

This equation has been successfully used in physics to describe relativistic particles having a spin 1/2.

The main unusual feature of the Dirac equation is the spectrum of  $D_c$  which is not bounded from below:

$$(6) \quad \sigma(D_c) = (-\infty, -mc^2] \cup [mc^2, \infty).$$

Compared with non-relativistic theories in which the Schrödinger operator  $-\Delta/(2m)$  appears instead of  $D_c$ , property (6) leads to important physical, mathematical and numerical difficulties. Indeed, if one simply replaces  $-\Delta/(2m)$  by  $D_c$  in the energies or operators which are commonly used in the non-relativistic case, one obtains energies which are not bounded from below.

Although there is no observable electron of negative energy, the negative spectrum plays an important role in physics. Dirac himself suspected that the negative spectrum of his operator could generate new interesting physical phenomena, and he proposed in 1930 the following interpretation [45, 46, 47]:

“We make the assumption that, in the world as we know it, nearly all the states of negative energy for the electrons are occupied, with just one electron in each state, and that a uniform filling of all the negative-energy states is completely unobservable to us.” [47]

Physically, one therefore has to imagine that the vacuum (called the *Dirac sea*) is filled with infinitely many virtual particles occupying the negative energy states. With this conjecture, a real free electron cannot be in a negative state due to the Pauli principle which forbids it to be in the same state as a virtual electron of the Dirac sea.

With this interpretation, Dirac was able to conjecture the existence of “holes” in the vacuum, interpreted as “anti-electrons” or positrons, having a positive charge and a positive energy. The positron was discovered in 1932 by Anderson [3]. Dirac also predicted the phenomenon of vacuum polarization: in the presence of an electric field, the virtual electrons are displaced, and the vacuum acquires a nonconstant density of charge. All these phenomena are now well known and well established in physics. They are direct consequences of the existence of the negative spectrum of  $D_c$ , showing the crucial role played by Dirac’s discovery.

Actually, in practical computations it is quite difficult to deal properly with the Dirac sea. As a consequence the notion of “ground state” (state of “lowest energy” which is supposed to be the most “stable” for the system under consideration) is problematic for many of the models found in the literature. Numerically, the unboundedness from below of the spectrum is also the source of important practical issues concerning the convergence of the considered algorithms, or the existence of spurious (unphysical) solutions.

Dirac’s interpretation of the negative energies will be an implicit assumption in all this review in the sense that we shall (almost) always look for positive energy solutions for the electrons. In the last section, we present a model from Quantum Electrodynamics (QED) in which the nonlinear behavior of the Dirac sea will be fully taken into account.

Mathematically, most of the energy functionals that we shall consider are strongly indefinite: they are unbounded from below and all their critical points have an infinite Morse index. Note that the mathematical methods allowing to deal with strongly indefinite functionals have their origin in the works of P. Rabinowitz concerning the study of nonlinear waves [134] and also the study of periodic solutions of Hamiltonian systems [135]. Many results have followed these pioneering works, and powerful theories have been devised, in particular in the field of periodic orbits of Hamiltonian systems: the linking theorem of Benci-Rabinowitz [17], Clarke-Ekeland’s dual action functional [39], Conley-Zehnder’s relative index [41], Floer’s homology [74]...

Another difficulty with the models presented in this review is the lack of compactness: the Palais-Smale condition is not satisfied due to the unboundedness of the domain  $\mathbb{R}^3$ . Variational problems with lack of compactness also have been extensively studied. Let us mention the work of Sacks-Uhlenbeck [145] on harmonic maps, Lieb’s Lemma [114], Brezis-Nirenberg’s study of elliptic PDEs with critical exponents [25], the concentration-compactness method of P.-L. Lions [118], Bahri-Coron’s critical points at infinity [10] and more recently Fang-Ghoussoub’s Palais-Smale sequences with Morse information [68, 77].

The combination of the two above types of difficulties poses a challenge in the Calculus of Variations. To prove the results presented in this review, it has been necessary to adapt some of the sophisticated tools mentioned above and to introduce new ideas. The novel variational methods that we have designed can be applied in general situations and in particular in the study of crystalline matter (nonlinear Schrödinger equations with periodic potentials).

The review contains four different parts which are almost independent. The common feature of all the problems addressed in this review is the variational study of linear and nonlinear eigenvalue problems with eigenvalues in spectral gaps. In the nonlinear case, there are two different classes of problems. Either we fix the

eigenvalue and let the  $L^2$ -norm of the solutions free. Or we look for normalized solutions, the eigenvalue is then a Lagrange multiplier which has to stay in the spectral gap.

In the first section, we describe the results that have been obtained for some models describing one self-interacting free relativistic spin-1/2 particle. The simplest case is when the interaction is “local”, i.e. represented by a nonlinear function  $F$  of the spinor  $\psi(t, x)$  of the particle. The general form for the equations that we consider in this part is:

$$D_c\psi - \omega\psi = \nabla F(\psi).$$

These models are phenomenological. A Lorentz-invariant interaction term  $F(\psi)$  is chosen in order to find a model of the free localized electron (or on another spin 1/2 particle), which fits with experimental data (see [138]).

At the end of the first section, we present two nonlocal models: the Maxwell-Dirac and the Klein-Gordon-Dirac equations in which the electron interacts with its own electromagnetic field. The Maxwell-Dirac equations take the following form:

$$\begin{cases} (D_c + v - \boldsymbol{\alpha} \cdot A)\psi = \omega\psi, \\ -4\pi\Delta v = |\psi|^2, \\ -4\pi\Delta A_k = (\psi, \alpha_k\psi), \quad k = 1, 2, 3. \end{cases}$$

From a mathematical viewpoint, the equations considered in the first section are nonlinear eigenvalue problems, in which the eigenvalue is fixed in a spectral gap, but the  $L^2$  norm of the solution is not known. They are the Euler-Lagrange equations of a strongly indefinite functional. Moreover, this functional does not satisfy the Palais-Smale condition and the classical Benci-Rabinowitz linking theorem [17] cannot be applied. The solutions are obtained by a “noncompact” linking argument inspired by the works of Hofer-Wysocki [93] and Séré [150] on homoclinic orbits of Hamiltonian systems. An additional difficulty is that the nonlinearity can vanish even for very large values of  $\psi$ , and this makes the a priori bounds on Palais-Smale sequences very delicate.

The second section is devoted to the study of min-max principles allowing to characterize the eigenvalues of Dirac operators with an external potential  $V$  in the gap of their essential spectrum. Such operators are commonly used to describe the dynamics of an electron which is subject to the action of an external electrostatic field with associated potential  $V$  (for instance an electron in the field created by a nucleus). It can also be used to describe many *non-interacting* electrons. For potentials  $V$  satisfying appropriate assumptions, the spectrum of the perturbed Dirac operator  $D_c + V$  takes the form

$$\sigma(D_c + V) = (-\infty, -mc^2] \cup \{\varepsilon_i\}_{i \in \mathbb{N}} \cup [mc^2, \infty)$$

where the  $\varepsilon_i$ 's are eigenvalues of finite multiplicity in  $(-mc^2, mc^2)$ , which can only accumulate at the thresholds  $-mc^2$  or  $mc^2$  (see [18, 19]). The min-max formulas presented in this section furnish a useful variational characterization of the  $\varepsilon_i$ 's and of the associated eigenfunctions:

$$(D_c + V)\varphi_i = \varepsilon_i\varphi_i.$$

The min-max formulas are general and can be used in other settings where eigenvalues in a gap of the essential spectrum have to be characterized or computed. Many consequences of the min-max principles are derived in this section, including an algorithm for the computation of the eigenvalues.

In Section 3, we present results concerning the Dirac-Fock model [156], allowing to describe  $N$  *interacting* electrons in an external electrostatic field. This is a nonlinear model which is used in Quantum Chemistry to compute the state of such electrons in heavy atoms. The energy functional is strongly indefinite and therefore it is really not obvious how to find an adequate definition of the ground state, and to prove the existence of critical points in general. Explaining how this can be achieved is the main goal of the section. The model consists of a system of  $N$  coupled nonlinear equations, posed on  $L^2(\mathbb{R}^3, \mathbb{C}^4)^N$ :

$$D_{c,\Phi} \varphi_i = \varepsilon_i \varphi_i, \quad 0 < \varepsilon_i < mc^2,$$

where

$$D_{c,\Phi} = D_c + V + \Gamma_\Phi,$$

$\Gamma_\Phi$  being an operator depending nonlinearly on  $\Phi := (\varphi_1, \dots, \varphi_N)$  and which models the interactions between the electrons. The functions  $\varphi_i$ 's, which are assumed to satisfy the constraints  $\int_{\mathbb{R}^3} (\varphi_i, \varphi_j) = \delta_{ij}$ , represent the states of the  $N$  electrons. Being a system of nonlinear eigenvalue problems with eigenvalues in a spectral gap, the Dirac-Fock equations carry some similarity with the equations studied in Section 1. But there is a big difference: the  $L^2$  norm of the solutions is now fixed *a priori*, and the eigenvalues  $\varepsilon_i$  are unknown Lagrange multipliers associated with these constraints. This makes the problem harder, the main difficulty being to keep the multipliers  $\varepsilon_i$  in the interval  $(0, mc^2)$ . The positivity of  $\varepsilon_i$  is obtained thanks to a new penalization method. See [26] for a generalization of this method, with applications to nonlinear periodic Schrödinger models for crystals. The inequality  $\varepsilon_i < mc^2$  follows from Morse-type estimates, as in the existence proof of Lions for the nonrelativistic Hartree-Fock model [119]. To obtain these Morse-type estimates, the easiest way is to use a general theorem of Fang-Ghoussoub [68]. Note that, since the functional is strongly indefinite, one has to work in fact with a relative Morse index.

Finally, in the last section we present a more involved physical model in which the behavior of the electrons is coupled to that of the Dirac sea, in the presence of an external electrostatic field  $V$ . In this model, Dirac's interpretation of the negative energies is really taken into account: the vacuum is considered as being an unknown physical object which can react to an external stimulation. The important feature of the model will be that the energy functional is *bounded from below*, as first proposed by Chaix and Iracane [36], showing the importance of the vacuum polarization effects. The main drawback will be that one necessarily has to deal with infinitely many interacting particles (the real ones and the virtual ones of the Dirac sea), which creates lots of mathematical difficulties. In particular, the main unknown of the model is, this time, an orthogonal projector  $P$  of infinite rank. The optimal projector  $P$  representing the ground state of the system is solution of a nonlinear equation of the form

$$(7) \quad P = \chi_{(-\infty, \mu]}(D_c + V + \Gamma'_P)$$

where  $\Gamma'_P$  is an operator depending on the projector  $P$  and describing the interactions between all particles (the real and the virtual ones). We have used the standard notation  $\chi_I(A)$  for the spectral projector of  $A$  associated with the interval  $I$ . Solutions of (7) are obtained by a minimization principle, on a set of compact operators. One has to be very careful in the choice of this set, and in the definition of the energy. A serious difficulty is the presence of ultraviolet divergencies.

**Acknowledgment.** First of all, M.J.E. and E.S. would like to thank T. Cazenave, I. Ekeland, P.-L. Lions and L. Vázquez for having introduced them to this field. We would like also to thank J. Dolbeault, V. Georgiev, C. Hainzl, M. Loss and J.-P. Solovej with whom we have collaborated on different parts of the works described in this review. Some years ago P. Chaix was very patient in explaining to us some models of QED on which we have worked later; the last section of this paper has its origin in those conversations. Several discussions with B. Buffoni were very inspiring to us when looking for new variational principles in the case of strongly indefiniteness functionals. We want also to acknowledge our participation in the european research projects “Analysis and Quantum” (HPRN-CT-2002-00277) and “HYKE” (HPRN-CT-2002-00282). The interaction with our partners in those two projects has been a source of inspiration and motivation for us along the years and we thank them for that. We also acknowledge support from the ANR “ACCQUAREL” of the french ministry of research. E.S. was partially supported by the “Institut Universitaire de France”. Finally, we wish to thank P. Rabinowitz, who has been very patient encouraging us to write this review article.

### Notations and basic properties of the Dirac operator.

Before going further, let us fix some notations. All throughout this review, the conjugate of  $z \in \mathbb{C}$  will be denoted by  $z^*$ . For  $X = (z_1, \dots, z_4)^T$  a column vector in  $\mathbb{C}^4$ , we denote by  $X^*$  the row co-vector  $(z_1^*, \dots, z_4^*)$ . Similarly, if  $A = (a_{ij})$  is a  $4 \times 4$  complex matrix, we denote by  $A^*$  its adjoint,  $(A^*)_{ij} = a_{ji}^*$ .

We denote by  $(X, X')$  the Hermitian product of two vectors  $X, X'$  in  $\mathbb{C}^4$ , and by  $|X|$ , the canonical hermitian norm of  $X$  in  $\mathbb{C}^4$ , i. e.  $|X|^2 = \sum_{i=1}^4 X_i^* X_i$ . The usual Hermitian product in  $L^2(\mathbb{R}^3, \mathbb{C}^4)$  is denoted as

$$(8) \quad (\psi, \psi')_{L^2} = \int_{\mathbb{R}^3} (\psi(x), \psi'(x)) d^3x.$$

For the sake of simplicity, we shall use a system of units in which

$$m = \hbar = 1.$$

Actually, by scaling one can also fix the value of another physical constant, like for instance the speed of light  $c$  or the charge of an electron  $e$ . We shall use both possibilities in this review (they are of course equivalent).

Let us now list some basic and important properties of the free Dirac operator. We refer to the book of Thaller [159] for details.

**Proposition 1** (Basic properties of the free Dirac operator). *The free Dirac operator  $D_c$  is a self-adjoint operator on  $L^2(\mathbb{R}^3, \mathbb{C}^4)$ , with domain  $H^1(\mathbb{R}^3, \mathbb{C}^4)$  and form-domain  $H^{1/2}(\mathbb{R}^3, \mathbb{C}^4)$ . Its spectrum is purely continuous,  $\sigma(D_c) = (-\infty, -c^2] \cup [c^2, +\infty)$ . Moreover, there are two orthogonal projectors (both having infinite rank) from  $L^2(\mathbb{R}^3, \mathbb{C}^4)$  into itself,  $P_{+,c}^0$  and  $P_{-,c}^0 = \mathbf{1}_{L^2} - P_{+,c}^0$ , such that*

$$(9) \quad \begin{cases} D_c P_{+,c}^0 = P_{+,c}^0 D_c = \sqrt{c^4 - c^2 \Delta} P_{+,c}^0 = P_{+,c}^0 \sqrt{c^4 - c^2 \Delta} \\ D_c P_{-,c}^0 = P_{-,c}^0 D_c = -\sqrt{c^4 - c^2 \Delta} P_{-,c}^0 = -P_{-,c}^0 \sqrt{c^4 - c^2 \Delta}. \end{cases}$$

The projectors  $P_{+,c}^0$  and  $P_{-,c}^0$  are multiplication operators in the Fourier domain, given by

$$(10) \quad P_{\pm,c}^0(p) = \frac{\pm D_c(p) + \sqrt{c^2 |p|^2 + c^4}}{2\sqrt{c^2 |p|^2 + c^4}}.$$

Note that Proposition 1 enables us to split the space

$$\mathfrak{H} := L^2(\mathbb{R}^3, \mathbb{C}^4)$$

as the direct sum of two infinite dimensional Hilbert spaces  $\mathfrak{H}_{\pm,c}^0 = P_{\pm,c}^0 \mathfrak{H}$ . The restriction of  $D_c$  to  $\mathfrak{H}_{\pm,c}^0$  is a self-adjoint operator in this subspace, with domain  $\mathfrak{H}_{\pm,c}^0 \cap H^1(\mathbb{R}^3, \mathbb{C}^4)$ . Furthermore, it will be convenient to use the following norm in  $\mathfrak{H}$ , equivalent to the usual norm of  $H^{1/2}(\mathbb{R}^3, \mathbb{C}^4)$ ,

$$(11) \quad \|\psi\|_{\mathfrak{H}} := \left( \psi, (D_c P_{+,c}^0 - D_c P_{-,c}^0) \psi \right)_{\mathfrak{H} \times \mathfrak{H}'}^{1/2} = \left( \psi, |D_c| \psi \right)^{1/2}$$

The subspaces  $\mathfrak{H}_{+,c}^0 \cap H^{1/2}(\mathbb{R}^3, \mathbb{C}^4)$  and  $\mathfrak{H}_{-,c}^0 \cap H^{1/2}(\mathbb{R}^3, \mathbb{C}^4)$  are orthogonal for this norm as well as for the  $L^2$  norm.

When  $c = 1$ , one recovers the usual  $H^{1/2}$  norm. In this case, we shall use the convenient notation  $P_{\pm,1}^0 = P_{\pm}^0$  and  $\mathfrak{H}_{\pm,1}^0 = \mathfrak{H}_{\pm}^0$ .

## 1. NONLINEAR DIRAC EQUATIONS FOR A FREE PARTICLE

In this section, we present some nonlinear Dirac equations for a free particle. We therefore do not consider any external potential (but possibly a self-generated one). Stationary solutions of such equations represent the state of a localized particle which can propagate without changing its shape. The first to propose and study models for the description of this phenomenon were Ivanenko [97], Weyl [169] and Heisenberg [90]. We refer to Rañada [138] for a very interesting review on the historical background of this kind of models.

In this section, we shall always assume that  $c = 1$ . A general form for the equations that we want to present is

$$(12) \quad i\partial_t \Psi - D_1 \Psi + \nabla F(\Psi) = 0,$$

where  $\Psi(t, \cdot) \in L^2(\mathbb{R}^3, \mathbb{C}^4)$ . Throughout this chapter we assume that  $F \in C^2(\mathbb{C}^4, \mathbb{R})$  satisfies

$$(13) \quad F(e^{i\theta} \Psi) = F(\Psi) \quad \text{for all } \theta \in \mathbb{R}.$$

The charge of the particle  $-e$  does not appear since it is incorporated into the nonlinear functional  $F$ .

The relativistic invariance requirement imposes severe restrictions on the possible nonlinearities  $F$ . In two very interesting papers [69, 70], Finkelstein *et al* proposed various models for extended particles corresponding to various fourth order self-couplings  $F$ . In those papers, they gave some numerical description of the structure of the set of solutions, for different values of the parameters. Among the considered functions  $F$ , in [69, 70] we find the family of general self-couplings

$$(14) \quad F_b(\Psi) := \lambda \left\{ (\bar{\Psi} \Psi)^2 + b (\bar{\Psi} \gamma^5 \Psi)^2 \right\}$$

where  $\gamma^5 = -i\alpha_1 \alpha_2 \alpha_3$ ,  $b$  is a real constant, and where we have used the notation

$$\bar{\Psi} \Psi := (\beta \Psi, \Psi).$$

In the sequel, without any loss of generality, we will assume that  $\lambda = 1/2$ .

Stationary solutions of (12) are functions of the type

$$(15) \quad \Psi(t, x) = e^{-i\omega t} \psi(x),$$

such that  $\psi$  is a non-zero localized solution of the following stationary nonlinear Dirac equation

$$(16) \quad D_1\psi - \omega\psi - \nabla F(\psi) = 0 \quad \text{in } \mathbb{R}^3.$$

It is interesting to note that the latter equation has a variational structure: it is indeed the Euler-Lagrange equation associated with the functional

$$(17) \quad I^\omega(\psi) = \int_{\mathbb{R}^3} \left( \frac{1}{2}(D_1\psi(x), \psi(x)) - \frac{\omega}{2}|\psi(x)|^2 - F(\psi(x)) \right) dx.$$

Hence, the solutions of (16) are formally the critical points of the “energy” functional  $I^\omega$ . In this context, we say that  $\psi$  is a *localized* solution if  $I^\omega(\psi)$  is well-defined, that is, if  $\psi \in H^{1/2}(\mathbb{R}^3, \mathbb{C}^4)$  and  $F(\psi) \in L^1(\mathbb{R}^3, \mathbb{R})$ . Due to the structure of the Dirac operator  $D_1$ , the functional  $I^\omega$  is not bounded below and solutions of (16) cannot be obtained by a minimization method.

**1.1. Soler models: existence by O.D.E. techniques.** The case  $b = 0$  in the definition (14) of  $F_b$  was proposed by Soler in [154] to describe elementary fermions. In this case, (17) reduces to

$$(18) \quad D_1\psi - \omega\psi - (\bar{\psi}\psi)\beta\psi = 0 \quad \text{in } \mathbb{R}^3$$

which is usually called the *Soler model*. Its more general version

$$(19) \quad D_1\psi - \omega\psi - g(\bar{\psi}\psi)\beta\psi = 0 \quad \text{in } \mathbb{R}^3$$

is often called the *generalized Soler equation*, and it is the Euler-Lagrange equation associated with  $I^\omega$  for  $F(\psi) = \frac{1}{2}G(\bar{\psi}\psi)$ ,  $G' = g$ ,  $G(0) = 0$ .

The main advantage of (19) is the existence of a special ansatz compatible with the equation:

$$(20) \quad \psi(x) = \begin{pmatrix} v(r) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ iu(r) \begin{pmatrix} \cos \theta \\ \sin \theta e^{i\xi} \end{pmatrix} \end{pmatrix}.$$

In this ansatz, equation (19) reduces to the O.D.E. system

$$(21) \quad \begin{cases} (u' + \frac{2u}{r}) & = v(g(v^2 - u^2) - (1 - \omega)) \\ v' & = u(g(v^2 - u^2) - (1 + \omega)) \end{cases}$$

The O.D.E. system (21) has been extensively studied. In [164] Vázquez proved some qualitative properties of the solutions in the case  $0 < \omega < 1$ , and showed the non-existence of localized solutions when  $|\omega| > 1$ . Cazenave and Vázquez obtained the first rigorous existence result for this problem in [33]. More precisely, in [33] they proved the existence of a solution  $(u, v)$  to (21) for nonlinearities  $g$  of class  $C^1$  satisfying:

$$(22) \quad g : [0, +\infty) \rightarrow [0, +\infty), \quad g(0) = 0, \quad g' > 0, \quad \lim_{s \rightarrow +\infty} g(s) = +\infty,$$

whenever  $0 < \omega < 1$ . Moreover, this solution  $(u, v)$  is such that  $u$  and  $v$  are positive on  $\mathbb{R}^+$ ,  $u(0) = 0$ ,  $v(0) > 0$ . Additionally,  $u$  and  $v$  (as well as their first derivatives) decay exponentially at infinity.

Later on, Merle [124] improved the above result by extending it to a more general class of nonlinearities  $g$ . Then, Balabane *et al* proved the following general multiplicity result:

**Theorem 2.** ([11]) *Assume that  $g : \mathbb{R} \rightarrow \mathbb{R}$  is a function of class  $C^1$  such that  $g(0) = 0$ ,  $g$  is increasing in  $(0, +\infty)$ ,  $g(s) > 1 + \omega$  for  $s$  large,  $g'(g^{-1}(1 - \omega)) > 0$  and  $g(s) \leq 0$  for  $s \leq 0$ . Then, for any  $\omega \in (0, 1)$ , there exists an increasing sequence of positive numbers  $\{x_n\}_{n \geq 1}$  such that for every  $n \geq 1$ , there is a solution  $(u_n, v_n)$  of (21) satisfying*

- $u_n(0) = 0, v_n(0) = x_n$ ,
- $u_n$  and  $v_n$  have  $n - 1$  zeros in  $\mathbb{R}^+$ ,
- $u_n$  and  $v_n$  decay exponentially at infinity, as well as their first derivatives.

Moreover, if for all  $s$ ,  $g(s) = s$ , then the sequence  $\{x_n\}$  is bounded.

In the case of singular nonlinearities, compactly supported solutions may exist. More concretely, the following result was proved in [12]:

**Theorem 3.** ([12]) *Assume that  $g : (0, +\infty) \rightarrow (-\infty, 0)$  is a function of class  $C^1$  which is nondecreasing and integrable near the origin. Suppose also that there exists a number  $a$  such that  $g(a^2) = 1 - \omega$ , while  $g'(a^2) > 0$ . Then, for every  $0 < \omega < 1$  there exists a positive solution  $(u, v)$  of (21). Moreover, the support of  $(u, v)$  is bounded if and only if*

$$\int_0^1 \frac{ds}{G(s)} < +\infty, \quad \text{where } G(s) := - \int_0^s g(x) dx.$$

**1.2. Soler models: existence by variational techniques.** All the above results were obtained by a dynamical systems approach. But it is also possible to exploit the variational structure of (19) (and also of the O.D.E. system (21)) to prove existence of solutions.

In the case of (21), the use of variational methods does not radically improve the results that were obtained by O.D.E. methods (see above). The assumptions needed to use variational techniques are slightly different. In [62], Esteban and Séré obtained the following result:

**Theorem 4.** ([62]) *Let  $F : \mathbb{C}^4 \rightarrow \mathbb{R}$  satisfy*

$$F(\psi) = \frac{1}{2}G(\bar{\psi}\psi), \quad G \in C^2(\mathbb{R}, \mathbb{R}), \quad G(0) = 0,$$

with  $G \in C^2(\mathbb{R}, \mathbb{R})$ . Denoting by  $g$  the first derivative of  $G$ , we make the following assumptions:

$$(23) \quad \forall x \in \mathbb{R}, \quad xg(x) \geq \theta G(x) \quad \text{for some } \theta > 1,$$

$$(24) \quad G(0) = g(0) = 0,$$

$$(25) \quad (\forall x \in \mathbb{R}, \quad G(x) \geq 0) \quad \text{and} \quad G(A_0) > 0 \quad \text{for some } A_0 > 0,$$

$$(26) \quad 0 < \omega < 1.$$

Then there exist infinitely many solutions of Equation (16) in  $\bigcap_{2 \leq q < +\infty} W^{1,q}(\mathbb{R}^3, \mathbb{C}^4)$ . Each of them solves a min-max variational problem on the functional  $I^\omega$ . They are of the form (20) and thus correspond to classical solutions of (21) on  $\mathbb{R}_+$ . Finally, they all decrease exponentially at infinity, together with their first derivatives.

The interest of using variational techniques appears much more clearly when one studies equations for which no simplifying ansatz is known, for instance in the case where  $F = F_b$  with  $b \neq 0$ . In that case, Equation (16) cannot be reduced to a system of O.D.Es similar to (21). A general result proved by Esteban and Séré in this context is the following:

**Theorem 5.** ([62]) *Let be  $F(\psi) = \lambda(|\psi\bar{\psi}|^{\alpha_1} + b|\bar{\psi}\gamma^5\psi|_2^{\alpha_2})$ , with  $1 < \alpha_1, \alpha_2 < \frac{3}{2}$ ;  $\lambda, b > 0$ . Then, for every  $\omega \in (0, 1)$ , there exists a non-zero solution of (16) in  $\bigcap_{1 \leq q < +\infty} W^{1,q}(\mathbb{R}^3, \mathbb{C}^4)$ .*

In fact, Theorem 5 is a consequence of the more general following result in [62]

**Theorem 6.** ([62]) *Assume that  $F : \mathbb{C}^2 \rightarrow \mathbb{R}$  satisfies :*

$$(27) \quad \forall \psi \in \mathbb{C}^4, \quad 0 \leq F(\psi) \leq a_1(|\psi|^{\alpha_1} + |\psi|^{\alpha_2}),$$

with  $a_1 > 0$  and  $2 < \alpha_1 \leq \alpha_2 < 3$ . Assume moreover that

$$(28) \quad \begin{cases} F \in C^2(\mathbb{C}^4, \mathbb{R}), & F'(0) = F''(0) = 0, \\ |F''(\psi)| \leq a_2|\psi|^{\alpha_2-2}, & a_2 > 0, \text{ for } |\psi| \text{ large;} \end{cases}$$

$$(29) \quad \forall \psi \in \mathbb{C}^4, \quad \nabla F(\psi) \cdot \psi \geq a_3 F(\psi), \quad a_3 > 2;$$

$$(30) \quad \exists a_4 > 3, \forall \delta > 0, \exists C_\delta > 0, \forall \psi \in \mathbb{C}^4, \quad |\nabla F(\psi)| \leq \left( \delta + C_\delta F(\psi)^{\frac{1}{a_4}} \right) |\psi|;$$

$$(31) \quad \forall \psi \in \mathbb{C}^4, \quad F(\psi) \geq a_5 |\psi\bar{\psi}|^\nu - a_6, \quad \nu > 1, a_5, a_6 > 0.$$

Then, for every  $\omega \in (0, 1)$ , there exists a non-zero solution of (16) in  $\bigcap_{2 \leq q < +\infty} W^{1,q}(\mathbb{R}^3, \mathbb{C}^4)$ .

**1.3. Existence of solutions by perturbation theory.** Another way of finding solutions to nonlinear Dirac equations is perturbation theory. In this approach, one uses previously known information about the nonlinear Schrödinger equation [168], which is approached in the non-relativistic limit. Ounaies proved in [130] that solutions of some nonlinear Dirac equations, when properly rescaled, are close to solutions of the nonlinear Schrödinger equation, with the same nonlinearity, when the phase  $\omega$  approaches 1. More precisely, assume for instance that

$$F(\psi) := \frac{1}{2} (G(\bar{\psi}\psi) + H(\bar{\psi}\gamma^5\psi)),$$

where  $G, H$  are two functions of class  $C^2$  such that  $G(0) = H(0) = 0$ , and such that  $g := G'$  and  $h := H'$  are homogeneous of degree  $\theta \in (0, 1]$ . Then, if we write any 4-spinor  $\psi$  as  $\psi = \begin{pmatrix} \varphi \\ \chi \end{pmatrix}$ , the main theorem in [130] states the following

**Theorem 7.** ([130]) *Under the above assumptions, let  $1 - \omega = a^{2\theta} = \lambda^2 = b^{\frac{2\theta}{\theta+1}} = \varepsilon$ . If we rescale the functions  $\varphi, \chi$  as follows*

$$\varphi(x) := a\bar{\varphi}(\lambda x), \quad \chi(x) := b\bar{\chi}(\lambda x),$$

then,  $\psi = \begin{pmatrix} \varphi \\ \chi \end{pmatrix}$  is a solution to (16) if and only if  $\bar{\varphi}, \bar{\chi}$  are solutions to the system

$$\begin{cases} (-i\sigma \cdot \nabla)\bar{\chi} + \bar{\varphi} - g(|\bar{\varphi}|^2)\bar{\varphi} + K_1(\varepsilon, \bar{\varphi}, \bar{\chi}) & = 0 \\ (-i\sigma \cdot \nabla)\bar{\varphi} - 2\bar{\chi} + K_2(\varepsilon, \bar{\varphi}, \bar{\chi}) & = 0. \end{cases}$$

Here  $K_1$  and  $K_2$  are small functions for small  $\varepsilon$ ,  $\bar{\varphi}$  and  $\bar{\chi}$  taking values in a bounded set of  $\mathbb{C}^2$ . Moreover, for  $\varepsilon$  small enough, there exist solutions to the above equation. They are close to a solution of the nonlinear Schrödinger equation

$$(32) \quad -\frac{1}{2}\Delta\bar{\varphi} + \bar{\varphi} - g(|\bar{\varphi}|^2)\bar{\varphi} = 0, \quad \bar{\chi} = -\frac{i}{2}(\sigma \cdot \nabla)\bar{\varphi}.$$

**Remark 8.** Note that the function  $h = H'$  does not appear in the above limiting equation.

The proof of this theorem makes use of the implicit function theorem in an appropriate manner. Important ingredients are the uniqueness (up to translation) of the solution to the elliptic equation (32) for  $\bar{\varphi}$  and its nondegeneracy [106, 40, 168].

**1.4. Nonlinear Dirac equations in the Schwarzschild metric.** All the above models correspond to the Dirac equation written in the Minkowski metric, this is, in flat space. But the space-time geometry plays an important role when one wants to take relativistic effects into account. For instance, when considering the Schwarzschild metric outside a massive star, the nonlinear Dirac equation appears to be different.

In [9] A. Bachelot-Motet has studied numerically this problem in the case of the symmetric solutions as above. One has to study a system of O.D.Es similar to (21) but with  $r$ -dependent coefficients instead of constant ones. More precisely, in the ansatz (20) and for the case  $F(s) = \lambda|s|^2$ , system (21) becomes

$$(33) \quad \begin{aligned} fu' + \frac{u}{r}(f + f^{1/2}) &= v[\lambda(v^2 - u^2) - (f^{1/2} - \omega)] \\ fv' + \frac{v}{r}(f - f^{1/2}) &= u[\lambda(v^2 - u^2) - (f^{1/2} + \omega)], \end{aligned}$$

where  $f(r) = 1 - \frac{1}{r}$ .

Notice that this problem is not to be considered in the whole space: since the physical situation corresponds to the outside of a massive star, the natural domain is the complement of a ball, or in radial coordinates, the interval  $r > r_0$ , for some  $r_0 > 0$ . In this case, the usual ‘‘MIT-bag’’ boundary condition reads

$$u(r_0) = -v(r_0).$$

The very interesting numerical results obtained by Bachelot-Motet suggested conditions for some existence and multiplicity results for (33) that were later rigorously proved by Patrel in [131]. Note that in [131] the solutions are found as critical points of a reduced energy functional by a mountain-pass argument, while as we see below, we use a linking method to produce our solutions.

**1.5. Solutions of the Maxwell-Dirac equations.** The nonlinear terms appearing in all the above models are local, that is, are functions of the spinor field  $\Psi$ . But in some cases, one has to introduce nonlocal terms, like for instance when considering the interaction of the Dirac field with a self-generated field. In this case, the equations become integro-differential.

Our first example is the Maxwell-Dirac system of classical field equations, describing the interaction of a particle with its self-generated electromagnetic field. In order to write the equations in relativistically covariant form, we introduce the usual four-dimensional notations: let be  $\gamma^0 := \beta$  and  $\gamma^k := \beta\alpha_k$ . For any wavefunction  $\Psi(x_0, x) : \mathbb{R} \times \mathbb{R}^3 \mapsto \mathbb{C}^4$  (note that  $x_0$  plays the role of the time  $t$ ), we

denote  $\bar{\Psi} := \beta\Psi$ . In the Lorentz gauge the Maxwell-Dirac equations can be written as follows

$$(34) \quad \begin{cases} (i\gamma^\mu \partial_\mu - \gamma^\mu A_\mu - 1)\Psi = 0 & \text{in } \mathbb{R} \times \mathbb{R}^3 \\ \partial_\mu A^\mu = 0, \quad 4\pi \partial_\mu \partial^\mu A_\nu = (\bar{\Psi}, \gamma^\nu \Psi) & \text{in } \mathbb{R} \times \mathbb{R}^3. \end{cases}$$

Notice that we have used Einstein's convention for the summation over  $\mu$ . We also introduce the electromagnetic current  $J^\nu := (\bar{\Psi}, \gamma^\nu \Psi)$ .

Finite energy stationary solutions of classical nonlinear wave equations have been sometimes used to describe extended particles. Of course the electromagnetic field should in principle be quantized like in Quantum Electrodynamics. In the Maxwell-Dirac model, the field is not quantized but it is believed that interesting qualitative results can be obtained by using classical fields (see, e.g. [79, Chapter 7]).

Another example of a self-interaction is the Klein-Gordon-Dirac system which arises in the so-called Yukawa model (see, e.g. [20]). It can be written as

$$(35) \quad \begin{cases} (i\gamma^\mu \partial_\mu - \chi - 1)\Psi = 0 & \text{in } \mathbb{R} \times \mathbb{R}^3 \\ \partial^\mu \partial_\mu \chi + M^2 \chi = \frac{1}{4\pi} (\bar{\Psi}, \Psi) & \text{in } \mathbb{R} \times \mathbb{R}^3. \end{cases}$$

Other related models, that we will not discuss, include the Einstein-Dirac-Maxwell equations, which have been investigated by F. Finster, J. Smoller and S.-T. Yau [71], [72]. The above systems (34) and (35) have been extensively studied and many results are available concerning the Cauchy problem (we refer to [61] and [79], chapter 7, for detailed references).

A stationary solution of the Maxwell-Dirac system (34) is a particular solution  $(\Psi, A) : \mathbb{R} \times \mathbb{R}^4 \mapsto \mathbb{C}^4 \times \mathbb{R}^4$  of the form

$$(36) \quad \begin{cases} \Psi(x_0, x) = e^{-i\omega x_0} \psi(x) & \text{with } \psi : \mathbb{R}^3 \rightarrow \mathbb{C}^4, \\ A^\mu(x_0, x) = J^\mu * \frac{1}{|x|} = \int_{\mathbb{R}^3} \frac{dy}{|x-y|} J^\mu(y). \end{cases}$$

The existence of such stationary solutions of (34) has been an open problem for a long time (see, e.g. [79, p. 235]). Indeed, the interaction between the spinor and its own electromagnetic field makes equations (34) nonlinear.

Concerning stationary solutions of (34), let us mention the pioneering works of Finkelstein et al [69] and Wakano [166]. The latter considered this system in the approximation  $A_0 \neq 0$ ,  $A_1 = A_2 = A_3 \equiv 0$ , the so-called Dirac-Poisson system. This problem can be reduced to a system of three coupled differential equations by using the spherical spinors (20). Wakano obtained numerical evidence for the existence of stationary solutions of the Dirac-Poisson equation. Further work in this direction (see [138]) yielded the same kind of numerical results for some modified Maxwell-Dirac equations which include some nonlinear self-coupling.

In [120] Lisi found numerical solutions of the Dirac-Poisson and of the Maxwell-Dirac systems. The computation of the magnetic part of the field  $A$  for these solutions showed that Wakano's approximation was reasonable, since the field components  $(A_1, A_2, A_3)$  stay small compared with  $A_0$ . See also [136, 137, 24] where various kinds of stationary solutions are considered, like the so-called *static solutions* which have no current "flow".

In the case  $-1 < \omega < 0$ , Esteban, Georgiev and Séré [61] used variational techniques to prove the existence of stationary solutions of (34). Any solution of (34) taking the form (36) corresponds (formally) to a critical point  $\psi$  of the following

functional:

$$I_\omega(\psi) = \int_{\mathbb{R}^3} \frac{1}{2} (i\alpha^k \partial_k \psi, \psi) - \frac{1}{2} (\bar{\psi}, \psi) - \frac{\omega}{2} |\psi|^2 - \frac{1}{4} \iint_{\mathbb{R}^3 \times \mathbb{R}^3} \frac{J^\mu(x) J_\mu(y)}{|x-y|} dx dy.$$

This remark was used in [61] to find a stationary solution of (34) in the appropriate space of functions.

**Theorem 9.** ([61]) *For any  $\omega$  strictly between  $-1$  and  $0$ , there exists a non-zero critical point  $\psi_\omega$  of  $I_\omega$ . This function  $\psi_\omega$  is smooth in  $x$  and exponentially decreasing at infinity. Finally, the fields  $\Psi(x_0, x) = e^{-i\omega x_0} \psi_\omega$ ,  $A^\mu(x_0, x) = J_\omega^\mu * \frac{1}{|x|}$  are solutions of the Maxwell-Dirac system (34).*

Later, using cylindrical coordinates, S. Abenda [1] extended the above result to the case  $-1 < \omega < 1$ . Indeed, in the class of cylindrically symmetric functions, the energy functional has better properties which allow to use the same variational procedure as in the work of Esteban-Séré, but in the more general case  $\omega \in (-1, 1)$ .

Many questions are still open about the existence of stationary solutions for (34). It is easy to see that they have all a negative “mass”. Wakano already observed this phenomenon for the soliton-like solutions of the Dirac-Poisson system. However, it was shown in [166] that a positive mass can be reached by taking into account the vacuum polarization effect.

For the case of the Klein-Gordon-Dirac equations the situation is slightly simpler because they are compatible with the ansatz (20) introduced above. So, in this case the authors of [61] did not only obtain existence of solutions, but also multiplicity:

**Theorem 10.** ([61]) *For any  $\omega$  strictly between  $-1$  and  $0$ , there exists infinitely many solutions to the Klein-Gordon-Dirac system (35). These solutions are all smooth and exponentially decreasing at infinity in  $x$ .*

We finish this section by explaining the general ideas of the proof of Theorem 9. The proofs of Theorems 4, 5 and 6 basically follow the same lines and we will skip them.

**Sketch of the proof of Theorem 9.** As already mentioned in the introduction, the presence of the negative spectrum for the Dirac operator forbids the use of a minimization argument to construct critical points. Instead, the solution will be obtained by means of a min-max variational method based on complicated topological arguments. This kind of method used to treat problems with infinite negative and positive spectrum have been already used under the name of *linking*. The linking method was introduced by V. Benci and P. Rabinowitz in a compact context [17]. The reasons making the use of variational arguments nonstandard in our case are : (1) the equations are translation invariant, which creates a lack of compactness; (2) the interaction term  $J^\mu A_\mu$  is not positive definite. Note that as we already pointed out, in some cases one can perform a reduction procedure and obtain a reduced functional for which critical points can be found by a mountain-pass argument [131].

**First step: Estimates.** Defining  $A^\mu = J^\mu * \frac{1}{|x|}$ , then one deduces  $J^\mu A_\mu = J^0 A^0 - \sum_{k=1}^3 J^k A^k$  and

$$L(\psi) := \iint_{\mathbb{R}^3 \times \mathbb{R}^3} \frac{J^\mu(x) J_\mu(y)}{|x-y|} dx dy = \int_{\mathbb{R}^3} J^\mu A_\mu dx.$$

Let us also introduce the functional

$$(2.2) \quad Q(\psi) = \iint_{\mathbb{R}^3 \times \mathbb{R}^3} \frac{(\bar{\psi}, \psi)(x) (\bar{\psi}, \psi)(y)}{|x-y|} dx dy.$$

It is easy to prove that  $Q$  is non-negative, continuous and convex on  $\mathfrak{H} = H^{1/2}(\mathbb{R}^3, \mathbb{C}^4)$ , and vanishes only when  $(\bar{\psi}, \psi)(x) = 0$  a.e. in  $\mathbb{R}^3$ .

Let us state a lemma giving some properties of the quadratic forms in  $I_\omega$ :

**Lemma 11.** *For any  $\psi \in \mathfrak{H}$ , the following inequalities hold:*

- (i)  $J^\mu A_\mu(x) \geq 0$ , a.e. in  $\mathbb{R}^3$ ,
- (ii)  $\int_{\mathbb{R}^3} J^\mu A_\mu \geq Q(\psi)$ ,
- (iii)  $A_0 \geq \left( \sum_{k=1}^3 |A_k|^2 \right)^{1/2}$ ,
- (iv)  $|\gamma^\mu A_\mu \psi| \leq C \sqrt{A^0} \sqrt{A^\mu J_\mu}$  a.e. in  $\mathbb{R}^3$ .

**Remark 12.** *Note that when the function  $\psi$  is cylindrically symmetric, the functional  $L$  defined above is not only non-negative, but actually controls from below  $\|\psi\|_{\mathfrak{H}}^4$  (see Lemma 1 in [1]). This is the reason why Abenda has been able to treat the case  $\omega \in (-1, 1)$ , extending Theorem 9.*

Another important information is given by the

**Lemma 13.** *Let be  $\mu > 0$ . There is a non-zero function  $e_+ : (0; \infty) \rightarrow \mathfrak{H}_1^+ = \Lambda_1^+ \mathfrak{H}$  such that, if  $\Lambda^+ \psi = e_+(\mu)$ , then*

$$\frac{1}{2} \int_{\mathbb{R}^3} (\psi, D_1 \psi) - \frac{1}{4} Q(\psi) \leq \frac{\mu}{2} \|\psi\|_{L^2}^2.$$

**Second step: Modified functional and variational argument.** In order to obtain some coercivity, a modified functional  $I_{\omega, \varepsilon}$  was considered in [61]. It reads

$$I_{\omega, \varepsilon}(\psi) = I_\omega(\psi) - \frac{2\varepsilon}{5} \|\psi\|_{L^{5/2}}^{5/2}.$$

where  $\varepsilon > 0$ . The critical points of  $I_{\omega, \varepsilon}(\psi)$  satisfy

$$(37) \quad \begin{cases} i\gamma^k \partial_k \psi - \psi - \omega \gamma^0 \psi - \gamma^\mu A_\mu \psi - \varepsilon \gamma^0 |\psi|^{\frac{1}{2}} \psi = 0 \\ -4\pi \Delta A_0 = J^0 = |\psi|^2, \quad -4\pi \Delta A_k = -J^k. \end{cases}$$

Let  $\theta$  be a smooth function satisfying  $\theta(s) = 0$  for  $s \leq -1$  and  $\theta(s) = 1$  for  $s \geq 0$ . The gradient being defined by  $-\nabla I_{\omega, \varepsilon} = -|D_1|^{-1} I'_{\omega, \varepsilon}$ , let us consider the flow for positive times  $t$ ,  $\eta_{\omega, \varepsilon}^t$ , of a modified gradient :

$$(38) \quad \begin{cases} \eta^0 = \mathbf{1}_{\mathfrak{H}} \\ \partial_t \eta_{\omega, \varepsilon}^t = -(\theta(I_{\omega, \varepsilon}) \nabla I_{\omega, \varepsilon}) \circ \eta_{\omega, \varepsilon}^t. \end{cases}$$

It can be seen that for  $\varepsilon > 0$  the functional  $I_{\omega, \varepsilon}$  enjoys the properties needed for the Benci-Rabinowitz linking argument [17], except that its gradient is not of the form  $L + K$  with  $L$  linear and  $K$  compact. Due to this lack of compactness,

one cannot use Leray-Schauder's degree. One can work instead with a generalized version of the Leray-Schauder  $\mathbb{Z}_2$ -degree, due to Smale [153] to show the existence of a positive critical level of  $I_{\omega,\varepsilon}$ . This idea was introduced by Hofer-Wysocki [93] in the study of homoclinic orbits of nonconvex Hamiltonian systems, where a similar lack of compactness occurs. Hofer and Wysocki worked with the unregularized  $L^2$ -gradient. This gradient does not have a well-defined flow, but for the linking argument it is sufficient to consider certain smooth gradient lines, which are pseudo-holomorphic curves satisfying boundary conditions. Later Séré [150], studying homoclinic orbits on singular energy hypersurfaces, worked with the  $H^{1/2}$ -regularized gradient, which has a well-defined flow leading to an easier and more flexible linking argument. This approach is adapted to nonlinear Dirac in [62] and to Dirac-Maxwell and Dirac-Klein-Gordon in [61]. Consider the sets

$$\mathcal{N}^- = \{\psi = \psi_- + \lambda e_+(\mu), \psi_- \in \mathfrak{H}_1^-, \|\psi_-\|_{\mathfrak{H}} \leq \|e_+(\mu)\|_{\mathfrak{H}}, \lambda \in [0, 1]\}$$

and

$$\Sigma_+ = \{\psi \in \mathfrak{H}_1^+ / \|\psi\|_{\mathfrak{H}} = \rho\}, \quad \rho > 0.$$

Then one can prove the

**Proposition 14.** *For any  $-1 < \omega < -\mu$ ,  $\varepsilon > 0$  and  $\Sigma_+$ ,  $\mathcal{N}_-$  constructed as above, there exists a positive constant  $c_\omega$ , such that the set  $\eta_{\omega,\varepsilon}^t(\mathcal{N}_-) \cap \Sigma_+$  is non-empty, for all  $t \geq 0$ . Moreover, the number*

$$c_{\omega,\varepsilon} = \inf_{t \geq 0} I_{\omega,\varepsilon} \circ \eta_{\omega,\varepsilon}^t(\mathcal{N}_-)$$

is strictly positive, it is a critical level for  $I_{\omega,\varepsilon}$  and  $c_{\omega,\varepsilon} \rightarrow c_\omega > 0$  as  $\varepsilon \rightarrow 0$ .

Additionally, for any  $\omega, \varepsilon$  fixed, there is a sequence  $\{\varphi_{\omega,\varepsilon}^n\}_{n \geq 0}$  such that as  $n \rightarrow +\infty$ ,

$$(39) \quad \begin{cases} I_{\omega,\varepsilon}(\varphi_{\omega,\varepsilon}^n) \rightarrow c_{\omega,\varepsilon}, \\ (1 + \|\varphi_{\omega,\varepsilon}^n\|) \nabla I_{\omega,\varepsilon}(\varphi_{\omega,\varepsilon}^n) \rightarrow 0. \end{cases}$$

**Remark 15.** *In [61] and [62], an easy regularization step is missing. Indeed, Smale's degree theory requires  $C^2$ -regularity for the flow, which corresponds to  $C^3$ -regularity for the functional. In the case of the local nonlinear Dirac equation, such a regularity can be easily achieved by a small perturbation of the function  $F(\psi) + \varepsilon|\psi|^{\alpha_2-1}\psi$ . Since all the estimates will be independent of the regularization parameter, the solutions of the non-regularized problem will be obtained by a limiting argument.*

Note that the linking argument of [150], [62] and [61] has inspired later work (see [162, 157]), where an abstract linking theorem in a noncompact setting is given, valid first for  $C^2$  and then for  $C^1$ -functionals.

**Third step: Properties of the critical sequences.** The concentration-compactness theory of P.-L. Lions [118] allows us to analyze the behavior of critical sequences of  $I_{\omega,\varepsilon}$  as follows:

**Proposition 16.** *Let  $\omega \in (-1, 0)$  and  $\varepsilon \geq 0$  be fixed. Let  $(\psi_n) \subset \mathfrak{H}$  be a sequence in  $\mathfrak{H}$  such that*

$$(40) \quad 0 < \inf_n \|\psi_n\|_{\mathfrak{H}} \leq \sup_n \|\psi_n\|_{\mathfrak{H}} < +\infty$$

and  $I'_{\omega,\varepsilon}(\psi_n) \rightarrow 0$  in  $\mathfrak{H}'$  as  $n$  goes to  $+\infty$ . Then we can find a finite integer  $p \geq 1$ ,  $p$  non-zero solutions  $\varphi^1, \dots, \varphi^p$  of (37) in  $\mathfrak{H}$  and  $p$  sequences  $(x_n^i) \subset \mathbb{R}^3$ ,  $i = 1, \dots, p$  such that for  $i \neq j$ ,  $|x_n^i - x_n^j| \xrightarrow{n \rightarrow +\infty} +\infty$ , and, up to extraction of a subsequence,

$$\left\| \psi_n - \sum_{i=1}^p \varphi^i(\cdot - x_n^i) \right\|_{\mathfrak{H}} \xrightarrow{n \rightarrow +\infty} 0.$$

Obtaining estimates in  $H^{1/2}(\mathbb{R}^3)$  for the sequence  $\{\varphi_{\omega,\varepsilon}^n\}_{n \geq 0}$  of Proposition 14 is quite easy because of the coercivity introduced by the perturbation term in  $\varepsilon$ . Moreover,  $c_{\omega,\varepsilon}$  being strictly positive, the sequence  $\{\varphi_{\omega,\varepsilon}^n\}_{n \geq 0}$  is also bounded from below away from 0. So, Proposition 16 applies to prove the existence of a solution to (37) for every  $\varepsilon > 0$ . Next, we want to pass to the limit when  $\varepsilon$  goes to 0. Note that we are doing so along a sequence of functions which are exact solutions of the approximate problem (37). This part of the proof is done by first proving the equivalent of the Pohozaev identity for equation (37),  $\varepsilon \geq 0$ , and then by introducing some special topologies in the spaces  $L^q$  which are related to the decomposition of  $\mathbb{R}^3$  as the union of unit cubes. Analyzing the solutions to (37) in those topologies, we find the following

**Theorem 17.** *There is a constant  $\kappa > 0$  such that if  $-1 < \omega < 0$  and  $0 < \varepsilon \leq 1$ , there is a function  $\psi_\varepsilon \in \mathfrak{H}$  such that  $I'_{\omega,\varepsilon}(\psi_\varepsilon) = 0$  and*

$$\kappa \leq I_{\omega,\varepsilon}(\psi_\varepsilon) \leq c_{\omega,\varepsilon}.$$

**Last step: Passing to the limit  $\varepsilon \rightarrow 0$ .** Eventually, we use Proposition 16 to pass to the limit  $\varepsilon \rightarrow 0$ . When obtaining the estimates (40) for the critical sequences of  $I_{\omega,\varepsilon}$ , we observe that the lower estimate for the norm  $\|\cdot\|_{\mathfrak{H}}$  is actually independent of  $\varepsilon$ . Assume, by contradiction that the upper estimates do not hold for the sequence  $(\psi_\varepsilon)$ . Then, we consider the normalized functions

$$\tilde{\psi}_\varepsilon = \|\psi_\varepsilon\|_{\mathfrak{H}}^{-1} \psi_\varepsilon$$

and apply Proposition 16 to the sequence  $(\tilde{\psi}_\varepsilon)$ . Under the assumption that  $\|\psi_\varepsilon\|_{\mathfrak{H}} \rightarrow +\infty$ , we use all the previous estimates to infer that for  $j = 1, \dots, p$ ,

$$\sum_{j=1}^p \int_{\mathbb{R}^3} \bar{\varphi}^j \varphi^j + \omega |\varphi|^2 dx = 0, \quad Q(\varphi^j) = 0.$$

But the latter implies that for every  $j$ ,  $\bar{\varphi}^j \varphi^j = 0$  a.e. and so, from the r.h.s. identity we obtain that  $\varphi^j = 0$  a.e. for all  $j$ . This contradicts Theorem 17.  $\square$

**1.6. Nonlinear Dirac evolution problems.** The results which we have mentioned so far are concerned with the existence of stationary solutions of various nonlinear Dirac evolution equations. These particular solutions are global and do not change their shape with time. The study of the nonlinear Dirac evolution problem

$$(41) \quad \begin{cases} i\partial_t \Psi - D_1 \Psi + G(\Psi) = 0 \\ \Psi(0) = \Psi_0 \end{cases}$$

is also interesting in itself and, even if this is not the aim of the present paper, let us mention some references.

For the case of local nonlinearities as the ones considered in this section, several works have proved well-posedness for small initial data in well chosen Sobolev spaces. For nonlinearities containing only powers of  $\Psi$  of order  $p \geq 4$ , Reed proved in [139] the global well-posedness for small initial data in  $H^s$ ,  $s > 3$ . A decay estimate at infinity was also obtained in this paper. Later, Dias and Figueira [43] improved this result to include powers of order  $p = 3$  and for  $s > 2$ . Najman [127] took the necessary regularity of the initial data down to  $H^2$ . In [60] Escobedo and Vega proved an “optimal result” which states that for the physically relevant nonlinearities of order  $p \geq 3$  of the type

$$(42) \quad G(\Psi) := \lambda \left\{ (\bar{\Psi}\Psi)^{\frac{p-1}{2}} \beta \Psi + b(\Psi, \gamma^5 \Psi)^{\frac{p-1}{2}} \gamma^5 \Psi \right\},$$

there is local well-posedness of the evolution equation in  $H^s$ , for  $s > \frac{3}{2} - \frac{1}{p-1}$ , when  $p$  is an odd integer, while  $s$  has to be in the interval  $(\frac{3}{2} - \frac{1}{p-1}, \frac{p-1}{2})$  otherwise. Moreover, if  $p > 3$ , then the problem is globally well-posed for small initial data in  $H^{s(p)}$ , with  $s(p) = \frac{3}{2} - \frac{1}{p-1}$ . For a more recent result, see for instance a paper of Machihara, Nakanishi and Ozawa [121], in which the existence of small global solutions is proved in  $H^s$  for  $s > 1$ , and the nonrelativistic limit is also considered.

An interesting question to ask is about the (linear or nonlinear) stability properties of the stationary solutions with respect to the flow generated by the evolution equation. At present this seems to be a widely open problem (see [138] and [155] for a discussion). Recently, Boussaid [21] has obtained the first stability results, for small stationary solutions of nonlinear Dirac equations with exterior potential.

Concerning the Cauchy problem for the Maxwell-Dirac equations, the first result about the local existence and uniqueness of solutions was obtained by L. Gross in [82]. Later developments were made by Chadam [34] and Chadam and Glassey [35] in  $1 + 1$  and  $2 + 1$  space-time dimensions and in  $3 + 1$  dimensions when the magnetic field is 0. Choquet-Bruhat studied in [38] the case of spinor fields of zero mass and Maxwell-Dirac equations in the Minkowski space were studied by Flato, Simon and Taffin in [73]. In [76], Georgiev obtained a class of initial values for which the Maxwell-Dirac equations have a global solution. This was performed by using a technique introduced by Klainerman (see [100, 101, 102]) to obtain  $L^\infty$  a priori estimates via the Lorentz invariance of the equations and a generalized version of the energy inequalities. The same method was used by Bachelot [8] to obtain a similar result for the Klein-Gordon-Dirac equation. Finally, more recent efforts have been directed to proving existence of solutions for the time-dependent Klein-Gordon-Dirac and Maxwell-Dirac equations in the energy space, namely  $C(-T, T; H^{1/2} \times H^1)$ . The existence and uniqueness of solutions to the Maxwell-Dirac system in the energy space has been proved by Masmoudi and Nakanishi in [122, 123], improving Bournaveas' result in [22], where the space considered was  $C(-T, T; H^{1/2+\varepsilon} \times H^{1+\varepsilon})$ .

Note that as mentioned above, the stationary states of the form (36) are particular solutions of the Maxwell-Dirac equations. Physically they correspond to bound states of the electron.

## 2. LINEAR DIRAC EQUATIONS FOR AN ELECTRON IN AN EXTERNAL FIELD

When looking for stationary states describing the dynamics of an electron moving in an external field generated by an electrostatic potential  $V$ , one is led to study the

eigenvalues and eigenfunctions of the operator  $D_c + V$ . If the electron has to enjoy some stability, the eigenvalues should also be away from the essential spectrum. In the case of not very strong potentials  $V$ , the essential spectrum of  $D_c + V$  is the same as that of  $D_c$ , that is, the set  $(-\infty, -c^2] \cup [c^2, +\infty)$ . So the eigenvalues which are of interest to us are those lying in the gap of the essential spectrum, i.e. in the interval  $(-c^2, c^2)$ . More precisely, in general a state describing an electron is always assumed to correspond to a positive eigenvalue. It is therefore important to be able to determine whether there are positive eigenvalues or not, and what is the behaviour of the ‘first’ eigenvalue when  $V$  varies (whether it crosses 0 or dives into the lower negative essential spectrum for instance). Note that one expects that for a reasonable potential there are no eigenvalues embedded in the essential spectrum. Very general conditions on  $V$  which ensure nonexistence of embedded eigenvalues have been given by [18, 19]. Note finally that in this section  $c$  is kept variable.

Formally, the eigenvalues of the operator  $D_c + V$  are critical values of the Rayleigh quotient

$$(43) \quad Q_V(\psi) := \frac{((D_c + V)\psi, \psi)}{(\psi, \psi)}$$

in the domain of  $D_c + V$ . Of course, one cannot use a minimizing argument to find such critical points since, due to the negative continuous spectrum of the free Dirac operator,  $Q_V$  is not bounded-below. Many works have been devoted to finding non-minimization variational problems yielding the eigenvalues of  $D_c + V$  in the interval  $(-c^2, c^2)$ . Another important issue is to avoid the appearance of spurious states (some eigenvalues of the finite dimensional problem may not approach the eigenvalues of the Dirac operator  $D_c + V$ ) as it has been the case in many proposed algorithms (see for instance [55]). W. Kutzelnigg has written two excellent reviews [104, 105] on this subject, where many relevant references can be found. The main techniques which have been developed so far and used in practice can be divided in three groups:

- (1) Use of effective Hamiltonians whose point spectra are close to the spectrum of the Dirac operator in the limit  $c \rightarrow +\infty$ . For instance, one can cut at a finite level some infinite asymptotic formal expansion in negative powers of  $c$ . To this category of works belong for instance [58, 59, 110, 111, 112, 105].
- (2) Use of a Galerkin approximation technique to approach the eigenvalues, and this without falling into the negative continuum  $(-\infty, -c^2)$ . This is equivalent to projecting the equation onto a well-chosen finite dimensional space. This procedure has been well explained for instance in [56, 57, 104].
- (3) Replacement of the problematic minimization of  $Q_V$  by another one. For instance, it was proposed to minimize the Rayleigh quotient for the squared Hamiltonian  $(D_c + V)^2$  (see, e.g. [167, 16]) or later on, to maximize the Rayleigh quotients for the “inverse Hamiltonian”  $\frac{D_c + V}{|D_c + V|^2}$  (see [92]).

Before we go further, let us recall some useful inequalities which are usually used to control the external field  $V$  and show that  $D_c + V$  is essentially self-adjoint. We recall that  $\mathfrak{H} = H^{1/2}(\mathbb{R}^3, \mathbb{C}^4)$  and that  $\mathfrak{H}_{\pm}^0$  are the positive and negative spectral subspaces of  $D_1$ .

**Proposition 18** (Hardy-like inequalities). *The Coulomb potential  $W(x) = \frac{1}{|x|}$  satisfies the following Hardy-type inequalities:*

$$(44) \quad W \leq \frac{\pi}{2} \sqrt{-\Delta} \leq \frac{\pi}{2c} |D_c|,$$

$$(45) \quad \forall \psi \in \mathfrak{H}_+^0 \cup \mathfrak{H}_-^0, \quad (\psi, W(x)\psi)_{L^2} \leq \frac{1}{2} \left( \frac{\pi}{2} + \frac{2}{\pi} \right) (\psi, |D_1|\psi)_{L^2},$$

$$(46) \quad W^2 \leq -4\Delta \leq 4|D_1|^2.$$

The inequalities of Proposition 18 are classical (see, e.g. [91, 98] for (44)), except for (45). The proof of the latter is based on a method of Evans, Perry and Siedentop [67] and is contained in the recent papers [29, 160, 161].

**2.1. A variational characterization of the eigenvalues of  $D_c + V$ .** Formally, the eigenvalues of  $D_c + V$  lying in the gap of the essential spectrum should be described by some kind of *min-max* argument. This was mentioned in several papers dealing with numerical computations of Dirac eigenvalues before it was formally addressed in different contexts in a series of papers [63, 81, 80, 50, 51].

For the sake of clarity, we are going to present only a particular version of those min-max arguments allowing to characterize eigenvalues of the operator  $D_c + V$  for appropriate potentials  $V$ . This method derives from a proposition made by Talman [158] and Datta, Deviah [42] and based on the decomposition of any spinor  $\psi = \begin{pmatrix} \varphi \\ \chi \end{pmatrix}$  as the sum of its upper and its lower components:

$$(47) \quad \psi = \begin{pmatrix} \varphi \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \chi \end{pmatrix}.$$

This proposal consisted in saying that the first eigenvalue of  $D_c + V$  could be obtained by solving the min-max problem

$$(48) \quad \min_{\varphi \neq 0} \max_{\chi} \frac{((D_c + V)\psi, \psi)}{(\psi, \psi)}.$$

The first rigorous result on this min-max principle was obtained by Griesemer and Siedentop [81], who proved that (48) yields indeed the first positive eigenvalue of  $D_c + V$  for potentials  $V$  which are in  $L^\infty$  and not too large. In [51], Dolbeault, Esteban and Séré proved that if  $V$  satisfies the assumptions

$$(49) \quad V(x) \xrightarrow{|x| \rightarrow +\infty} 0,$$

$$(50) \quad -\frac{\nu}{|x|} - K_1 \leq V \leq K_2 = \sup_{x \in \mathbb{R}^3} V(x),$$

$$(51) \quad K_1, K_2 \geq 0, \quad K_1 + K_2 - c^2 < \sqrt{c^4 - \nu^2 c^2}$$

with  $\nu \in (0, c)$ ,  $K_1, K_2 \in \mathbb{R}$ , then the first eigenvalue  $\lambda_1(V)$  of  $D_c + V$  in the interval  $(-c^2, c^2)$  is given by the formula

$$(52) \quad \lambda_1(V) = \inf_{\varphi \neq 0} \sup_{\chi} \frac{(\psi, (D_c + V)\psi)}{(\psi, \psi)}, \quad \psi = \begin{pmatrix} \varphi \\ \chi \end{pmatrix}.$$

Actually, under the conditions (49)-(50)-(51), it can be seen that  $D_c + V$  has an infinite sequence of eigenvalues  $\{\lambda_k(V)\}_{k \geq 1}$  converging to 1, and it was proved in [51] that each of them can be obtained by a min-max procedure:

**Theorem 19.** (Min-max characterization of the eigenvalues of  $D_c + V$  [51]). *Let  $V$  be a scalar potential satisfying (49)-(50)-(51). Then, for all  $k \geq 1$ , the  $k$ -th eigenvalue  $\lambda_k(V)$  of the operator  $D_c + V$  is given by the following min-max formula*

$$(53) \quad \lambda_k(V) = \inf_{\substack{Y \text{ subspace of } C_0^\infty(\mathbb{R}^3, \mathbb{C}^2) \\ \dim Y = k}} \sup_{\varphi \in Y \setminus \{0\}} \lambda^T(V, \varphi),$$

where

$$(54) \quad \lambda^T(V, \varphi) := \sup_{\substack{\psi = \begin{pmatrix} \varphi \\ \chi \end{pmatrix} \\ \chi \in C_0^\infty(\mathbb{R}^3, \mathbb{C}^2)}} \frac{((D_c + V)\psi, \psi)}{(\psi, \psi)}$$

is the unique number in  $(K_2 - c^2, +\infty)$  such that

$$(55) \quad \lambda^T(V, \varphi) \int_{\mathbb{R}^3} |\varphi|^2 dx = \int_{\mathbb{R}^3} \left( \frac{c^2 |(\sigma \cdot \nabla)\varphi|^2}{c^2 - V + \lambda^T(V, \varphi)} + (c^2 + V)|\varphi|^2 \right) dx.$$

The above result is optimal for Coulomb potentials for which all the cases  $\nu \in (0, c)$  are included. But note that assumptions (50)-(51) can be replaced by weaker ones allowing in particular to treat potentials which have a finite number of isolated singularities, even of different signs. We describe some of these extensions at the end of this subsection.

Theorem 19 is a useful tool from a practical point of view in the sense that the first eigenvalue (case  $k = 1$ ) of  $D_c + V$  can be obtained by a *minimization procedure* over the (bounded-below) nonlinear functional  $\varphi \mapsto \lambda^T(V, \varphi)$ . Higher eigenvalues are obtained by the usual Rayleigh-Ritz minimax principle on the same nonlinear functional. As we shall see below, this has important consequences from a numerical point of view.

Theorem 19 is a direct consequence of an abstract theorem proved by Dolbeault, Esteban and Séré [51], providing variational characterizations for the eigenvalues of self-adjoint operators in the gaps of their essential spectrum.

**Theorem 20.** (Min-max principle for eigenvalues of operators with gaps [51]) *Let  $\mathcal{H}$  be a Hilbert space and  $A : D(A) \subset \mathcal{H} \rightarrow \mathcal{H}$  a self-adjoint operator. We denote by  $\mathcal{F}(A)$  the form-domain of  $A$ . Let  $\mathcal{H}_+$ ,  $\mathcal{H}_-$  be two orthogonal Hilbert subspaces of  $\mathcal{H}$  such that  $\mathcal{H} = \mathcal{H}_+ \oplus \mathcal{H}_-$  and let  $\Lambda_\pm$  be the projectors associated with  $\mathcal{H}_\pm$ . We assume the existence of a subspace of  $D(A)$ ,  $F$ , which is dense in  $D(A)$  and such that :*

(i)  $F_+ = \Lambda_+ F$  and  $F_- = \Lambda_- F$  are two subspaces of  $\mathcal{F}(A)$ .

(ii)  $a_- = \sup_{x_- \in F_- \setminus \{0\}} \frac{(x_-, Ax_-)}{\|x_-\|_{\mathcal{H}}^2} < +\infty$ .

Moreover we define the sequence of min-max levels

$$(56) \quad c_k(A) = \inf_{\substack{V \text{ subspace of } F_+ \\ \dim V = k}} \sup_{x \in (V \oplus F_-) \setminus \{0\}} \frac{(x, Ax)}{\|x\|_{\mathcal{H}}^2}, \quad k \geq 1,$$

and assume that

(iii)  $c_1(A) > a_-$ .

Then

$$\forall k \geq 1, \quad c_k(A) = \mu_k,$$

where, if  $b = \inf(\sigma_{\text{ess}}(A) \cap (a_-, +\infty)) \in (a_-, +\infty]$ ,  $\mu_k$  denotes the  $k^{\text{th}}$  eigenvalue of  $A$  (counted with multiplicity) in the interval  $(a_-, b)$  if it exists, or  $\mu_k = b$  if there is no  $k^{\text{th}}$  eigenvalue.

As a consequence,  $b = \lim_{k \rightarrow \infty} c_k(A) = \sup_k c_k(A) > a_-$ .

An important feature of this min-max principle is that the min-max levels do not depend on the splitting  $\mathcal{H} = \mathcal{H}_+ \oplus \mathcal{H}_-$  provided assumptions (i), (ii) and (iii) hold true. In practice, one can find many different splittings satisfying these assumptions and choose the most convenient one for a given application.

In order to treat families of operators without checking the assumptions of the above theorem for every case, there is a “continuous” version of Theorem 20 in [51] which we shall present now.

Let us start with a self-adjoint operator  $A_0 : D(A_0) \subset \mathcal{H} \rightarrow \mathcal{H}$ . and denote by  $\mathcal{F}(A_0)$  the form-domain of  $A_0$ . Now, for  $\nu$  in an interval  $[0, \bar{\nu})$  we define  $A_\nu = A_0 + \nu W$  where  $W$  is a bounded operator. The operator  $A_\nu$  is self-adjoint with  $\mathcal{D}(A_\nu) = \mathcal{D}(A_0)$ ,  $\mathcal{F}(A_\nu) = \mathcal{F}(A_0)$ . Let  $\mathcal{H} = \mathcal{H}_+ \oplus \mathcal{H}_-$  be an orthogonal splitting of  $\mathcal{H}$ , and  $P_+$ ,  $P_-$  the associated projectors, as in Section 1. We assume the existence of a subspace of  $D(A_0)$ ,  $F$ , dense in  $D(A_0)$  and such that:

(j)  $F_+ = P_+F$  and  $F_- = P_-F$  are two subspaces of  $\mathcal{F}(A_0)$ ;

(jj) there is  $a_- \in \mathbb{R}$  such that for all  $\nu \in (0, \bar{\nu})$ ,

$$a_\nu := \sup_{x_- \in F_- \setminus \{0\}} \frac{(x_-, A_\nu x_-)}{\|x_-\|_{\mathcal{H}}^2} \leq a_-.$$

For  $\nu \in (0, \bar{\nu})$ , let  $b_\nu := \inf(\sigma_{\text{ess}}(A_\nu) \cap (a_\nu, +\infty))$ , and for  $k \geq 1$ , let  $\mu_{k,\nu}$  be the  $k$ -th eigenvalue of  $A_\nu$  in the interval  $(a_\nu, b_\nu)$ , counted with multiplicity, if it exists. If it does not exist, we simply let  $\mu_{k,\nu} := b_\nu$ . Our next assumption is

(jjj) there is  $a_+ > a_-$  such that for all  $\nu \in (0, \bar{\nu})$ ,  $\mu_{1,\nu} \geq a_+$ .

Finally, we define the min-max levels

$$(57) \quad c_{k,\nu} := \inf_{\substack{V \text{ subspace of } F_+ \\ \dim V = k}} \sup_{x \in (V \oplus F_-) \setminus \{0\}} \frac{(x, A_\nu x)}{\|x\|_{\mathcal{H}}^2}, \quad k \geq 1,$$

and assume that

(jv)  $c_{1,0} > a_-$ .

Then, we have the

**Theorem 21.** ([51]) *Under conditions (j) to (jv),  $A_\nu$  satisfies the assumptions (i) to (iii) of Theorem 20 for all  $\nu \in [0, \bar{\nu})$ , and  $c_{k,\nu} = \mu_{k,\nu} \geq a_+$ , for all  $k \geq 1$ .*

Theorems 20 and 21 are very good tools in the study of the point-spectrum of Dirac operators  $D_c + \nu V$ , where  $V$  is a potential which has singularities not stronger than  $c/|x - x_0|$  ( $0 < c < 1$ ). Of course, Theorem 21 cannot be directly applied to the case of unbounded potentials, but this can actually be done by first truncating the potential and then passing to the limit in the truncation parameter, as we did in the proof of Theorem 19.

Theorem 21 is an easy consequence of the proof of Theorem 20. In contrast, the proof of Theorem 20 is more involved. We sketch now its main steps.

**Sketch of the proof of Theorem 20.** For  $E > a$  and  $x_+ \in F_+$ , let us define

$$\begin{aligned} \varphi_{E,x_+} : \quad F_- &\rightarrow \mathbb{R} \\ y_- &\mapsto \varphi_{E,x_+}(y_-) = \left( (x_+ + y_-), A(x_+ + y_-) \right) - E \|x_+ + y_-\|_{\mathcal{H}}^2. \end{aligned}$$

From assumptions (i) – (ii),  $N(y_-) = \sqrt{(a+1)\|y_-\|_{\mathcal{H}}^2 - (y_-, Ay_-)}$  is a norm on  $F_-$ . Let  $\overline{F_-}^N$  be the completion of  $F_-$  for this norm. Since  $\|\cdot\|_{\mathcal{H}} \leq N$  on  $F_-$ , we have  $\overline{F_-}^N \subset \mathcal{H}_-$ . For all  $x_+ \in F_+$ , there is an  $x \in F$  such that  $\Lambda_+x = x_+$ . If we consider the new variable  $z_- = y_- - \Lambda_-x$ , we can define

$$\psi_{E,x}(z_-) := \varphi_{E,\Lambda_+x}(z_- + \Lambda_-x) = (A(x + z_-), x + z_-) - E(x + z_-, x + z_-).$$

Since  $F$  is a subspace of  $D(A)$ ,  $\psi_{E,x}$  (hence  $\varphi_{E,x_+}$ ) is well-defined and continuous for  $N$ , uniformly on bounded sets. So,  $\varphi_{E,x_+}$  has a unique continuous extension  $\overline{\varphi}_{E,x_+}$  on  $\overline{F_-}^N$ , which is continuous for the extended norm  $\overline{N}$ . It is well-known (see e.g. [140]) that there is a unique self-adjoint operator  $B : D(B) \subset \mathcal{H}_- \rightarrow \mathcal{H}_-$  such that  $D(B)$  is a subspace of  $\overline{F_-}^N$ , and

$$\forall x_- \in D(B), \quad \overline{N}(x_-)^2 = (a+1)\|x_-\|_{\mathcal{H}}^2 + (x_-, Bx_-).$$

Now,  $\overline{\varphi}_{E,x_+}$  is of class  $C^2$  on  $\overline{F_-}^N$  and

$$\begin{aligned} D^2\overline{\varphi}_{E,x_+}(x_-) \cdot (y_-, y_-) &= -2(y_-, By_-) - 2E\|y_-\|_{\mathcal{H}}^2 \\ (58) \quad &\leq -2\min(1, E)\overline{N}(y_-)^2. \end{aligned}$$

So  $\overline{\varphi}_{E,x_+}$  has a unique maximum, at the point  $y_- = L_E(x_+)$ . The Euler-Lagrange equations associated to this maximization problem are :

$$(59) \quad \Lambda_-Ax_+ - (B + E)y_- = 0.$$

The above arguments allow us, for any  $E > a$ , to define a map

$$\begin{aligned} Q_E : \quad F_+ &\rightarrow \mathbb{R} \\ (60) \quad x_+ &\mapsto Q_E(x_+) = \sup_{x_- \in F_-} \varphi_{E,x_+}(x_-) = \overline{\varphi}_{E,x_+}(L_E x_+) \\ &= (x_+, (A - E)x_+) + (\Lambda_-Ax_+, (B + E)^{-1}\Lambda_-Ax_+). \end{aligned}$$

It is easy to see that  $Q_E$  is a quadratic form with domain  $F_+ \subset \mathcal{H}_+$  and it is monotone nonincreasing in  $E > a$ .

We may also, for  $E > a$  given, define the norm  $n_E(x_+) = \|x_+ + L_E x_+\|_{\mathcal{H}}$ . We consider the completion  $X$  of  $F_+$  for the norm  $n_E$  and denote by  $\overline{n}_E$  the extended norm. Then, we define another norm on  $F_+$  by

$$\mathcal{N}_E(x_+) = \sqrt{Q_E(x_+) + (K_E + 1)(n_E(x_+))^2}$$

with  $K_E = \max(0, \frac{E^2(E-\lambda_1)}{\lambda_1^2})$  and consider the completion  $G$  of  $F_+$  for the norm  $\mathcal{N}_E$ . Finally, we use the monotonicity of the map  $E \mapsto Q_E$  and classical tools of spectral theory to prove that the  $k$ -th eigenvalue of  $A$  in the interval  $(0, \inf(\sigma_{\text{ess}}(A) \cap (a, +\infty)))$  is the unique  $E_k > a$  such that

$$(61) \quad \inf_{\substack{V \text{ subspace of } G \\ \dim V = k}} \sup_{x_+ \in V \setminus \{0\}} \frac{\overline{Q}_{E_k}(x_+)}{(\overline{n}_{E_k}(x_+))^2} = 0.$$

Note that since the above min-maxs correspond to the eigenvalues of the operator  $T^E$  associated to the quadratic form  $Q_E$ , (61) is actually equivalent to

$$\lambda_k(T^{E_k}) = 0.$$

and the fact that this inequality defines a unique  $E_k$  relies on the monotonicity of  $Q_E$  w.r.t.  $E$ .  $\square$

In the application of Theorem 20 to prove Theorems 19 and 21, various decompositions  $\mathcal{H} = \tilde{\mathcal{H}}^+ \oplus \tilde{\mathcal{H}}^-$  could be considered. One that gives excellent results is defined by

$$(62) \quad \psi = \begin{pmatrix} \varphi \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \chi \end{pmatrix}.$$

This decomposition yields optimal results about the point spectrum for some potentials  $V$ . There are cases for which this is not anymore true. For instance, this happens when the potential  $V$  has “large” positive and negative parts, case which is not dealt with in the previous results.

Recently, Dolbeault, Esteban and Séré [53] have considered the case where a potential can give rise to two different types of eigenvalues, not only those appearing in Theorems 19 and 21. More precisely, if  $V$  satisfies (49), assume that it is continuous everywhere except at two finite sets of isolated points,  $\{x_i^+\}$ ,  $\{x_j^-\}$ ,  $i = 1, \dots, I$ ,  $j = 1, \dots, J$ , where

$$(63) \quad \begin{aligned} \lim_{x \rightarrow x_i^+} V(x) &= +\infty, & \lim_{x \rightarrow x_i^+} V(x) |x - x_i^+| &\leq \nu_i, \\ \lim_{x \rightarrow x_j^-} V(x) &= -\infty, & \lim_{x \rightarrow x_j^-} V(x) |x - x_j^-| &\geq -\nu_j, \end{aligned}$$

with  $\nu_i, \nu_j \in (0, c)$  for all  $i, j$ . Under the above assumptions, as above,  $D_c + V$  has a distinguished self-adjoint extension  $A$  with domain  $\mathcal{D}(A)$  such that

$$H^1(\mathbb{R}^3, \mathbb{C}^4) \subset \mathcal{D}(A) \subset H^{1/2}(\mathbb{R}^3, \mathbb{C}^4).$$

The essential spectrum of  $A$  is the same as that of  $D_c$  :

$$\sigma_{\text{ess}}(A) = (-\infty, -c^2] \cup [c^2, +\infty),$$

see [159, 146, 128, 103]. Finally,  $V$  maps  $\mathcal{D}(A)$  into its dual, since (49) and (63) imply that for all  $\varphi \in H^{1/2}(\mathbb{R}^3)$ ,  $V\varphi \in H^{-1/2}(\mathbb{R}^3)$ . The decomposition of  $\mathfrak{H}$  considered here is related to the positive/negative spectral decomposition of the free Dirac operator  $D_c$  :

$$\mathfrak{H} = \mathfrak{H}_+^0 \oplus \mathfrak{H}_-^0,$$

with  $\mathfrak{H}_\pm^0 = P_\pm^0 \mathfrak{H}$ , where  $P_\pm^0$  are the positive/negative spectral projectors of the free Dirac operator  $D_c$ .

As above, we assume the existence of a core  $F$  (i.e. a subspace of  $\mathcal{D}(A)$  which is dense for the norm  $\|\cdot\|_{\mathcal{D}(A)}$ ), such that :

- (i)  $F_+ = P_+^0 F$  and  $F_- = P_-^0 F$  are two subspaces of  $\mathcal{D}(A)$ ;
- (ii<sup>-</sup>)  $a^- := \sup_{x_- \in F_- \setminus \{0\}} \frac{(x_-, Ax_-)}{\|x_-\|_{\mathfrak{H}}^2} < +\infty$ ;
- (ii<sup>+</sup>)  $a^+ := \inf_{x_+ \in F_+ \setminus \{0\}} \frac{(x_+, Ax_+)}{\|x_+\|_{\mathfrak{H}}^2} > -\infty$ .

We consider the two sequences of min-max and max-min levels  $\{\lambda_k^+\}_{k \geq 1}$  and  $\{\lambda_k^-\}_{k \geq 1}$  defined by

$$(64) \quad \lambda_k^+ := \inf_{\substack{V \text{ subspace of } F_+ \\ \dim V = k}} \sup_{x \in (V \oplus F_-) \setminus \{0\}} \frac{(x, Ax)}{\|x\|_5^2},$$

$$(65) \quad \lambda_k^- := \sup_{\substack{V \text{ subspace of } F_- \\ \dim V = k}} \inf_{x \in (V \oplus F_+) \setminus \{0\}} \frac{(x, Ax)}{\|x\|_5^2}.$$

**Theorem 22.** *Take a positive integer  $k_0$  and any  $k \geq k_0$  and let  $A$  be the self-adjoint extension of  $D_c + V$  defined above, where  $V$  is a scalar potential satisfying (49) and (63).*

*If  $a^- < \lambda_{k_0}^+ < c^2$ , then  $\lambda_k^+$  is either an eigenvalue of  $D_c + V$  in the interval  $(a^-, c^2)$ , or  $\lambda_k^+ = c^2$ . If additionally  $V \geq 0$ , then  $a^- = c^2$  and  $\lambda_k^+ = c^2$ .*

*If  $-c^2 < \lambda_{k_0}^- < a^+$ , then  $\lambda_k^-$  is either an eigenvalue of  $D_c + V$  in the interval  $(-c^2, a^+)$  or  $\lambda_k^- = -c^2$ . If additionally  $V \leq 0$ , then  $a^+ = -c^2$  and  $\lambda_k^- = -c^2$ .*

The sequences  $\{\lambda_k^+\}_{k \geq 1}$  and  $\{\lambda_k^-\}_{k \geq 1}$  are respectively nondecreasing and nonincreasing. As a consequence of their definitions we have:

$$(66) \quad \text{for all } k \geq 1, \quad \lambda_k^+ \geq \max\{a^-, a^+\} \text{ and } \lambda_k^- \leq \min\{a^-, a^+\},$$

and if  $a^- \geq a^+$ , we do not state anything about the possible eigenvalues in the interval  $[a^+, a^-]$ . Note that, as it is showed in [53], there are operators for which *all* or *almost all* the eigenvalues lie in the interval  $[a^+, a^-]$  and thus, they are not given by the variational procedures defining the  $\lambda_k^\pm$ 's.

Finally, let us remark that if we apply Theorem 22 to deal with a family of operators  $D_c + \tau V$ , with  $V$  satisfying (49)-(63), then we see that the eigenvalues  $\lambda_k^+$ 's and  $\lambda_k^-$ 's are of a "different" kind, since

$$\lim_{\tau \rightarrow 0^+} \lambda_k^{\tau, \pm} = \pm c^2, \quad \text{for all } k \geq 1.$$

In physical words, we could say that the  $\lambda_k^+$ 's correspond to electronic states, and the  $\lambda_k^-$ 's to positronic ones.

**2.2. Numerical method based on the min-max formula.** Let us now come back to the case  $k = 1$  of Theorem 19. Note that from (52) and (54) we see that (under the right assumptions on  $V$ ) the first eigenvalue  $\lambda_1(V)$  of  $D_c + V$  in the gap  $(-c^2, c^2)$  can be seen as the solution of a minimization problem, that is,

$$(67) \quad \lambda_1(V) = \min_{\substack{\varphi \in C^\infty(\mathbb{R}^3, \mathbb{C}^2) \\ \varphi \neq 0}} \lambda^T(V, \varphi),$$

where  $\varphi \mapsto \lambda^T(V, \varphi)$  is a nonlinear functional implicitly defined by

$$(68) \quad \lambda^T(V, \varphi) \int_{\mathbb{R}^3} |\varphi|^2 dx = \int_{\mathbb{R}^3} \left( \frac{c^2 |(\sigma \cdot \nabla)\varphi|^2}{c^2 - V + \lambda^T(V, \varphi)} + (c^2 + V)|\varphi|^2 \right) dx.$$

The idea of characterizing the first eigenvalue in a gap of the essential spectrum as the solution of a minimization problem is not completely new. It has for instance already been used by Bayliss and Peel [16] in another context. It is also close to the Feshbach method and to some techniques used in Pencil Theories.

The fact that one can reduce the computation of  $\lambda_1(V)$  to that of a minimization problem (67)-(68) has an important practical consequence: these problems (67)-(68)

can now be easily discretized to construct an algorithm, allowing us to approximate  $\lambda_1(V)$  in an efficient manner. Indeed, the functional  $\lambda^T(V, \cdot)$  to be minimized is *bounded from below* in the whole space  $H^{1/2}(\mathbb{R}^3, \mathbb{C}^2)$ . This is a huge advantage compared to other methods in which the total Rayleigh quotient  $Q_v$  is minimized on particular finite dimensional subspaces of  $H^{1/2}(\mathbb{R}^3, \mathbb{C}^4)$ : the latter are prone to variational collapse (entering into the negative continuum) and can even furnish spurious solutions (see, e.g. [36]).

The discretization method based on (67) is completely free of all these complications and satisfactory numerical tests for atomic and molecular models have been presented in [54, 52]. Notice that molecular simulations are more complicated to carry on because one cannot use the rotational symmetry like in the atomic case. In contrast to the one-dimensional radially symmetric problem, the discretization has to be made in  $\mathbb{R}^2$  when axial symmetry is present, or in  $\mathbb{R}^3$  in the general case. Below we describe the algorithm that was used in [54, 52] to find eigenvalues of  $D_c + V$  by the minimization method presented above.

Consider the following approximation procedure for  $\lambda_k(V)$ ,  $k \geq 1$ . Take any complete countable basis set  $\mathcal{B}$  in the space of admissible 2-spinors  $X$  and let  $\mathcal{B}_n$  be an  $n$ -dimensional subset of  $\mathcal{B}$  generating the space  $X_n$ . We assume that  $\mathcal{B}_n$  is monotone increasing in the sense that if  $n < n'$ , then  $\mathcal{B}_n$  is contained in  $\mathcal{B}_{n'}$ . Denote by  $\varphi_1, \varphi_2, \dots, \varphi_n$  the elements of  $\mathcal{B}_n$ . For all  $1 \leq i, j \leq n$ , we define the  $n \times n$  matrix  $A_n(\lambda)$  whose entries are

$$(69) \quad A_n^{i,j}(\lambda) = \int_{\mathbb{R}^3} \left( \frac{(c(\sigma \cdot \nabla) \varphi_i, c(\sigma \cdot \nabla) \varphi_j)}{\lambda + c^2 - V} + (V + c^2 - \lambda) (\varphi_i, \varphi_j) \right) dx.$$

The matrix  $A_n(\lambda)$  is selfadjoint and has therefore  $n$  real eigenvalues. For  $1 \leq k \leq n$ , we compute  $\lambda_{k,n}$  as the solution of the equation

$$(70) \quad \mu_{k,n}(\lambda) = 0,$$

where  $\mu_{k,n}(\lambda)$  is the  $k$ -th eigenvalue of  $A_n(\lambda)$ . Note that the uniqueness of such a  $\lambda$  comes from the monotonicity of the r.h.s. of equation (69) with respect to  $\lambda$ . Moreover, since for a fixed  $\lambda$

$$(71) \quad \mu_{k,n}(\lambda) \searrow \mu_k(\lambda) \quad \text{as } n \rightarrow +\infty,$$

we also have

$$(72) \quad \lambda_{k,n} \searrow \lambda_k(V) \quad \text{as } n \rightarrow +\infty.$$

The elements of the basis set used in [54] were Hermite functions. In [52] more efficient numerical results have been obtained by means of  $B$ -spline functions. The interest of using well-localized basis set functions is the sparseness and the nice structure of the corresponding discretized matrix  $A_n(\lambda)$ . If the degree of the basis of  $B$ -splines increases, the number of filled diagonals will also increase. So, a good balance has to be found between the smoothness of elements of the approximating basis set and the speed of the corresponding numerical computations. In [52] the simple choice of considering second order spline functions on a variable length grid was made. In the atomic case, when 1-dimensional  $B$ -splines are used, very quick and accurate results can be obtained. In [52] numerical tests were provided for some axially symmetric diatomic molecules.

In [109] we can find an algorithm which has some analogy with the algorithm described above.

**2.3. New Hardy-like inequalities.** Another byproduct of the minimization characterization of the first eigenvalue of  $D_c + V$  given in (67) and of (55) is the following: for all  $\varphi \in D(D_c + V)$ ,

$$(73) \quad \int_{\mathbb{R}^3} \left( \frac{c^2 |(\sigma \cdot \nabla) \varphi|^2}{c^2 - V + \lambda_1(V)} + (c^2 + V - \lambda_1(V)) |\varphi|^2 \right) dx \geq 0.$$

In the particular case  $V = -\nu/|x|$ ,  $\nu \in (0, c)$ , (73) means that for all  $\varphi \in H^1(\mathbb{R}^3, \mathbb{C}^2)$ ,

$$(74) \quad \int_{\mathbb{R}^3} \left( \frac{c^2 |(\sigma \cdot \nabla) \varphi|^2}{c^2 + \nu/|x| + \sqrt{c^4 - \nu^2 c^2}} + (c^2 - \sqrt{c^4 - \nu^2 c^2}) |\varphi|^2 \right) dx \geq \nu \int_{\mathbb{R}^3} \frac{|\varphi|^2}{|x|} dx.$$

By scaling, one finds that for all  $\varphi \in H^1(\mathbb{R}^3, \mathbb{C}^2)$ , and for all  $\nu \in (0, 1)$ ,

$$(75) \quad \int_{\mathbb{R}^3} \left( \frac{|(\sigma \cdot \nabla) \varphi|^2}{1 + \nu/|x| + \sqrt{1 - \nu^2}} + (1 - \sqrt{1 - \nu^2}) |\varphi|^2 \right) dx \geq \nu \int_{\mathbb{R}^3} \frac{|\varphi|^2}{|x|} dx,$$

and, passing to the limit when  $\nu$  tends to 1, we get:

$$(76) \quad \int_{\mathbb{R}^3} \left( \frac{|(\sigma \cdot \nabla) \varphi|^2}{1 + 1/|x|} + |\varphi|^2 \right) dx \geq \int_{\mathbb{R}^3} \frac{|\varphi|^2}{|x|} dx.$$

This inequality is a Hardy-like inequality related to the Dirac operator. It is not invariant under dilation, which corresponds to the fact that the Dirac operator  $D_c$  is not homogeneous. But by another scaling argument, (76) yields, as a corollary, an inequality which is invariant by dilation,

$$(77) \quad \int_{\mathbb{R}^3} |x| |(\sigma \cdot \nabla) \varphi|^2 dx \geq \int_{\mathbb{R}^3} \frac{|\varphi|^2}{|x|} dx,$$

which is actually equivalent to the ‘‘classical’’ Hardy inequality

$$(78) \quad \int_{\mathbb{R}^3} |\nabla \varphi|^2 dx \geq \frac{1}{4} \int_{\mathbb{R}^3} \frac{|\varphi|^2}{|x|^2} dx.$$

For a 4-dimensional version of (77), see [170, 99].

Finally, note that in [49] the Hardy-like inequality (76) (and slightly more general ones also) has been proved by analytical means, without using any previous knowledge about the Coulomb-Dirac operator’s spectrum.

**2.4. The nonrelativistic limit.** Let us now indicate how we can relate the eigenvalues of the Dirac operator to those of the Schrödinger operator. This relation is established by taking the limit  $c \rightarrow +\infty$ , so by passing to the nonrelativistic limit.

A  $\psi$  with values in  $\mathbb{C}^4$  satisfies the eigenvalue equation

$$(79) \quad (D_c + V) \psi = \lambda \psi$$

if and only if, writing  $\psi = \begin{pmatrix} \varphi \\ \chi \end{pmatrix}$  with  $\varphi, \chi$  taking values in  $\mathbb{C}^2$ ,

$$(80) \quad \begin{cases} R_c \chi = (\lambda - c^2 - V) \varphi, \\ R_c \varphi = (\lambda + c^2 - V) \chi, \end{cases}$$

with

$$R_c = -i c (\vec{\sigma} \cdot \vec{\nabla}) = \sum_{j=1}^3 -i c \sigma_j \frac{\partial}{\partial x_j}.$$

Recall that  $\sigma_j$ ,  $j = 1, 2, 3$ , are the Pauli matrices. As long as  $\lambda + c^2 - V \neq 0$ , the system (80) can be written as

$$(81) \quad R_c \begin{pmatrix} R_c \varphi \\ g_\mu \end{pmatrix} + V \varphi = \mu \varphi, \quad \chi = \frac{R_c \varphi}{g_\mu}$$

where  $g_\mu = \mu + 2c^2 - V$  and  $\mu = \lambda - c^2$ .

Assume now that  $\psi_c = \begin{pmatrix} \varphi_c \\ \chi_c \end{pmatrix}$  is an eigenfunction of the operator  $(D_c + V)$  associated with the eigenvalue  $\lambda_c$  which satisfies

$$\liminf_{c \rightarrow \infty} (\lambda_c - c^2) > -\infty, \quad \limsup_{c \rightarrow \infty} (\lambda_c - c^2) < 0.$$

The system (81) can be written as

$$(82) \quad \frac{-c^2 \Delta \varphi_c}{\mu_c + 2c^2 - V} + \frac{c^2 (\vec{\sigma} \cdot \vec{\nabla}) \varphi_c \cdot (\vec{\sigma} \cdot \vec{\nabla}) V}{(\mu_c + 2c^2 - V)^2} + V \varphi_c = \mu_c \varphi_c, \quad \chi_c = \frac{-i c (\vec{\sigma} \cdot \vec{\nabla}) \varphi_c}{\mu_c + 2c^2 - V},$$

with  $\mu_c = \lambda_c - c^2$ . It is then easy to prove (see [65]) that for  $c$  large, the functions  $\varphi_c$ , which in principle are only in  $H^{1/2}(\mathbb{R}^3, \mathbb{C}^2)$ , actually belong to the space  $H^1(\mathbb{R}^3, \mathbb{C}^2)$  and are uniformly bounded for the  $H^1(\mathbb{R}^3, \mathbb{C}^2)$  norm. Moreover, after taking subsequences, we can find  $\bar{\varphi} \in H^1(\mathbb{R}^3, \mathbb{C}^2)$  and  $\bar{\mu} < 0$  such that

$$\lim_{c \rightarrow +\infty} \|\varphi_c - \bar{\varphi}\|_{H^1(\mathbb{R}^3, \mathbb{C}^2)} = 0, \quad \lim_{c \rightarrow +\infty} \mu_c = \bar{\mu},$$

and

$$-\frac{\Delta \bar{\varphi}}{2} + V \bar{\varphi} = \bar{\mu} \bar{\varphi}.$$

Note that  $\chi_c$ , the lower component of the eigenfunction  $\psi_c$ , converges to 0 in the nonrelativistic limit.

It can be proved that for all the potentials  $V$  considered in the theorems of this section, all the eigenvalues in the gap  $(-c^2, c^2)$  satisfy the above conditions, and converge, when shifted by the quantity  $-c^2$ , to the associated eigenvalues of the Schrödinger operator  $-\frac{\Delta}{2}$  perturbed by the same potential  $V$ .

**2.5. Introduction of a constant external magnetic field.** The previous results were devoted to the case of a scalar electrostatic field  $V$ . The Dirac operator for a hydrogenic atom interacting with a constant magnetic field in the  $x_3$ -direction is given by

$$(83) \quad D^B := \boldsymbol{\alpha} \cdot \left[ -i\nabla + \frac{B}{2}(-x_2, x_1, 0) \right] + \beta - \frac{\nu}{|x|},$$

where  $\nu = Z\alpha > 0$ ,  $Z$  being the nuclear charge number (we fix the speed of light  $c = 1$  in this subsection) and  $B$  is a constant.

The magnetic Dirac operator without the Coulomb potential  $\nu/|x|$  has essential spectrum  $(-\infty, -1] \cup [1, \infty)$  and no eigenvalue in the gap  $(-1, 1)$  for any  $B \in \mathbb{R}$ . The operator  $D^B$  has the same essential spectrum and possibly some eigenvalues in the gap. The ground state energy  $\lambda_1(\nu, B)$  is the smallest among these. As the field gets large enough, one expects that the ground state energy of the Dirac operator decreases and eventually penetrates the lower continuum. The implication of this for a second quantized model is that electron–positron pair creation comes into the picture [129, 132]. The intuition comes from the Pauli equation, where the magnetic field tends to lower the energy because of the spin. It is therefore reasonable to define the *critical field strength*  $B(\nu)$  as the supremum of the  $B$ 's for

which  $\lambda_1(\nu, b)$  is in the gap  $(-1, 1)$  for all  $b < B$ . As a function of  $\nu$ ,  $\lambda_1(\nu, B)$  is non-increasing, and as a result the function  $B(\nu)$  is also non-increasing. Estimates on this critical field as a function of the nuclear charge  $\nu$  can be found in [48]. They have been obtained by adapting to this case the variational arguments of Theorem 19. One of the first results in this paper states that for all  $\nu \in (0, 1)$ ,

$$(84) \quad \frac{0.75}{\nu^2} \leq B(\nu) \leq \min \left( \frac{18\pi\nu^2}{[3\nu^2 - 2]_+^2}, e^{C/\nu^2} \right).$$

As a corollary we see that as  $\nu \rightarrow 1$  the critical field  $B(\nu)$  stays strictly positive. This is somewhat remarkable, since in the case without magnetic field the ground state energy  $\lambda_1(\nu, 0)$ , as a function of  $\nu$  tends to 0 as  $\nu \rightarrow 1$  but *with an infinite slope*. Thus, one might expect very large variations of the eigenvalue at  $\nu = 1$  as the magnetic field is turned on, in particular one might naively expect that the ground state energy leaves the gap for small fields  $B$ . This is not the case.

Next, again by using the min-max characterization of  $\lambda_1(\nu, B)$ , it is shown in [48] that for  $\nu > 0$  small enough, and  $B$  not too large  $\lambda_1(\nu, B)$  is asymptotically close to the ground state energy of the Coulomb-Dirac magnetic operator in the lowest relativistic Landau level  $c_0(\nu, B)$ . This constant is proved to be given by

$$(85) \quad c_0(\nu, B) = \inf_{f \in C_0^\infty(\mathbb{R}, \mathbb{C}) \setminus \{0\}} \lambda_0^B(f),$$

where  $\lambda_0^B(f)$  is implicitly defined by

$$(86) \quad \lambda_0^B(f) \int_{-\infty}^{+\infty} |f(z)|^2 dz = \int_{-\infty}^{+\infty} \left( \frac{|f'(z)|^2}{1 + \lambda_0^B(f) + \nu a_0^B(z)} + (1 - \nu a_0^B(z)) |f(z)|^2 \right) dz,$$

and

$$a_0^B(z) = B \int_0^{+\infty} \frac{s e^{-\frac{Bs^2}{2}}}{\sqrt{s^2 + z^2}} ds.$$

In [48] it is proved that for  $B$  not too small and  $\nu$  small enough,

$$(87) \quad c_0(\nu + \nu^{3/2}, B) \leq \lambda_1(\nu, B) \leq c_0(\nu, B).$$

and that since for  $\nu$  small,  $\nu^{3/2} \ll \nu$ ,

$$c_0(\nu + \nu^{3/2}, B) \sim c_0(\nu, B) \sim \lambda_1(\nu, B) \quad \text{as } \nu \rightarrow 0.$$

The one dimensional  $c_0(\nu, B)$  problem, while not trivial, is simpler to calculate than the  $\lambda_1(\nu, B)$  problem. As a result, in the limit as  $\nu \rightarrow 0$ , this new theory yields the first term in the asymptotics of the logarithm of the critical field. In particular we have the following result,

$$\lim_{\nu \rightarrow 0} \nu \log(B(\nu)) = \pi$$

### 3. THE DIRAC-FOCK EQUATIONS FOR ATOMS AND MOLECULES

In the previous two sections we described some results concerning the solutions of nonlinear or linear Dirac equations in  $\mathbb{R}^3$ , which represent the state of one electron only (or possibly many *non-interacting* electrons). We now want to present the *Dirac-Fock (DF) model* which allows to describe the state of *interacting* electrons, like for instance  $N$  electrons in a molecule. The DF model is very often used in quantum chemistry computations and usually gives very good numbers when the correlation between the electrons is negligible. It is the relativistic counterpart of

the better known non-relativistic Hartree-Fock equations, which can indeed be seen as the non-relativistic limit ( $c \rightarrow \infty$ ) of the Dirac-Fock model as explained below. For this reason, we start by recalling briefly the Hartree-Fock model.

**3.1. The (non-relativistic) Hartree-Fock equations.** The Hartree-Fock equations are easily derived from the linear Schrödinger model in which one considers the following operator

$$(88) \quad H := \sum_{i=1}^N \left( -\frac{\Delta_{x_i}}{2} + V(x_i) \right) + \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|}$$

whose associated quadratic form describes the energy of  $N$  interacting electrons in the potential field  $V$ . Most often,  $V$  is the Coulomb electrostatic potential created by a positive distribution of charge  $\nu \geq 0$ , of total charge  $Z$ :

$$V = -\nu * \frac{1}{|\cdot|}, \quad \int_{\mathbb{R}^3} \nu = Z.$$

In the case of  $M$  pointwise nuclei of charges  $z_1, \dots, z_M$  and located at  $\bar{x}_1, \dots, \bar{x}_M$ , one takes  $\nu = \sum_{m=1}^M z_m \delta_{\bar{x}_m}$  and  $Z = \sum_{m=1}^M z_m$ . But extended nuclei can also be considered in which case  $\nu$  is assumed to be a smooth  $L^1$  non-negative function.

Due to the Pauli principle, the operator  $H$  acts on  $\bigwedge_{i=1}^N L^2(\mathbb{R}^3 \times \{\pm\}, \mathbb{C})$ , that is to say the space of  $L^2$  functions  $\Psi(x_1, \sigma_1, \dots, x_N, \sigma_N)$  which are antisymmetric with respect to the permutations of the  $(x_i, \sigma_i)$ 's. When  $Z > N - 1$ , it is known [171, 172] that the spectrum of  $H$  has the form  $\sigma(H) = \{\lambda_i\} \cup [\Sigma, \infty)$  where  $\{\lambda_i\}$  is an increasing sequence of eigenvalues with finite multiplicity converging to the bottom of the essential spectrum  $\Sigma$ . We notice that the condition  $Z > N - 1$  plays a special role even for the linear theory based on the operator (88), as one knows [165, 96] that only *finitely many* eigenvalues exist below  $\Sigma$  when  $N \geq Z + 1$ , and that there is *no* eigenvalue below  $\Sigma$  when  $N \gg Z$  [144, 151, 152, 115]. In the following, we shall always assume that  $Z > N - 1$ .

In the Hartree-Fock approximation, one computes an approximation of the first eigenvalue  $\lambda_1$  of  $H$  by restricting the quadratic form  $\Psi \mapsto \langle \Psi, H\Psi \rangle$  to the class of the functions  $\Psi$  which are a simple (Slater) determinant:

$$(89) \quad \Psi = \varphi_1 \wedge \dots \wedge \varphi_N$$

where  $(\varphi_1, \dots, \varphi_N)$  is an orthonormal system of  $L^2(\mathbb{R}^3 \times \{\pm\}, \mathbb{C}) = L^2(\mathbb{R}^3, \mathbb{C}^2)$ ,  $\int_{\mathbb{R}^3} (\varphi_i, \varphi_j)_{\mathbb{C}^2} = \delta_{ij}$ . Denoting  $\varphi_i = \begin{pmatrix} \varphi_i^+ \\ \varphi_i^- \end{pmatrix}$ , (89) means more precisely

$$\Psi(x_1, \sigma_1, \dots, x_N, \sigma_N) = \frac{1}{\sqrt{N!}} \det(\varphi_i^{\sigma_j}(x_j)).$$

Since the set of all the  $\Psi$ 's having the form (89) is not a vector subspace of  $\bigwedge_{i=1}^N L^2(\mathbb{R}^3 \times \{\pm\}, \mathbb{C})$ , one then obtains an energy functional which is nonlinear in terms of  $\varphi_1, \dots, \varphi_N$ . The associated Euler-Lagrange equations form a system of  $N$  coupled nonlinear PDEs:

$$(90) \quad H_{\Phi} \varphi_k = \lambda_k \varphi_k, \quad k = 1, \dots, N$$

where  $H_{\Phi}$  is the so-called *mean-field operator* seen by each of the  $N$  electrons

$$(91) \quad H_{\Phi} = -\frac{\Delta}{2} + (\rho_{\Phi} - \nu) * \frac{1}{|\cdot|} - \frac{\gamma_{\Phi}(x, y)}{|x - y|},$$

with  $\rho_\Phi$  being the (scalar) *electronic density* and  $\gamma_\Phi$  the so-called *density matrix* of the  $N$  electrons (this is a  $2 \times 2$  matrix for any  $(x, y) \in \mathbb{R}^3 \times \mathbb{R}^3$ ):

$$(92) \quad \rho_\Phi := \sum_{i=1}^N |\varphi_i|^2 \quad \text{and} \quad \gamma_\Phi(x, y) := \sum_{i=1}^N \varphi_i(x) \otimes \varphi_i(y)^*.$$

We notice that (91) means

$$(H_\Phi \psi)(x) = -\frac{\Delta \psi(x)}{2} + \left( (\rho_\Phi - \nu) * \frac{1}{|\cdot|} \right) (x) \psi(x) - \int \frac{\gamma_\Phi(x, y) \psi(y)}{|x - y|} dy$$

for any  $\psi \in H^2(\mathbb{R}^3, \mathbb{C}^2)$ . The existence of solutions to (90) when  $\int_{\mathbb{R}^3} \nu = Z > N - 1$  has been proved first by Lieb and Simon [117] by a minimization method, and then by Lions [119] by general min-max arguments. See also [108] for a recent survey.

**3.2. Existence of solutions to the Dirac-Fock equations.** The relativistic Dirac-Fock equations were first introduced by Swirles in [156]. They take the same form as the Hartree-Fock equations (90), with  $-\Delta/2$  replaced by the Dirac operator  $D_c$ . They are of course posed for functions taking values in  $\mathbb{C}^4$  instead of  $\mathbb{C}^2$ . Note however that when  $-\Delta/2$  is replaced by  $D_c$  in the formula of the  $N$ -body Hamiltonian (88), one obtains an operator whose spectrum is the whole line  $\mathbb{R}$  as soon as  $N \geq 2$ . To our knowledge, it is not known whether there exist or not eigenvalues which are embedded in the essential spectrum. In any case, the relativistic  $N$ -body problem is not well-posed. This somehow restricts the physical interpretation of the Dirac-Fock model, compared to its non-relativistic counterpart. We refer to the next section in which a better model deduced from Quantum Electrodynamics is presented. Despite this issue, the Dirac-Fock equations have been widely used in computational atomic physics and quantum chemistry to study atoms and molecules involving heavy nuclei, and they seem to provide very good results when the correlation between the electrons is negligible.

In the case of  $N$  electrons, the Dirac-Fock equations read

$$(93) \quad D_{c, \Phi} \varphi_k = \varepsilon_k \varphi_k, \quad k = 1, \dots, N,$$

where  $\Phi = (\varphi_1, \dots, \varphi_N)$  satisfies  $\int_{\mathbb{R}^3} (\varphi_i(x), \varphi_j(x)) dx = \delta_{ij}$ , i.e.

$$(94) \quad \text{Gram}(\Phi) = \mathbf{1},$$

and

$$(95) \quad D_{c, \Phi} = D_c + (\rho_\Phi - \nu) * \frac{1}{|x|} - \frac{\gamma_\Phi(x, y)}{|x - y|},$$

$$(96) \quad \gamma_\Phi(x, y) = \sum_{\ell=1}^N \varphi_\ell(x) \otimes \varphi_\ell(y)^*, \quad \rho_\Phi(x) = \text{tr}_{\mathbb{C}^4}(\gamma_\Phi(x, x)) = \sum_{\ell=1}^N |\varphi_\ell(x)|^2.$$

Notice that  $\gamma_\Phi(x, y)$  is a  $4 \times 4$  complex matrix, and that the operator whose kernel is  $\gamma_\Phi(x, y)$ , is nothing but the orthogonal projector onto the space spanned by  $\varphi_1, \dots, \varphi_N$ . We also denote it by  $\gamma_\Phi$ .

Indeed, like for the Hartree-Fock case, equations (93) are the Euler-Lagrange equations of the Dirac-Fock functional

$$(97) \quad \mathcal{E}_{\text{DF}}^{\nu,c}(\Phi) = \sum_{\ell=1}^N \left( \varphi_\ell, D_c \varphi_\ell \right)_{L^2} - \sum_{\ell=1}^N \left( \varphi_\ell, \left( \nu * \frac{1}{|x|} \right) \varphi_\ell \right)_{L^2} \\ + \frac{1}{2} \iint_{\mathbb{R}^3 \times \mathbb{R}^3} \frac{\rho_\Phi(x) \rho_\Phi(y) - \text{tr} \left( \gamma_\Phi(x, y) \gamma_\Phi(y, x) \right)}{|x - y|} dx dy$$

on the manifold

$$\mathcal{M} := \left\{ \Phi = (\varphi_1, \dots, \varphi_N) \in (H^{1/2}(\mathbb{R}^3, \mathbb{C}^4))^N, \quad \text{Gram}(\Phi) = \mathbf{1} \right\}.$$

It will be important to notice that the functional  $\mathcal{E}_{\text{DF}}^{\nu,c}$  only depends on the projector  $\gamma_\Phi$  defined in (96):

$$(98) \quad \mathcal{E}_{\text{DF}}^{\nu,c}(\Phi) = \text{“tr}((D_c + V)\gamma_\Phi)\text{”} + \frac{1}{2} \iint_{\mathbb{R}^3 \times \mathbb{R}^3} \frac{\rho_\Phi(x) \rho_\Phi(y) - |\gamma_\Phi(x, y)|^2}{|x - y|} dx dy,$$

where by “tr(( $D_c + V$ ) $\gamma_\Phi$ )” we denote  $\sum_{i=1}^N ((D_c + V)\varphi_i, \varphi_i)$ . Note that this expression is really a trace if the  $\varphi_i$ 's are in  $H^1(\mathbb{R}^3, \mathbb{C}^4)$ .

As a matter of fact, the Euler-Lagrange equations of  $\mathcal{E}_{\text{DF}}^{\nu,c}$  on  $\mathcal{M}$  only depend on the space spanned by  $(\varphi_1, \dots, \varphi_N)$ . This explains why, up to a rotation of the  $\varphi_i$ 's, one can always assume that the Lagrange multiplier matrix associated with the constraint (94) is diagonal, with diagonal elements  $(\varepsilon_1, \dots, \varepsilon_N)$  appearing in (93).

Finding solutions of (93) is then reduced to finding critical points of the functional  $\mathcal{E}_{\text{DF}}^{\nu,c}$  on the manifold  $\mathcal{M}$ . Once again, the unboundedness (from above and below) of the spectrum of the free Dirac operator makes the functional  $\mathcal{E}_{\text{DF}}^{\nu,c}$  totally indefinite. This together with the a priori lack of compactness of the problem posed in  $\mathbb{R}^3$  and the fact that we have to work on a manifold  $\mathcal{M}$  and not in the whole functional space, makes the variational problem difficult. A minimization procedure is once again impossible and another method has to be found. In [64], Esteban and Séré defined a penalized variational problem (see below for details) which can be solved by first maximizing on some part of the spinor functions  $\varphi_i$  and then defining a more standard min-max argument for the remaining functional, together with Morse index considerations.

The theorem proved in [64] and improved later by Paturel in [131] states the following:

**Theorem 23.** (Existence of solutions to the Dirac-Fock equations [64, 131]) *With the above notations, assume that  $N$  and  $Z = \int_{\mathbb{R}^3} \nu$  are two positive integers satisfying  $\max(Z, N) < \frac{2c}{\pi/2+2/\pi}$  and  $N - 1 < Z$ . Then, there exists an infinite sequence  $(\Phi^{c,j})_{j \geq 0}$  of critical points of the Dirac-Fock functional  $\mathcal{E}_{\text{DF}}^{\nu,c}$  on  $\mathcal{M}$ . The functions  $\varphi_1^{c,j}, \dots, \varphi_N^{c,j}$  satisfy the normalization constraints (94) and they are strong solutions, in  $H^{1/2}(\mathbb{R}^3, \mathbb{C}^4) \cap \bigcap_{1 \leq q < 3/2} W^{1,q}(\mathbb{R}^3, \mathbb{C}^4)$ , of the Dirac-Fock equations*

$$(99) \quad D_{c, \Phi^{c,j}} \varphi_k^{c,j} = \varepsilon_k^{c,j} \varphi_k^{c,j}, \quad 1 \leq k \leq N,$$

$$(100) \quad 0 < \varepsilon_1^{c,j} \leq \dots \leq \varepsilon_N^{c,j} < c^2 .$$

Moreover,

$$(101) \quad 0 < \mathcal{E}_{\text{DF}}^{\nu,c}(\Phi^{c,j}) < Nc^2 ,$$

$$(102) \quad \lim_{j \rightarrow \infty} \mathcal{E}_{\text{DF}}^{\nu,c}(\Phi^{c,j}) = Nc^2 .$$

**Remark 24.** *In our units, taking into account the physical value of the speed of light  $c$ , the above conditions become*

$$N \leq Z \leq 124.$$

The proof of the above theorem is done by defining a sequence of min-max principles providing critical points of increasing Morse index. We notice that the solution  $\Phi^{c,0}$  obtained by Theorem 23 when  $j = 0$  will play an important role, since for  $c$  large it will be actually interpreted as an “electronic ground state” (see below).

The condition  $Z, N < 2c/(2\pi + \pi/2)$  is not that unnatural, since already in the linear case such a condition was necessary to use Hardy-like inequalities ensuring the existence of a gap of the spectrum of  $D_c + V$  around 0, see (45). Notice however that in [64] the following additional technical assumption was used:

$$(103) \quad 3N - 1 < \frac{2c}{\pi/2 + 2/\pi}.$$

This assumption was removed by Paturel shortly afterwards in [131]. He did so by studying a finite dimensional reduction of the problem and then passing to the limit, after having obtained the necessary bounds. This proof was done in the spirit of the Conley-Zehnder proof of the Arnold’s conjecture [41]. The sketch of the proof that we give below is that of [64] because the variational arguments are easier to explain in that case.

**Sketch of the proof of Theorem 23 when (103) holds.** The first (and smallest) difficulty here is that  $\nu * \frac{1}{|x|}$  is not a compact perturbation of  $D_c$  when  $\nu = Z\delta_0$ . This creates some technical problems. They are easily solved, replacing the Coulomb potential  $\frac{1}{|x|}$  by a regularized potential  $V_\delta$ . The modified energy functional is denoted now by  $\mathcal{E}_{\text{DF}}^{\nu,c,\eta}$ . At the end of the proof, we shall be able to pass to the limit  $\eta \rightarrow 0$ .

The second difficulty is that the Morse index estimates can only give upper bounds on the multipliers  $\varepsilon_k^{c,j}$  in (99). But we also want to ensure that  $\varepsilon_k^{c,j} > 0$ , since these multipliers are interpreted as the energies of the different electrons. To overcome this problem, we replace the constraint  $\text{Gram}(\Phi) = \mathbf{1}$  by a penalization term  $\pi_p(\Phi)$ , subtracted from the energy functional:

$$(104) \quad \pi_p(\Phi) = \text{tr} \left[ \left( \text{Gram} \Phi \right)^p \left( \mathbf{1} - \text{Gram} \Phi \right)^{-1} \right] .$$

In this way we obtain a new functional  $\mathcal{F}_{c,\eta,p} = \mathcal{E}_{\text{DF}}^{\nu,c,\eta} - \pi_p$ , defined now on the set of  $\Phi$ ’s satisfying

$$\mathbf{0} < \text{Gram} \Phi < \mathbf{1} .$$

Since in the basis in which

$$\text{Gram} \Phi = \text{Diag}(\sigma_1, \dots, \sigma_N), \quad 0 < \sigma_1 \leq \dots \leq \sigma_N < c^2 ,$$

the matrix  $(\text{Gram } \Phi)^p (\mathbf{1} - \text{Gram } \Phi)^{-1}$  is also diagonal and equals

$$\text{Diag}(f_p(\sigma_1), \dots, f_p(\sigma_N)), \quad \text{with } f_p(x) := \frac{x^p}{1-x},$$

the corresponding Euler-Lagrange equations are then

$$D_{c, \Phi^{c,j}} \varphi_k^{c,j} = f'_p(\sigma_k^{c,j}) \varphi_k^{c,j}.$$

The numbers  $\varepsilon_k^{c,j} := f'_p(\sigma_k^{c,j})$  are now explicit functions of  $\varphi_k^{c,j}$ . Thus,  $f_p$  being an increasing function, we automatically get  $\varepsilon_k^{c,j} > 0$ .

The third difficulty with DF, is that all critical points have an infinite Morse index. This kind of problem is often encountered in the theory of Hamiltonian systems and in certain elliptic PDEs. One way of dealing with it is to use a concavity property of the functional, to get rid of the “negative directions”, see e.g. [2, 27, 28, 31]. This method was used in [64]. Doing so, we get a reduced functional  $I_{c,\eta,p}$ . A min-max argument gives us Palais-Smale sequences  $(\Phi_n^{c,\eta,p})_{n \geq 1}$  for  $I_{c,\eta,p}$  with a Morse index “at most  $j$ ” (up to an error which converges to 0 as  $n \rightarrow \infty$ ), thanks to [68, 77]. Moreover, adapting the arguments of [119], we prove that the corresponding  $\varepsilon_k^{c,j,\eta,p,n}$  are bounded away from  $c^2$ . Finally, the assumptions made on  $Z, N$  guarantee that the  $\varepsilon_k^{c,j,\eta,p,n}$  are also bounded away from 0, uniformly on  $\eta, p$  and  $n$ . Then we pass to the limit  $(\eta, p, n) \rightarrow (0, \infty, \infty)$ , and get the desired solutions of DF, with  $0 < \varepsilon_k^{c,j} < c^2$ .

The fact that we recover at the limit  $p \rightarrow +\infty$  the constraint  $\text{Gram } \Phi = \mathbf{1}$  is a consequence of the *a priori* estimates

$$0 < \varepsilon < \varepsilon_k^{c,j,\eta,p} = f'_p(\sigma_k^{c,j,\eta,p}),$$

with  $\varepsilon$  independent of  $\eta, p$ . The properties of the function  $f_p$  and the above inequality imply that as  $p \rightarrow +\infty$ , one necessarily has  $\sigma_k^{c,j,\eta,p} \rightarrow 1$ , which of course is equivalent to saying that in the limit there is no loss of charge:  $\int_{\mathbb{R}^3} |\varphi_k^{c,j}|^2 = 1$ .

The method that we have just described can be generalized. In [26] an abstract version is provided, with applications to nonlinear periodic Schrödinger models arising in the physics of crystalline matter.

The concavity argument of [64] works only if (103) holds. In his theorem, Paturel [131] got rid of this assumption by making a finite dimensional reduction first. This allowed him to deal with finite Morse indices again.  $\square$

### 3.3. Nonrelativistic limit and definition of the Dirac-Fock “ground state”.

As in the case of linear Dirac equations of Section 3, it is interesting here to see what is the nonrelativistic limit of the Dirac-Fock equations. We shall recover in the limit the Hartree-Fock equations (90) presented above, for the two-dimensional upper component of the  $\varphi_i$ 's (recall that  $\varphi_1, \dots, \varphi_N$  are  $\mathbb{C}^4$ -valued functions), the lower component converging to zero. This was proved rigorously in [65]. This result has been of importance to better understand the variational structure of the Dirac-Fock problem and in particular to obtain a good definition of an electronic ground-state energy, which is *a priori* not clear because of the unboundedness of the Dirac-Fock energy.

**Theorem 25.** (Non-relativistic limit of the Dirac-Fock equations [65]) *Let be  $N < Z+1$ . Consider a sequence of numbers  $c_n \rightarrow +\infty$  and a sequence  $(\Phi^n)_n$  of solutions*

of (93), i.e.  $\Phi^n = (\varphi_1^n, \dots, \varphi_N^n)$ , each  $\varphi_k^n$  being in  $H^{1/2}(\mathbb{R}^3, \mathbb{C}^4)$ , with  $\text{Gram } \Phi^n = \mathbb{1}$  and  $D_{c_n, \Phi^n} \varphi_k^n = \varepsilon_k^n \varphi_k^n$ . Assume that the multipliers  $\varepsilon_k^n$ ,  $k = 1, \dots, N$ , satisfy

$$0 < \varepsilon < \varepsilon_1^n \leq \dots \leq \varepsilon_N^n < c^2 - \varepsilon', \quad \text{with } \varepsilon, \varepsilon' > 0 \text{ independent of } n.$$

Then for  $n$  large enough, each  $\varphi_k^n$  is in  $H^1(\mathbb{R}^3, \mathbb{C}^4)$ , and there exists a solution of the Hartree-Fock equations (90),  $\bar{\Phi} = (\bar{\varphi}_1, \dots, \bar{\varphi}_N) \in H^1(\mathbb{R}^3, \mathbb{C}^2)^N$ , with negative multipliers,  $\bar{\lambda}_1, \dots, \bar{\lambda}_N$ , such that, after extraction of a subsequence,

$$(105) \quad \lim_{n \rightarrow \infty} (\varepsilon_k^n - (c_n)^2) = \bar{\lambda}_k,$$

$$(106) \quad \varphi_k^n = \begin{pmatrix} \bar{\varphi}_k^n \\ \chi_k^n \end{pmatrix} \xrightarrow{n \rightarrow +\infty} \begin{pmatrix} \bar{\varphi}_k \\ 0 \end{pmatrix} \quad \text{in } H^1(\mathbb{R}^3, \mathbb{C}^2) \times H^1(\mathbb{R}^3, \mathbb{C}^2),$$

$$(107) \quad \left\| \chi_k^n + \frac{i}{2c_n} (\sigma \cdot \nabla) \bar{\varphi}_k^n \right\|_{L^2(\mathbb{R}^3, \mathbb{C}^2)} = O(1/(c_n)^3),$$

for all  $k = 1, \dots, N$ , and

$$(108) \quad \mathcal{E}_{\text{DF}}^{\nu, c_n}(\Phi^n) - Nc_n^2 \xrightarrow{n \rightarrow +\infty} \mathcal{E}_{\text{HF}}(\bar{\Phi}).$$

The Hartree-Fock energy  $\mathcal{E}_{\text{HF}}$  appearing in (108) is the same as (97), but with  $D_c$  replaced by  $-\Delta/2$  and  $\mathbb{C}^4$  by  $\mathbb{C}^2$ . It can be proved [65] that the critical points constructed in Theorem 23 all satisfy the assumptions of Theorem 25 for any fixed  $j$ . Therefore, all the  $\Phi^{c, j}$  converge as  $c \rightarrow \infty$  to a state whose upper component is a solution of the Hartree-Fock equations, and whose lower component vanishes. This result can even be made more precise in the case of one of the “first” solutions (i.e. corresponding to  $j = 0$ ): the critical point  $\Phi^{c, 0}$  does not converge to *any* solution of the HF equations, but actually to a Hartree-Fock *ground state*, as stated in the following

**Theorem 26.** (Non-relativistic limit of the Dirac-Fock “ground state” [65]) *Assume that  $N < Z + 1$  and  $\nu$  are fixed, and that  $c_n \rightarrow \infty$ . Then the critical point  $\Phi^{c_n, 0}$  constructed in Theorem 23 for  $j = 0$  satisfies*

$$\lim_{c_n \rightarrow \infty} \{ \mathcal{E}_{\text{DF}}^{\nu, c}(\Phi^{c_n, 0}) - Nc_n^2 \} = \min_{\substack{\bar{\Phi} \in H^1(\mathbb{R}^3, \mathbb{C}^2)^N \\ \text{Gram } \bar{\Phi} = \mathbb{1}}} \mathcal{E}_{\text{HF}}(\bar{\Phi}).$$

Up to a subsequence,  $(\Phi^{c_n, 0})$  converges as  $c_n \rightarrow \infty$  to  $\begin{pmatrix} \bar{\Phi}_0 \\ 0 \end{pmatrix}$  where  $\bar{\Phi}_0$  is a minimizer of  $\mathcal{E}_{\text{HF}}$ .

Furthermore, for  $c_n$  large enough, the  $\varepsilon_k^{c_n, 0}$  are the first positive eigenvalues of  $D_{c_n, \Phi^{c_n, 0}}$ :

$$(109) \quad \gamma_{\Phi^{c_n, 0}} = \chi_{[0, \varepsilon_N^{c_n, 0}]}(D_{c_n, \Phi^{c_n, 0}})$$

We notice that (109) means that the last level  $\varepsilon_N^{c_n, 0}$  is necessarily totally filled. In other words, similarly to the Hartree-Fock case [7], “there are no unfilled shell in the Dirac-Fock theory for  $c \gg 1$ ”.

Although the Dirac-Fock functional  $\mathcal{E}_{\text{DF}}^{\nu, c}$  is not bounded-below, Theorem 26 allows to interpret the first min-max solution  $\Phi^{c, 0}$  (any of them, since there is no uniqueness) as an electronic ground state, since it converges to a Hartree-Fock ground state in the nonrelativistic limit. Actually, more has been proved in [65]:  $\Phi^{c, 0}$  indeed minimizes the Dirac-Fock functional among all the  $\Phi = (\varphi_1, \dots, \varphi_N)$

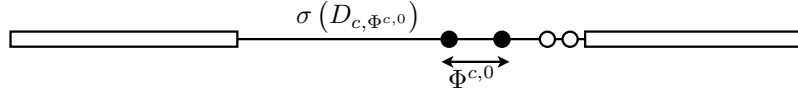


FIGURE 1. For  $c \gg 1$ , the Dirac-Fock ‘ground state’  $\Phi^{c,0}$  contains the eigenfunctions associated with the  $N$  first positive eigenvalues of the mean-field operator  $D_{c, \Phi^{c,0}}$ .

such that each  $\varphi_i$  belongs to the positive spectral subspace of the mean-field operator  $D_{c, \Phi}$ :

**Theorem 27.** (Variational interpretation of the Dirac-Fock “ground state” [65])  
*Assume that  $N < Z + 1$  and  $\nu$  are fixed, and that  $c$  is sufficiently large. Then  $\Phi^{c,0}$  is a solution of the following minimization problem:*

$$(110) \quad \inf \{ \mathcal{E}_{\text{DF}}^{\nu, c}(\Phi) \mid \Phi = (\varphi_1, \dots, \varphi_N), \text{Gram } \Phi = \mathbf{1}, \chi_{(-\infty, 0)}(D_{c, \Phi}) \Phi = 0 \}$$

where  $\chi_{(-\infty, 0)}(D_{c, \Phi})$  denotes the negative spectral projector of the operator  $D_{c, \Phi}$ , and

$$\chi_{(-\infty, 0)}(D_{c, \Phi}) \Phi = 0$$

means

$$\chi_{(-\infty, 0)}(D_{c, \Phi}) \varphi_k = 0,$$

for any  $k = 1, \dots, N$ .

The interpretation of the theorem is the following: although the Dirac-Fock energy is unbounded from below, the critical points  $\Phi^{c,0}$  are, for  $c$  large enough, the minimizers of  $\mathcal{E}_{\text{DF}}^{\nu, c}$  on the set of functions satisfying the *nonlinear* condition  $\chi_{(-\infty, 0)}(D_{c, \Phi}) \Phi = 0$ . Calling these functions *electronic*, we obtain that for  $c$  large,  $\Phi^{c,0}$  is really an *electronic ground state*. In particular, it is also a minimizer of  $\mathcal{E}_{\text{DF}}^{\nu, c}$  among all the solutions of Dirac-Fock equations with positive Lagrange multipliers.

**Remark 28.** *Recently, explicit bounds on  $c$  under which Theorem 27 is valid have been provided by Huber and Siedentop [94].*

#### 4. THE MEAN-FIELD APPROXIMATION IN QUANTUM ELECTRODYNAMICS

In this last section, we want to present some progress that has been made recently concerning models from Quantum Electrodynamics (QED) in which, instead of trying to ‘avoid’ the negative spectrum of  $D_c$ , the latter is reinterpreted as Dirac’s vacuum and completely ‘incorporated’ into the model. The somehow surprising consequence will be that, contrary to the previous sections, the energy functional will be *bounded from below*. The price to pay is that one has to deal with *infinitely many particles* (many of them being ‘virtual’) instead of the finite number of electrons as previously.

As mentioned in the introduction, Dirac interpreted the negative spectrum of his operator  $D_c$  as follows [45, 46, 47]. Since there exists no electron of negative kinetic energy, one has to find a way of avoiding the negative spectrum. This is done by assuming that the negative spectrum energies are all occupied by virtual electrons, one in each energy state, and that this (virtual) distribution of charge is not felt by the real particles on account of its uniformity. Then, the real electrons can in general only have positive energies due to the Pauli principle which prevents them to be in the same state as a virtual particle.

Mathematically, Dirac's postulate is interpreted as follows: the states of the real free electrons necessarily belong to the positive spectral subspace  $\mathfrak{H}_{+,c}^0 = P_{+,c}^0 \mathfrak{H}$  defined in Proposition 1, and the vacuum (Dirac sea) is described by the infinite rank spectral projector  $P_{-,c}^0$ . Recall that in Hartree or Dirac-Fock theories, the density matrix  $\gamma_\Phi$  of the  $N$  electron state  $\Psi = \varphi_1 \wedge \cdots \wedge \varphi_N$  is precisely the orthogonal projector onto  $\text{span}(\varphi_1, \dots, \varphi_N)$ , see, e.g. formulas (89) and (92). Therefore, when one represents the Dirac sea by the projector  $P_{-,c}^0$ , one describes formally the vacuum as an infinite Slater determinant

$$(111) \quad \Omega^0 = \varphi_1 \wedge \cdots \wedge \varphi_i \wedge \cdots$$

where  $(\varphi_i)_{i \geq 1}$  is an orthonormal basis of  $\mathfrak{H}_{-,c}^0 = P_{-,c}^0 \mathfrak{H}$ . Dirac's postulate is based on the important invariance by translation of the projector  $P_{-,c}^0$ , which is used to neglect the physical influence of this 'constant background'.

It was realized just after Dirac's discovery that, for consistency of the theory, the vacuum should not be considered as a totally virtual physical object which does not interact with the real particles. Dirac himself [45, 46, 47] conjectured the existence of surprising physical effects as a consequence of his theory, which were then experimentally confirmed. First, the virtual electrons of the Dirac sea can feel an external field and they will react to this field accordingly, i.e. the vacuum will become *polarized*. This polarization is then felt by the real particles and one therefore is led to consider a coupled system 'Dirac sea + real particles'. From the experimental viewpoint, vacuum polarization plays a rather small role for the calculation of the Lamb shift of hydrogen but it is important for high- $Z$  atoms [126] and it is even a crucial physical effect for muonic atoms [75, 78]. Second, in the presence of strong external fields, the vacuum could react so importantly that an electron-positron pair can be spontaneously created [129, 141, 142, 143].

The mathematical difficulties of a model aiming at describing both the Dirac sea and the real particles are important, for one has to deal at the same time with infinitely many particles (the real ones and the virtual ones of the Dirac sea). In the following, we present a Hartree-Fock (mean-field) type model for this problem, which has been mathematically studied by Hainzl, Lewin, Séré and Solovej [83, 84, 85, 86]. The model under consideration is inspired of an important physical article by Chaix and Iracane [36] in which the possibility that a bounded-below energy could be obtained by adding vacuum polarization was first proposed. But the equations of this so-called Bogoliubov-Dirac-Fock model were already known in Quantum Electrodynamics (QED) [142]. For the sake of simplicity, we shall not explain how the model is derived from the QED Hamiltonian and we refer to [83, 84, 85, 86] and the review [87]. In the version studied in these works, the electromagnetic field is not quantized (photons are not considered).

Let us now describe *formally* the mean-field approximation in QED, following mainly [86, 87]. The state of our system will be represented by an infinite rank projector  $P$ . This projector should be seen as the density matrix of an infinite Slater determinant made of an orthonormal basis of the subspace  $P\mathfrak{H}$ , like (111). The projector  $P$  describes not only Dirac's vacuum but the whole system consisting of the infinitely many virtual particles of the vacuum, together with the finitely many real particles (they could be electrons or positrons). Indeed, it is important to realize that in this model there is no *a priori* possible distinction between the real and the virtual particles. For  $N$  electrons, this would correspond to a decomposition

of the form  $P = P_{\text{vac}} + \gamma$  where  $\gamma$  is an orthogonal projector of rank  $N$  satisfying  $P_{\text{vac}}\gamma = \gamma P_{\text{vac}} = 0$  (for  $N$  positrons, this becomes  $P = P_{\text{vac}} - \gamma$ ). But there are infinitely many such decompositions for a given  $P$  and a given  $N$ : it will only be for the final solution of our equation that this decomposition will be done in a natural way.

For the sake of simplicity, we take  $c = 1$  except at the very end of this section. We recall that in this case an additional parameter  $\alpha = e^2$  appears, where  $e$  is the charge of the electron. The energy in the state  $P$  of our system in the presence of an external field  $V = -\alpha\nu * |\cdot|^{-1}$  is then *formally* given by [86]

$$(112) \quad \mathcal{E}_{\text{QED}}^\nu(P) := \mathcal{E}^\nu(P - 1/2)$$

where  $\mathcal{E}^\nu$  is the Dirac-Fock functional expressed in term of the density matrix, see (98),

$$(113) \quad \mathcal{E}^\nu(\Gamma) := \text{tr}(D_1\Gamma) - \alpha \int_{\mathbb{R}^3} \left( \nu * \frac{1}{|\cdot|} \right) (x) \rho_\Gamma(x) dx \\ + \frac{\alpha}{2} \iint_{\mathbb{R}^3 \times \mathbb{R}^3} \frac{\rho_\Gamma(x) \rho_\Gamma(y)}{|x-y|} dx dy - \frac{\alpha}{2} \iint_{\mathbb{R}^3 \times \mathbb{R}^3} \frac{|\Gamma(x,y)|^2}{|x-y|} dx dy.$$

The density of charge  $\nu$  which creates the field  $V$  can for instance represent a system of nuclei in a molecule. But in the following we shall not allow pointwise nuclei as we did for the Dirac-Fock model, and  $\nu$  will essentially be an  $L^1_{\text{loc}}$  function. This is not important for pointwise particles do not exist in nature.

The subtraction of half the identity in (112) is a kind of *renormalization* which was introduced by Heisenberg [89] and has been widely used by Schwinger (see [147, Eq. (1.14)], [148, Eq. (1.69)] and [149, Eq. (2.3)]) as a necessity for a covariant formulation of QED. The importance of this renormalization will be clarified below.

Before we come to the problems of definition of the above energy (which are numerous), let us mention briefly how  $\mathcal{E}_{\text{QED}}^\nu$  is supposed to be used. There are two possibilities. If one wants to find the state of the vacuum alone in the field  $V$  (no particle at all), then one has to minimize  $P \mapsto \mathcal{E}^\nu(P - 1/2)$  on the whole set of orthogonal projectors. When  $\nu = 0$ , one should obtain the *free vacuum*, a translation-invariant projector which is supposed to be physically unimportant. When  $\nu \neq 0$ , one should obtain the *polarized vacuum* in the presence of the field  $V$ . It can formally be seen that such a minimizer  $P$  should be a solution of the following nonlinear equation:

$$(114) \quad P = \chi_{(-\infty, 0)} \left( D_1 + \alpha(\rho_{[P-1/2]} - \nu) * \frac{1}{|\cdot|} - \alpha \frac{(P-1/2)(x,y)}{|x-y|} \right).$$

If one wants to describe a system containing  $N$  real particles (or more correctly of total charge  $-eN$ ), then one has to minimize  $P \mapsto \mathcal{E}^\nu(P - 1/2)$  under the additional constraint  $\text{tr}(P - 1/2) = N$ . Usually this will provide us with a projector  $P$  which describes adequately the state of a system of charge  $-eN$  interacting with Dirac's vacuum. It is formally solution of the following nonlinear equation

$$(115) \quad P = \chi_{(-\infty, \mu]} \left( D_1 + \alpha(\rho_{[P-1/2]} - \nu) * \frac{1}{|\cdot|} - \alpha \frac{(P-1/2)(x,y)}{|x-y|} \right),$$

$\mu$  being a Lagrange multiplier due to the charge constraint. If  $\nu$  is not too strong, the mean-field operator

$$\mathcal{D} := D_1 + \alpha(\rho_{[P-1/2]} - \nu) * \frac{1}{|\cdot|} - \alpha \frac{(P-1/2)(x, y)}{|x-y|}$$

will have exactly  $N$  eigenvalues in  $[0, \mu]$  (counted with their multiplicity). Therefore, there will be a natural decomposition

$$P = \chi_{(-\infty, 0)}(\mathcal{D}) + \chi_{[0, \mu]}(\mathcal{D}) = \chi_{(-\infty, 0)}(\mathcal{D}) + \sum_{\ell=1}^N \varphi_\ell \otimes \varphi_\ell$$

where the  $\varphi_i$ 's are solutions of

$$(116) \quad \mathcal{D}\varphi_k = \varepsilon_k \varphi_k, \quad \varepsilon_k > 0,$$

which is the natural decomposition between the real and the virtual particles mentioned above. It will be shown below that Equations (116) are very close to the Dirac-Fock equations (93).

The goal is now to give a mathematical meaning to this program.

First, one has to face an important difficulty: when  $P$  is an orthogonal projector and  $\mathfrak{H}$  is infinite dimensional, then  $P - 1/2$  is never compact. Therefore, none of the terms of (113) makes sense. In [86], Hainzl, Lewin and Solovej tackled this problem in the following way: they considered a box  $\mathcal{C}_L := [-L/2, L/2]^3$  of volume  $L^3$  and replaced the ambient space  $\mathfrak{H}$  by the space  $L^2_{\text{per}}(\mathcal{C}_L)$  of  $L^2$  functions on  $\mathcal{C}_L$  with periodic boundary conditions, with moreover a cut-off in Fourier space. Doing so, the problem becomes finite dimensional and everything makes perfect sense. Then, when no charge constraint is imposed, they took the thermodynamic limit  $L \rightarrow \infty$  with the Fourier cut-off fixed and proved that the sequence of minimizers converges to a limiting projector  $P$  satisfying equation (114). In this way, they could give a meaning to the minimization of the QED functional (112) and to the equation (114). Moreover, they also obtained the so-called Bogoliubov-Dirac-Fock (BDF) energy which attains its minimum at the limit state  $P$ . This energy has been studied by Hainzl, Lewin and Séré in [83, 84, 85]. In the following, we describe all these results in more detail.

Notice that the Fourier cut-off will *not* be removed in this study. Indeed it is well-known that QED contains important divergences which are difficult to remove. Therefore  $\mathfrak{H}$  will be replaced by the functional space

$$\mathfrak{H}_\Lambda := \left\{ f \in L^2(\mathbb{R}^3, \mathbb{C}^4), \text{supp}(\widehat{f}) \subset B(0, \Lambda) \right\}.$$

The ideas described in this section have recently been adapted [30] to a nonlinear model of solid state physics describing nonrelativistic electrons in a crystal with a defect.

**4.1. Definition of the free vacuum.** We start by explaining how the free vacuum was constructed in [86], in the case  $\nu = 0$ . This is done by defining the energy (113) for projectors  $P$  acting on the finite-dimensional space

$$\begin{aligned} \mathfrak{H}_\Lambda^L &:= \left\{ f \in L^2_{\text{per}}([-L/2, L/2]^3, \mathbb{C}^4), \text{supp}(\widehat{f}) \subset B(0, \Lambda) \right\} \\ &= \text{span} \left\{ \exp(ik \cdot x), k \in (2\pi/L)\mathbb{Z}^3 \cap B(0, \Lambda) \right\}. \end{aligned}$$

To define the energy properly, it is necessary to periodize the Coulomb potential as follows:

$$W_L(x) = \frac{1}{L^3} \left( \sum_{\substack{k \in (2\pi)\mathbb{Z}^3/L \\ k \neq 0}} \frac{4\pi}{|k|^2} e^{ik \cdot x} + wL^2 \right),$$

where  $w$  is some constant which is chosen such that  $\min_{\mathcal{C}_L} W_L = 0$  for any  $L$ . The Dirac operator  $D_1$  is also easily defined on  $\mathfrak{H}_\Lambda^L$ : it is just the multiplication of the Fourier coefficients by  $(D_1(k))_{k \in 2\pi\mathbb{Z}^3/L}$ . Then, one introduces

$$(117) \quad \mathcal{E}_L^0(\Gamma) := \text{tr}_{\mathfrak{H}_\Lambda^L}(D_1\Gamma) + \frac{\alpha}{2} \iint_{\mathcal{C}_L \times \mathcal{C}_L} W_L(x-y) \rho_\Gamma(x) \rho_\Gamma(y) dx dy \\ - \frac{\alpha}{2} \iint_{\mathcal{C}_L \times \mathcal{C}_L} W_L(x-y) |\Gamma(x,y)|^2 dx dy,$$

for any self-adjoint operator  $\Gamma$  acting on  $\mathfrak{H}_\Lambda^L$ . The kernel  $\Gamma(x,y)$  of  $\Gamma$  is easily defined since  $\mathfrak{H}_\Lambda^L$  is finite-dimensional. Its density  $\rho_\Gamma$  is then defined as  $\rho_\Gamma(x) := \text{tr}_{\mathbb{C}^4}(\Gamma(x,x))$ . A translation-invariant operator  $T$  acting on  $\mathfrak{H}_\Lambda^L$  is by definition a multiplication operator in the Fourier domain. In this case, one has  $T(x,y) = f(x-y)$  for some  $f$  and therefore  $\rho_T$  is constant. The identity of  $\mathfrak{H}_\Lambda^L$ , denoted by  $I_\Lambda^L$  is an example of a translation-invariant operator.

The first result proved in [86] is the following:

**Theorem 29.** (QED mean-field minimizer in a box [86]) *Assume that  $0 \leq \alpha < 4/\pi$ ,  $\Lambda > 0$  and that  $L$  is large enough. Then the functional  $\mathcal{E}_L^0$  has a unique minimizer  $\Gamma_L^0$  on the convex set*

$$\mathcal{G}_\Lambda^L := \{\Gamma \in \mathcal{L}(\mathfrak{H}_\Lambda^L), \Gamma^* = \Gamma, -I_\Lambda^L/2 \leq \Gamma \leq I_\Lambda^L/2\}.$$

*It is invariant by translation and satisfies  $\rho_{\Gamma_L^0} \equiv 0$ . Moreover, it takes the form  $\Gamma_L^0 = \mathcal{P}_L^0 - I_\Lambda^L/2$  where  $\mathcal{P}_L^0$  is an orthogonal projector on  $\mathfrak{H}_\Lambda^L$ .*

Notice that in the definition of the variational set  $\mathcal{G}_\Lambda^L$ , we did not consider only operators taking the form  $P - I_\Lambda^L/2$ , where  $P$  is an orthogonal projector as suggested by (112), but we indeed extended the energy functional to the convex hull of this set. But, as usual in Hartree-Fock type theories [113], the global minimizer is always an extremal point, i.e. a state taking the form  $P - I_\Lambda^L/2$ .

Of course, it can easily be shown that  $\mathcal{P}_L^0$  satisfies an equation similar to (114) with  $\nu$  removed and  $1/|x|$  replaced by  $W_L$ . But we do not give the details since we are more interested in the limit of  $\mathcal{P}_L^0$  as  $L \rightarrow \infty$ .

To be able to state the thermodynamic limit correctly, one needs first to introduce the translation-invariant projector  $\mathcal{P}_\Lambda^0$  acting on  $\mathfrak{H}_\Lambda$ , which will be the limit of the sequence  $(\mathcal{P}_L^0)_L$ . The identity of  $\mathfrak{H}_\Lambda$  is denoted by  $I_\Lambda$ . We introduce

$$(118) \quad \mathcal{T}(A) = \frac{1}{(2\pi)^3} \int_{B(0,\Lambda)} \text{tr}_{\mathbb{C}^4}[D^0(p)A(p)] dp - \frac{\alpha}{(2\pi)^5} \iint_{B(0,\Lambda)^2} \frac{\text{tr}_{\mathbb{C}^4}[A(p)A(q)]}{|p-q|^2} dp dq$$

for any  $A$  belonging to the convex set

$$\mathcal{A}_\Lambda := \{A \text{ translation-invariant on } \mathfrak{H}_\Lambda, A^* = A, -I_\Lambda/2 \leq A \leq I_\Lambda/2\}.$$

It will be shown in Theorem 31 below that  $\mathcal{T}$  represents the energy per unit volume of translation-invariant operators. For this reason, one now considers the minimization of  $\mathcal{T}$  on  $\mathcal{A}_\Lambda$ . The following was proved in [86]:

**Theorem 30.** (Definition of the free vacuum [86]) *Assume that  $0 \leq \alpha < 4/\pi$  and  $\Lambda > 0$ . Then  $\mathcal{T}$  possesses a unique global minimizer  $\Gamma^0$  on  $\mathcal{A}_\Lambda$ . It satisfies the self-consistent equation*

$$(119) \quad \begin{cases} \Gamma^0 = -\frac{\text{sgn}(\mathcal{D}^0)}{2}, \\ \mathcal{D}^0 = D_1 - \alpha \frac{\Gamma^0(x, y)}{|x - y|} \end{cases}$$

or, written in terms of the translation-invariant projector  $\mathcal{P}_-^0 = \Gamma^0 + I_\Lambda/2$ ,

$$(120) \quad \mathcal{P}_-^0 = \chi_{(-\infty, 0)}(\mathcal{D}^0).$$

Moreover,  $\mathcal{D}^0$  takes the special form, in the Fourier domain,

$$(121) \quad \mathcal{D}^0(p) = g_1(|p|)\boldsymbol{\alpha} \cdot p + g_0(|p|)\beta$$

where  $g_0, g_1 \in L^\infty([0, \Lambda], \mathbb{R})$  are such that  $1 \leq g_1(x) \leq g_0(x)$  for any  $x \in [0, \Lambda]$ , and therefore

$$(122) \quad |D_1(p)|^2 \leq |\mathcal{D}^0(p)|^2 \leq g_0(|p|)|D_1(p)|^2.$$

The self-consistent equation (119) has already been solved by Lieb and Siedentop in a different context [116]. They used a fixed point method only valid when  $\alpha \log \Lambda \leq C$  for some constant  $C$ .

As shown by the next result, the negative spectral projector  $\mathcal{P}_-^0$  of the Dirac-type operator  $\mathcal{D}^0$  represents the *free vacuum*, as it is the limit of the sequence  $\mathcal{P}_L^0$  when  $L \rightarrow \infty$ . An important property of  $\Gamma^0$  showing the usefulness of the subtraction of half the identity in (112) is the following. Due to

$$\mathcal{P}_-^0(p) - I_\Lambda(p)/2 = \Gamma^0(p) = -\frac{g_1(|p|)\boldsymbol{\alpha} \cdot p + g_0(|p|)\beta}{2|\mathcal{D}^0(p)|},$$

one infers

$$\text{tr}_{\mathbb{C}^4}(\Gamma^0(p)) = \text{tr}_{\mathbb{C}^4}[(\mathcal{P}_-^0 - I_\Lambda/2)(p)] = 0,$$

for any  $p \in B(0, \Lambda)$ , the Pauli matrices being trace-less. This has the important consequence that the (constant) density of charge of the free vacuum vanishes:

$$\rho_{\Gamma^0} \equiv (2\pi)^{-3} \int_{B(0, \Lambda)} \text{tr}_{\mathbb{C}^4}(\Gamma^0(p)) dp = 0,$$

which is physically meaningful. This formally means that

$$(123) \quad \text{“tr}(\mathcal{P}_-^0 - I_\Lambda/2) = \int_{\mathbb{R}^3} \rho_{\Gamma^0} dx = 0\text{”}.$$

We notice that  $\mathcal{P}_-^0$  is *not* Dirac's original choice  $P_-^0$  (except when  $\alpha = 0$ ) because the interaction between the particles (the virtual and the real ones) is taken into account by the model. Notice also that Equation (120) is exactly the same as (114) with  $\nu = 0$ , due to (119).

As a consequence of (121), the spectrum of  $\mathcal{D}^0$  is

$$\sigma(\mathcal{D}^0) = \left\{ \pm \sqrt{g_0(|p|)^2 + g_1(|p|)^2 |p|^2}, p \in B(0, \Lambda) \right\}.$$

It has a gap which is greater than the one of  $D^0$ , by (122):

$$(124) \quad 1 \leq m(\alpha) := \min \sigma(|\mathcal{D}^0|).$$

In [85], it is proved that when  $\alpha \ll 1$ , then  $m(\alpha) = g_0(0)$  and conjectured this is true for any  $0 \leq \alpha < 4/\pi$ . Notice that the following expansion is known [116, 86]:  $g_0(0) = 1 + \frac{\alpha}{\pi} \operatorname{arcsinh}(\Lambda) + O(\alpha^2)$ .

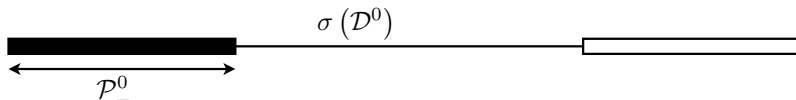


FIGURE 2. The free vacuum  $\mathcal{P}_-^0$  is the negative spectral projector of the translation-invariant operator  $\mathcal{D}^0$ .

We are now able to state the thermodynamic limit, as proved in [86]

**Theorem 31.** (Thermodynamic limit in the free case [86]) *Assume that  $0 \leq \alpha < 4/\pi$  and  $\Lambda > 0$ . Then, one has*

$$\lim_{L \rightarrow \infty} \frac{\mathcal{E}_L^0(\Gamma_L^0)}{L^3} = \min_{\mathcal{A}_\Lambda} \mathcal{T},$$

where we recall that  $\Gamma_L^0$  is the unique minimizer of  $\mathcal{E}_L^0$  defined in Theorem 29. Moreover,  $\mathcal{P}_L^0 = \Gamma_L^0 + I_\Lambda^L/2$  converges to  $\mathcal{P}_-^0$  in the following sense:

$$\lim_{L \rightarrow \infty} \|\mathcal{P}_L^0 - \mathcal{P}_-^0\|_{\mathfrak{S}_\infty(\mathfrak{H}_\Lambda^L)} = \lim_{L \rightarrow \infty} \sup_{p \in (2\pi\mathbb{Z}^3/L) \cap B(0, \Lambda)} |\mathcal{P}_L^0(p) - \mathcal{P}_-^0(p)| = 0.$$

**4.2. The Bogoliubov-Dirac-Fock model.** Now that the free vacuum  $\mathcal{P}_-^0$  has been correctly defined, we will be able to introduce the Bogoliubov-Dirac-Fock (BDF) energy as studied in [83, 84, 85, 86]. Formally, it measures the energy (112) of a state  $P$ , relatively to the (infinite) energy of the free vacuum  $\mathcal{P}_-^0$ . It will only depend on  $P - \mathcal{P}_-^0$  and reads formally:

$$(125) \quad \begin{aligned} \mathcal{E}_{\text{BDF}}^\nu(P - \mathcal{P}_-^0) &= \mathcal{E}_{\text{QED}}^\nu(P) - \mathcal{E}_{\text{QED}}^0(\mathcal{P}_-^0) \\ &= \mathcal{E}^\nu(P - I_\Lambda/2) - \mathcal{E}^0(\mathcal{P}_-^0 - I_\Lambda/2) \\ &= \operatorname{tr}(\mathcal{D}^0 Q) - \alpha \iint_{\mathbb{R}^3 \times \mathbb{R}^3} \frac{\rho_Q(x)\nu(y)}{|x-y|} dx dy \end{aligned}$$

$$(126) \quad + \frac{\alpha}{2} \iint_{\mathbb{R}^3 \times \mathbb{R}^3} \frac{\rho_Q(x)\rho_Q(y)}{|x-y|} dx dy - \frac{\alpha}{2} \iint_{\mathbb{R}^3 \times \mathbb{R}^3} \frac{|Q(x,y)|^2}{|x-y|} dx dy,$$

with  $Q = P - \mathcal{P}_-^0$ . This new energy looks again like a Hartree-Fock type functional except that our main variable is  $Q = P - \mathcal{P}_-^0$ , which measures the difference between our state and the (physically unobservable) translation-invariant free vacuum  $\mathcal{P}_-^0$ .

The energy (126) was introduced and studied by Chaix-Iracane in [36] (see also Chaix-Iracane-Lions [37]). An adequate mathematical formalism was then provided by Bach, Barbaroux, Helffer and Siedentop [6] in the free case  $\nu = 0$ , and by Hainzl, Lewin and Séré [83, 84] in the external field case  $\nu \neq 0$ . However, in all these works a simplified version was considered:  $\mathcal{D}^0$  and  $\mathcal{P}_-^0$  were replaced by Dirac's choice  $D_1$  and  $P_-^0$ . As mentioned in [86], although the choice of  $\mathcal{P}_-^0$  for the free vacuum is better physically, the two models are essentially the same from the mathematical point of view: the main results of [6, 83, 84] can be easily generalized to treat the model in which  $\mathcal{D}^0$  and  $\mathcal{P}_-^0$  are used.

What is gained with (126) is that  $Q$  can now be a compact operator (it will be Hilbert-Schmidt, indeed) and, thanks to the Fourier cut-off  $\Lambda$ , many of the terms in (126) will be mathematically well-defined. However, it will be necessary to generalize the trace functional to define correctly the kinetic energy  $\text{tr}(\mathcal{D}^0 Q)$ . This is done by introducing the following space

$$\mathfrak{S}_1^{\mathcal{P}^0}(\mathfrak{H}_\Lambda) := \{Q \in \mathfrak{S}_2(\mathfrak{H}_\Lambda) \mid \mathcal{P}_-^0 Q \mathcal{P}_-^0, \mathcal{P}_+^0 Q \mathcal{P}_+^0 \in \mathfrak{S}_1(\mathfrak{H}_\Lambda)\},$$

with the usual notation

$$\mathfrak{S}_p(\mathfrak{H}_\Lambda) := \{A \in \mathcal{L}(\mathfrak{H}_\Lambda), \text{tr}(|A|^p) < \infty\}.$$

An operator  $Q$  belonging to  $\mathfrak{S}_1^{\mathcal{P}^0}(\mathfrak{H}_\Lambda)$  is said to be  $\mathcal{P}_-^0$ -trace class. For any such  $Q \in \mathfrak{S}_1^{\mathcal{P}^0}(\mathfrak{H}_\Lambda)$ , we then define its  $\mathcal{P}_-^0$ -trace as

$$\text{tr}_{\mathcal{P}_-^0}(Q) := \text{tr}(\mathcal{P}_-^0 Q \mathcal{P}_-^0) + \text{tr}(\mathcal{P}_+^0 Q \mathcal{P}_+^0).$$

Due to the fact that the free vacuum has a vanishing charge (123),  $\text{tr}_{\mathcal{P}_-^0}(Q)$  can be interpreted as the charge of our state  $P = Q + \mathcal{P}_-^0$ . We refer to [83] for interesting general properties of spaces of the form  $\mathfrak{S}_1^{P'}(H)$  for any projector  $P'$  and infinite-dimensional Hilbert space  $H$ .

Thanks to the cut-off in Fourier space, the charge density  $\rho_Q$  of an operator  $Q \in \mathfrak{S}_1^{\mathcal{P}^0}$  is well-defined in  $L^2(\mathbb{R}^3, \mathbb{R})$ , via

$$\widehat{\rho}_Q(k) = (2\pi)^{-3/2} \int_{\substack{|p+k/2| \leq L \\ |p-k/2| \leq L}} \text{tr}_{\mathbb{C}^4} \left( \widehat{Q}(p+k/2, p-k/2) \right) dp.$$

Finally, the following notation is introduced

$$D(f, g) = 4\pi \int \frac{\widehat{f}(k) \widehat{g}(k)}{|k|^2} dk,$$

for any  $(f, g) \in L^2(\mathbb{R}^3, \mathbb{R})^2$ , which coincides with  $\iint_{\mathbb{R}^3 \times \mathbb{R}^3} f(x)g(y)|x-y|^{-1} dx dy$  when  $f$  and  $g$  are smooth enough.

It is possible now to define the Bogoliubov-Dirac-Fock energy as [83, 84, 86]

$$(127) \quad \mathcal{E}_{\text{BDF}}^\nu(Q) := \text{tr}_{\mathcal{P}_-^0}(\mathcal{D}^0 Q) - \alpha D(\rho_Q, \nu) + \frac{\alpha}{2} D(\rho_Q, \rho_Q) - \frac{\alpha}{2} \iint_{\mathbb{R}^6} \frac{|Q(x, y)|^2}{|x-y|} dx dy,$$

where

$$(128) \quad Q \in \mathcal{Q}_\Lambda := \left\{ Q \in \mathfrak{S}_1^{\mathcal{P}^0}(\mathfrak{H}_\Lambda), -\mathcal{P}_-^0 \leq Q \leq \mathcal{P}_+^0 \right\}.$$

It is proved in [85] that any  $Q \in \mathfrak{S}_1^{\mathcal{P}^0}(\mathfrak{H}_\Lambda)$  automatically has its density  $\rho_Q$  in the following so-called Coulomb space:

$$(129) \quad \mathcal{C} = \{f \mid D(f, f) < \infty\},$$

which is the natural space for defining the terms depending on  $\rho_Q$  in (127). Notice that the set  $\mathcal{Q}_\Lambda$  is one more time the convex hull of our initial states  $P - \mathcal{P}_-^0$  where  $P$  is an orthogonal projector on  $\mathfrak{H}_\Lambda$ . The following was proved:

**Theorem 32** (The BDF energy is bounded-below [37, 6, 83, 84, 86]). *Assume that  $0 \leq \alpha < 4/\pi$ ,  $\Lambda > 0$  and that  $\nu \in \mathcal{C}$ .*

(i) *One has*

$$\forall Q \in \mathcal{Q}_\Lambda, \quad \mathcal{E}_{\text{BDF}}^\nu(Q) + \frac{\alpha}{2}D(\nu, \nu) \geq 0,$$

*and therefore  $\mathcal{E}_{\text{BDF}}^\nu$  is bounded from below on  $\mathcal{Q}_\Lambda$ .*

(ii) *If moreover  $\nu = 0$ , then  $\mathcal{E}_{\text{BDF}}^0$  is non-negative on  $\mathcal{Q}_\Lambda$ , 0 being its unique minimizer.*

The boundedness from below of the BDF energy is an essential feature of the theory. It shows the usefulness of the inclusion of the vacuum effects in the model. The interpretation of (ii) is the following: by (125), it proves that the free vacuum  $\mathcal{P}_-^0$  is the unique minimizer of the (formal) QED energy in the set of all the projectors  $P$  which are such that  $P - \mathcal{P}_-^0 \in \mathcal{Q}_\Lambda$ . In the previous subsection (Theorem 30), it was also proved that  $\mathcal{P}_-^0$  is the unique minimizer of the energy per unit volume  $\mathcal{T}$ . These are two different ways of giving a mathematical meaning to the fact that  $\mathcal{P}_-^0$  is the unique minimizer of the QED energy when no external field is present.

The case  $\nu = 0$ , (ii) in Theorem 32, was proved by Bach, Barbaroux, Helffer and Siedentop [6]. In this paper, the authors also study a relativistic model, but with vacuum polarization neglected (see Section 4.5 below). They were inspired of a paper by Chaix, Iracane and Lions [37]. Then, it has been argued in [83, 84, 86] that the proof of the case  $\nu \neq 0$ , (i) in Theorem 32, is a trivial adaptation of [6].

Now that  $\mathcal{E}_{\text{BDF}}^\nu$  has been shown to be bounded-below, it is natural to try to minimize it. Actually, we shall be interested in two minimization problems. The first is the global minimization of  $\mathcal{E}_{\text{BDF}}^\nu$  in the whole set  $\mathcal{Q}_\Lambda$ . As mentioned above, a global minimizer of the mean-field QED energy  $\mathcal{E}_{\text{QED}}^\nu$  and therefore of the BDF energy  $\mathcal{E}_{\text{BDF}}^\nu$  (they formally differ by an infinite constant !) is interpreted as the *polarized vacuum* in the external electrostatic field  $-\alpha\nu * |\cdot|^{-1}$ . If one wants to describe a system of charge  $-eN$ , one has to minimize  $\mathcal{E}_{\text{BDF}}^\nu$  in the  $N$ th charge sector:

$$\mathcal{Q}_\Lambda(N) := \left\{ Q \in \mathcal{Q}_\Lambda \mid \text{tr}_{\mathcal{P}_-^0}(Q) = N \right\}.$$

These two minimization problems will be tackled in the following two subsections.

Let us also notice that the boundedness from below of the BDF energy (127) has been used by Hainzl, Lewin and Sparber in [88] to prove the existence of *global-in-time* solutions to the time-dependent nonlinear equation associated with the BDF functional:

$$(130) \quad i \frac{d}{dt} P(t) = [\mathcal{D}_{(P(t) - \mathcal{P}_-^0)}, P(t)]$$

where  $[\cdot, \cdot]$  is the usual commutator and

$$\mathcal{D}_Q := \mathcal{D}^0 + \alpha(\rho_Q - \nu) * \frac{1}{|\cdot|} - \alpha \frac{Q(x, y)}{|x - y|}$$

(indeed, like in [83, 84], the model for which  $\mathcal{P}_-^0$  and  $\mathcal{D}^0$  are replaced by  $P_-^0$  and  $D_1$  was considered, but the proof holds similarly in the case of (130)). Global solutions are shown to exist for *any* initial orthogonal projector

$$P(0) \in \mathcal{P}_-^0 + \mathfrak{S}_2(\mathfrak{H}_\Lambda).$$

For nonlinear Dirac theories, it is usually quite difficult to prove the existence of global solutions (see the comments in Section 1.6). Here, the proof is highly simplified by the important property that the BDF energy  $\mathcal{E}_{\text{BDF}}^\nu$ , which is conserved along solutions of (130), is bounded-below and coercive.

Let us now give the proof of Theorem 32, which is a simple adaptation of arguments of [6].

**Proof of Theorem 32.** Let be  $Q \in \mathcal{Q}_\Lambda$ , and therefore satisfying the operator inequality

$$(131) \quad -\mathcal{P}_-^0 \leq Q \leq \mathcal{P}_+^0.$$

It is easily proved that (131) is equivalent to

$$(132) \quad Q^2 \leq Q^{++} - Q^{--}$$

where  $Q^{++} := \mathcal{P}_+^0 Q \mathcal{P}_+^0$  and  $Q^{--} := \mathcal{P}_-^0 Q \mathcal{P}_-^0$ . This now implies that

$$(133) \quad 0 \leq \text{tr}(|\mathcal{D}^0|Q^2) \leq \text{tr}_{\mathcal{P}_-^0}(\mathcal{D}^0 Q)$$

which shows that the kinetic energy is non-negative. We now use Kato's inequality (44) and Equation (122) to obtain

$$\iint_{\mathbb{R}^3 \times \mathbb{R}^3} \frac{|Q(x, y)|^2}{|x - y|} dx dy \leq \frac{\pi}{2} \text{tr}(|D_1|Q^2) \leq \frac{\pi}{2} \text{tr}(|\mathcal{D}^0|Q^2).$$

Together with (133), this implies

$$\mathcal{E}_{\text{BDF}}^\nu(Q) \geq \left(1 - \alpha \frac{\pi}{4}\right) \text{tr}(|\mathcal{D}^0|Q^2) - \frac{\alpha}{2} D(\nu, \nu),$$

which easily ends the proof of Theorem 32.  $\square$

**4.3. Global minimization of  $\mathcal{E}_{\text{BDF}}^\nu$ : the polarized vacuum.** The existence of a global minimizer of  $\mathcal{E}_{\text{BDF}}^\nu$  has been proved by Hainzl, Lewin and Séré, first in [83] by a fixed-point argument valid only when  $\alpha\sqrt{\log \Lambda} \leq C_1$  and  $\alpha D(\nu, \nu)^{1/2} \leq C_2$ , and then by a global minimization procedure in [84], valid for any cut-off  $\Lambda$  and  $0 \leq \alpha < 4/\pi$ . The precise statement of the latter is the following:

**Theorem 33** (Definition of the polarized vacuum [83, 84, 86]). *Assume that  $0 \leq \alpha < 4/\pi$ ,  $\Lambda > 0$  and that  $\nu \in \mathcal{C}$ . Then  $\mathcal{E}_{\text{BDF}}^\nu$  possesses a minimizer  $\bar{Q}$  on  $\mathcal{Q}_\Lambda$  such that  $\mathcal{P} = \bar{Q} + \mathcal{P}_-^0$  is an orthogonal projector satisfying the self-consistent equation*

$$(134) \quad \mathcal{P} = \chi_{(-\infty, 0)}(\mathcal{D}_{\bar{Q}}),$$

$$(135) \quad \mathcal{D}_{\bar{Q}} = \mathcal{D}^0 + \alpha(\rho_{\bar{Q}} - \nu) * \frac{1}{|\cdot|} - \alpha \frac{\bar{Q}(x, y)}{|x - y|}$$

$$(136) \quad = D_1 + \alpha(\rho_{[\bar{\mathcal{P}}_- - 1/2]} - \nu) * \frac{1}{|\cdot|} - \alpha \frac{(\bar{\mathcal{P}}_- - 1/2)(x, y)}{|x - y|}.$$

Additionally, if  $\alpha$  and  $\nu$  satisfy

$$(137) \quad 0 \leq \alpha \frac{\pi}{4} \left\{ 1 - \alpha \left( \frac{\pi}{2} \sqrt{\frac{\alpha/2}{1 - \alpha\pi/4}} + \pi^{1/6} 2^{11/6} \right) D(\nu, \nu)^{1/2} \right\}^{-1} \leq 1,$$

then this global minimizer  $\bar{Q}$  is unique and the associated polarized vacuum is neutral, i.e.  $\bar{Q} \in \mathcal{Q}_\Lambda(0)$ :

$$(138) \quad \text{tr}_{\mathcal{P}_-^0}(\bar{Q}) = \text{tr}_{\mathcal{P}_-^0}(\bar{\mathcal{P}}_- - \mathcal{P}_-^0) = 0.$$

The proof consists in showing that  $\mathcal{E}_{\text{BDF}}^\nu$  is lower semi-continuous for the weak-\* topology of  $\mathcal{Q}_\Lambda$ . For this purpose, one shows that, in the electron-positron field, any mass escaping to infinity takes away a positive energy. This is the so-called dichotomy case of the concentration-compactness principle [118]. To prove (138), one first shows that  $\text{tr}_{\mathcal{P}_-^0}(\bar{\mathcal{P}}_- - \mathcal{P}_-^0)$  is always an integer, then one applies a continuation argument.

Notice that the definition (119) of  $\mathcal{D}^0$  has been used to obtain (136) from (135). Of course, equations (134) and (136) are exactly the one we wanted to solve in the beginning (114). For not too strong external densities  $\nu$ , a neutral vacuum is necessarily obtained, as shown by (138). But in general, a charged polarized vacuum could be found.

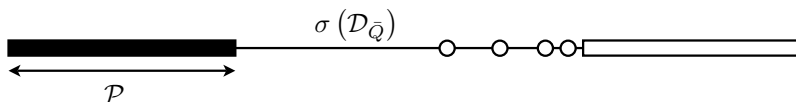


FIGURE 3. The polarized vacuum  $\mathcal{P}$  in the presence of the external density  $\nu$  is the negative spectral projector of the mean-field operator  $\mathcal{D}_{\bar{Q}}$ .

In [86], a thermodynamic limit was considered as for the free case  $\nu = 0$ , to justify the formal computation (126) when  $\nu \neq 0$ . As before, the QED energy is well-defined in a box  $\mathcal{C}_L = [-L/2, L/2]^3$  with periodic boundary conditions and a cut-off in Fourier space by

$$(139) \quad \mathcal{E}_L^\nu(\Gamma) := \mathcal{E}_L^0(\Gamma) - \alpha \iint_{\mathcal{C}_L \times \mathcal{C}_L} W_L(x-y) \rho_\Gamma(x) \nu_L(y) dx dy,$$

where

$$(140) \quad \nu_L(x) = \frac{(2\pi)^{3/2}}{L^3} \sum_{k \in (2\pi\mathbb{Z}^3)/L} \hat{\nu}(k) e^{ik \cdot x},$$

(for simplicity, it is assumed that  $\hat{\nu}$  is a smooth function). Then the following was proved in [86].

**Theorem 34** (Thermodynamic limit with external field [86]). *Assume that  $0 \leq \alpha < 4/\pi$ ,  $\Lambda > 0$ ,  $\nu \in \mathcal{C}$  and that  $\hat{\nu}$  is continuous on  $B(0, \Lambda)$ . Then for any  $L$ ,  $\mathcal{E}_L^\nu$  possesses a minimizer  $\Gamma_L = \mathcal{P}_L - I_\Lambda^L/2$  on  $\mathcal{G}_\Lambda^L$  where  $\mathcal{P}_L$  is an orthogonal projector, and one has*

$$(141) \quad \lim_{L \rightarrow \infty} \{\mathcal{E}_L^\nu(\Gamma_L) - \mathcal{E}_L^0(\Gamma_L^0)\} = \min \{\mathcal{E}_{\text{BDF}}^\nu(Q), Q \in \mathcal{Q}_\Lambda\}.$$

Moreover, up to a subsequence,  $Q_L(x, y) := (\Gamma_L - \Gamma_L^0)(x, y) = (\mathcal{P}_L - \mathcal{P}_L^0)(x, y)$  converges uniformly on compact subsets of  $\mathbb{R}^6$  to  $\bar{Q}(x, y)$ , a minimizer of  $\mathcal{E}_{\text{BDF}}^\nu$  on  $\mathcal{Q}_\Lambda$ .

Since all the previous results hold for any fixed  $\Lambda$ , it would be natural to consider the limit  $\Lambda \rightarrow \infty$ . This was done in [84] where it is argued that this limit is highly

unphysical: the vacuum polarization density *totally cancels the external density*  $\nu$ . In QED, this “nullification” of the theory as the cut-off  $\Lambda$  diverges was first suggested by Landau [107] and later studied by Pomeranchuk *et al.* [133].

Notice that in QED, the procedure of renormalization is often used to formally remove the cut-off  $\Lambda$  and the divergence of the theory. It consists in assuming that the parameter  $\alpha$  is not the physical one but the *bare* one. The physical  $\alpha_{\text{phys}} \simeq 1/137$  is related to  $\alpha$  by a formula of the form

$$\alpha_{\text{phys}} \simeq \frac{\alpha}{1 + 2\alpha \log \Lambda / (3\pi)}.$$

When  $\alpha_{\text{phys}}$  is fixed at its physical value, then necessarily  $\Lambda \leq 10^{280}$ , meaning that the large  $\Lambda$  limit should not be considered in principle. In the case where the exchange term is neglected in the energy (the so-called *reduced* BDF model), this renormalization procedure has been studied in detail in [84].

**4.4. Minimization of  $\mathcal{E}_{\text{BDF}}^\nu$  in charge sectors.** The previous subsection was devoted to the global minimization of the BDF energy. We now mention some results that have been obtained in [85] for the minimization with a charge constraint. It is believed that the charge constrained BDF model can be obtained as the thermodynamical limit of the full QED model in a fixed charge sector and posed in a box with periodic boundary conditions, but this has not been shown yet.

Due to the charge constraint and like for the Hartree-Fock model for instance, minimizers will not always exist for the BDF functional: it depends whether the external electrostatic potential created by the charge distribution  $\nu$  is strong enough to be able to bind the  $N$  particles in the presence of the Dirac sea. On the other hand, it must not be too strong otherwise electron-positron pairs could appear.

We start with a general result proved in [85] providing the form of a minimizer, if it exists. To this end, we introduce the minimum energy in the  $N$ th charge sector:

$$(142) \quad E^\nu(N) := \inf \left\{ \mathcal{E}_{\text{BDF}}^\nu(Q) \mid Q \in \mathcal{Q}_\Lambda, \text{tr}_{\mathcal{P}_-^0}(Q) = N \right\}.$$

In principle  $N$  could be any real number, but here, for simplicity, we shall restrict ourselves to integers.

**Theorem 35** (Self-Consistent Equation of a BDF Minimizer [85]). *Let be  $0 \leq \alpha < 4/\pi$ ,  $\Lambda > 0$ ,  $\nu \in \mathcal{C}$  and  $N \in \mathbb{Z}$ . Then any minimizer  $Q$  solution of the minimization problem (142), if it exists, takes the form  $Q = P - \mathcal{P}_-^0$  where*

$$(143) \quad P = \chi_{(-\infty, \mu]}(\mathcal{D}_Q) = \chi_{(-\infty, \mu]} \left( \mathcal{D}^0 + \alpha(\rho_Q - \nu) * 1/|\cdot| - \alpha \frac{Q(x, y)}{|x - y|} \right),$$

for some  $\mu \in [-m(\alpha), m(\alpha)]$ .

Recall that  $m(\alpha)$  is the threshold of the free operator  $\mathcal{D}^0$  defined in (124). We remark that (143) implicitly means that the last eigenvalue below  $\mu$  of the mean-field operator  $\mathcal{D}_Q$  is necessarily totally filled. As already mentioned in the Dirac-Fock case, this is a general fact for Hartree-Fock type theories [7]. For a minimizer of the form (143) and when  $N, \mu > 0$ , it is natural to consider the decomposition

$$P = \Pi + \chi_{[0, \mu]}(\mathcal{D}_Q),$$

where  $\Pi$  is the polarized Dirac sea:

$$\Pi := \chi_{(-\infty, 0)}(\mathcal{D}_Q).$$

For not too strong external potentials, the vacuum will be neutral, *i.e.*

$$\mathrm{tr}_{\mathcal{P}_-^0}(\Pi - \mathcal{P}_-^0) = 0,$$

and therefore  $\chi_{[0, \mu]}(\mathcal{D}_Q)$  will be a projector of rank  $N$ :

$$\chi_{[0, \mu]}(\mathcal{D}_Q) = \sum_{n=1}^N \varphi_n \otimes \varphi_n^*, \quad \mathcal{D}_Q \varphi_n = \varepsilon_n \varphi_n,$$

where  $\varepsilon_1 \leq \dots \leq \varepsilon_N$  are the first  $N$  positive eigenvalues of  $\mathcal{D}_Q$  counted with their multiplicity. Notice that

$$(144) \quad \mathcal{D}_Q = D_1 + \alpha(\rho_\Phi - \nu) * \frac{1}{|\cdot|} - \alpha \frac{\gamma_\Phi(x, y)}{|x - y|} \\ + \alpha \rho_{[\Pi - I_\Lambda/2]} * \frac{1}{|\cdot|} - \alpha \frac{(\Pi - I_\Lambda/2)(x, y)}{|x - y|},$$

where

$$\gamma_\Phi := \chi_{[0, \mu]}(\mathcal{D}_Q) = \sum_{n=1}^N \varphi_n \otimes \varphi_n^*, \quad \rho_\Phi(x) := \mathrm{tr}_{\mathbb{C}^4}(\gamma_\Phi(x, x)) = \sum_{n=1}^N |\varphi_n(x)|^2.$$

In the first line of (144), the Dirac-Fock operator associated with  $(\varphi_1, \dots, \varphi_N)$  appears, see (95). This shows that the electronic orbitals  $\varphi_i$  are solutions of a Dirac-Fock type equation in which the mean-field operator  $\mathcal{D}_\Phi$  is perturbed by the (self-consistent) potentials of the Dirac sea  $\Pi - I_\Lambda/2$ . In practice, these potentials are small, and the DF equations are a good approximation of the BDF equations for the electronic orbitals. But the energy functionals behave in a completely different way: as we have seen, the DF energy is strongly indefinite while the BDF energy is bounded below. The Dirac-Fock model is thus interpreted as a *non-variational* approximation of the mean-field model of no-photon QED [36].

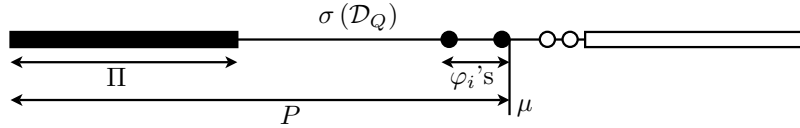


FIGURE 4. Decomposition of the system ‘vacuum +  $N$  electrons’ for the solution  $P = \Pi + \gamma_\Phi$  in the  $N$ th charge sector.

Concerning the existence of a minimizer, solution of (143), the following result was proved in [85]:

**Theorem 36** (Binding Conditions and Existence of a BDF Minimizer [85]). *Let be  $0 \leq \alpha < 4/\pi$ ,  $\Lambda > 0$ ,  $\nu \in \mathcal{C}$  and  $N \in \mathbb{Z}$ . Then the following two assertions are equivalent:*

$$(H_1) \quad E^\nu(N) < \min \{E^\nu(N - K) + E^0(K), K \in \mathbb{Z} \setminus \{0\}\}.$$

(H<sub>2</sub>) *Each minimizing sequence  $(Q_n)_{n \geq 1}$  for  $E^\nu(N)$  is precompact in  $\mathcal{Q}_\Lambda$  and converges, up to a subsequence, to a minimizer  $Q$  of  $E^\nu(N)$ .*

Conditions like  $(H_1)$  appear classically when analyzing the compactness properties of minimizing sequences, for instance by using the concentration-compactness principle of P.-L. Lions [118]. They are also very classical for linear models in which the bottom of the essential spectrum has the form of the minimum in the right hand side of  $(H_1)$ , as expressed by the HVZ Theorem [95, 163, 171]. Assume  $N > 0$  for simplicity. When  $0 < K \leq N$ ,  $(H_1)$  means that it is not favorable to let  $K$  electrons escape to infinity, while keeping  $N - K$  electrons near the nuclei. When  $K < 0$ , it means that it is not favorable to let  $|K|$  positrons escape to infinity, while keeping  $N + |K|$  electrons near the nuclei. When  $K > N$ , it means that it is not favorable to let  $K$  electrons escape to infinity, while keeping  $K - N$  positrons near the nuclei. When  $\alpha$  is small enough and  $N > 0$ , it was shown in [85] that the separation of electron-positron pairs is not energetically favorable, so that one just needs to check  $(H_1)$  for  $K = 1, 2, \dots, N$ .

To prove the existence of a minimizer, one can therefore prove that  $(H_1)$  holds. Two situations in which  $(H_1)$  is true have been provided by Hainzl, Lewin and Séré in [85]. The first one is the case of weak coupling  $\alpha \ll 1$  and  $\alpha\nu = \bar{\nu}$  fixed (the charge  $N$  is also fixed). The following was proved:

**Theorem 37** (Existence of a minimizer in the weak coupling limit [85]). *Assume that  $\Lambda > 0$ , that  $N$  is a non-negative integer, and that  $\bar{\nu} \in \mathcal{C}$  is such that*

- (1) *the spectrum  $\sigma(D^0 - \bar{\nu} * |\cdot|^{-1})$  contains at least  $N$  positive eigenvalues below 1,*
- (2)  *$\ker(D^0 - t\bar{\nu} * |\cdot|^{-1}) = \{0\}$  for any  $t \in [0, 1]$ .*

*Then  $(H_1)$  holds in Theorem 36 for  $\alpha$  small enough and  $\alpha\nu = \bar{\nu}$ , and therefore there exists a minimizer  $Q_\alpha$  of  $E^{\bar{\nu}/\alpha}(N)$ . It takes the form*

$$(145) \quad Q_\alpha = \chi_{(-\infty, 0]}(\mathcal{D}_{Q_\alpha}) - \mathcal{P}_-^0 + \chi_{(0, \mu_\alpha]}(\mathcal{D}_{Q_\alpha}) := Q_\alpha^{\text{vac}} + \sum_{i=1}^N |\varphi_i^\alpha\rangle \langle \varphi_i^\alpha|$$

$$(146) \quad \mathcal{D}_{Q_\alpha} \varphi_i^\alpha = \varepsilon_i^\alpha \varphi_i^\alpha$$

*where  $\varepsilon_1^\alpha \leq \dots \leq \varepsilon_N^\alpha$  are the  $N$  first positive eigenvalues of  $\mathcal{D}_{Q_\alpha}$ . Finally, for any sequence  $\alpha_n \rightarrow 0$ ,  $(\varphi_1^{\alpha_n}, \dots, \varphi_N^{\alpha_n})$  converges (up to a subsequence) in  $\mathfrak{H}_\Lambda$  to  $(\varphi_1, \dots, \varphi_N)$  which are  $N$  first eigenfunctions of  $D^0 - \bar{\nu} * |\cdot|^{-1}$  and  $Q_{\alpha_n}^{\text{vac}}$  converges to  $\chi_{(-\infty; 0)}(D^0 - \bar{\nu} * |\cdot|^{-1}) - \mathcal{P}_-^0$  in  $\mathfrak{S}_2(\mathfrak{H}_\Lambda)$ .*

The second situation provided in [85] is the case of the non-relativistic regime  $c \gg 1$ . To state the result correctly, we reintroduce the speed of light  $c$  in the model (of course, we shall then take  $\alpha = 1$ ). The expression of the energy and the definition of the free vacuum  $\mathcal{P}_-^0$  (which of course then depends on  $c$  and the ultraviolet cut-off  $\Lambda$ ) are straightforward. We denote by  $E_{\alpha, c, \Lambda}^\nu(N)$  the minimum energy of the BDF functional depending on the parameters  $(\alpha, c, \Lambda)$ . The following was proved:

**Theorem 38** (Existence of a minimizer in the non-relativistic limit [85]). *Assume that  $\alpha = 1$  and that the ultraviolet cut-off is  $\Lambda = \Lambda_0 c$  for some fixed  $\Lambda_0$ . Let be  $\nu \in \mathcal{C} \cap L^1(\mathbb{R}^3, \mathbb{R}^+)$  with  $\int_{\mathbb{R}^3} \nu = Z$ , and  $N$  a positive integer which is such that  $Z > N - 1$ . Then, for  $c$  large enough,  $(H_1)$  holds in Theorem 36 and therefore there*

exists a minimizer  $Q_c$  for  $E_{1,c,\Lambda_0c}^\nu(N)$ . It takes the following form:

$$Q_c = \chi_{(-\infty,0)}(\mathcal{D}_Q) - \mathcal{P}_-^0 + \chi_{[0,\mu_c)}(\mathcal{D}_Q) = Q_c^{\text{vac}} + \sum_{i=1}^N |\varphi_i^c\rangle\langle\varphi_i^c|,$$

and one has

$$\lim_{c \rightarrow \infty} \{E_{1,c,\Lambda_0c}^\nu(N) - Ng_0(0)\} = \min_{\substack{\bar{\Phi} \in H^1(\mathbb{R}^3, \mathbb{C}^2)^N \\ \text{Gram } \bar{\Phi} = \mathbf{1}}} \mathcal{E}_{\text{HF}}(\bar{\Phi}).$$

Moreover, for any sequence  $c_n \rightarrow \infty$ ,  $(\varphi_1^{c_n}, \dots, \varphi_N^{c_n})$  converges in  $H^1(\mathbb{R}^3, \mathbb{C}^4)^N$  (up to a subsequence) towards  $(\varphi_1, \dots, \varphi_N)$  with  $\varphi_i = \begin{pmatrix} \bar{\varphi}^i \\ 0 \end{pmatrix}$ , and where  $\bar{\Phi}_0 = (\bar{\varphi}_1, \dots, \bar{\varphi}_N)$  is a global minimizer of the Hartree-Fock energy.

We notice that this result is very similar to Theorem 26 providing the convergence of the Dirac-Fock ‘ground state’ in the non-relativistic limit.

**4.5. Neglecting Vacuum Polarization: Mittleman’s conjecture.** In view of the complications introduced by the Dirac sea, some authors [5, 6, 14] considered an approximate model in which the vacuum polarization is neglected, in the spirit of a paper by Mittleman [125]. In this subsection, we shall keep the speed of light  $c$  and therefore take  $\alpha = 1$ .

In the vacuum case, a possible way to describe Mittleman’s approach is first to write that the global minimizer  $\mathcal{P} - 1/2$  constructed in Theorem 33 is formally a solution of the following tautological max-min principle:

$$(147) \quad e = \sup_{P^2=P} \inf_{-P \leq \gamma \leq 1-P} \{ \mathcal{E}_{\text{QED}}^\nu(P + \gamma) - \mathcal{E}_{\text{QED}}^\nu(P) \}.$$

Indeed taking  $\gamma = 0$ , one finds that  $e \leq 0$ . Saying that  $\mathcal{P}$  is a global QED minimizer exactly means that

$$\min_{-P \leq \gamma \leq 1-P} \{ \mathcal{E}_{\text{QED}}^\nu(P + \gamma) - \mathcal{E}_{\text{QED}}^\nu(P) \} = 0.$$

Then, the idea is to approximate (147) by neglecting the vacuum polarization terms coming from  $P$ . One has formally

$$(148) \quad \begin{aligned} \mathcal{E}_{\text{QED}}^\nu(P + \gamma) - \mathcal{E}_{\text{QED}}^\nu(P) &= \text{tr}(D^P \gamma) - \iint_{\mathbb{R}^3 \times \mathbb{R}^3} \frac{\rho_\gamma(x)\nu(y)}{|x-y|} dx dy \\ &+ \frac{1}{2} \iint_{\mathbb{R}^3 \times \mathbb{R}^3} \frac{\rho_\gamma(x)\rho_\gamma(y)}{|x-y|} dx dy - \frac{1}{2} \iint_{\mathbb{R}^3 \times \mathbb{R}^3} \frac{|\gamma(x,y)|^2}{|x-y|} dx dy \end{aligned}$$

with

$$D^P := D_c + \rho_{[P-1/2]} * \frac{1}{|\cdot|} - \frac{(P-1/2)(x,y)}{|x-y|}.$$

Neglecting the vacuum polarization potentials then simply amounts to replacing  $D^P$  by the free Dirac operator  $D_c$ . The following max-min principle was studied by Bach, Barbaroux, Helffer and Siedentop in [6], inspired by Mittleman [125]:

$$(149) \quad e_{\text{Mitt}}^{\nu,c} = \sup_{P^2=P=P^*} \inf_{\substack{\gamma \in \mathfrak{S}_1(\mathfrak{H}_\Lambda), \\ -P \leq \gamma \leq 1-P}} \mathcal{E}_P^{\nu,c}(\gamma)$$

where  $\mathcal{E}_P^{\nu,c}$  is defined as

$$(150) \quad \mathcal{E}_P^{\nu,c}(\gamma) := \text{tr}(D_c \gamma) - D(\nu, \rho_\gamma) + \frac{1}{2}D(\rho_\gamma, \rho_\gamma) - \frac{1}{2} \iint_{\mathbb{R}^6} \frac{|\gamma(x, y)|^2}{|x - y|} dx dy$$

on the set depending on  $P$ :

$$\Gamma(P) := \{\gamma \in \mathfrak{S}_1(\mathfrak{H}), |\nabla|\gamma \in \mathfrak{S}_1(\mathfrak{H}), -P \leq \gamma \leq 1 - P\}.$$

The advantage of this formulation is that, since the vacuum polarization has been neglected, no divergence problem is encountered and (149) can be studied without any Fourier cut-off (i.e.  $\mathfrak{H} = H^{1/2}(\mathbb{R}^3, \mathbb{C}^4)$ ), and with pointwise external charges (i.e.  $\nu = Z\delta_0$  where  $\delta_0$  is the Dirac distribution at the point  $0 \in \mathbb{R}^3$ ). The following was proved in [6]:

**Theorem 39** (Mittleman Principle for the Vacuum [6]). *Assume that  $\nu = Z\delta_0$  for some  $Z \geq 0$  and that  $c > 0$  satisfies  $4c(1 - 2Z/c)/\pi \geq 1$ . Then*

$$\bar{P}_c := \chi_{(-\infty, 0)}(D_c - Z/|x|)$$

is the unique solution of Mittleman's max-min principle

$$(151) \quad e_{\text{Mitt}}^{\nu,c} = \sup_{P \in \mathcal{S}} \inf_{\gamma \in \Gamma(P)} \mathcal{E}_P^{\nu,c}(\gamma),$$

where  $\mathcal{S}$  denotes the set of all the orthogonal projectors  $P$  which are such that  $P$  and  $1 - P$  leave the domain of  $D_c - Z/|x|$  invariant. Moreover,  $e_{\text{Mitt}}^{\nu,c} = 0$ .

As a consequence, when vacuum polarization is neglected, the Dirac sea is represented by the negative spectral projector of the operator  $D_c - Z/|x|$ .

The  $N$  electron case was studied by Mittleman in [125]. His main idea was to justify the validity of the Dirac-Fock approximation by a type of max-min principle from QED in which vacuum polarization is neglected. Let us introduce the following max-min principle:

$$(152) \quad e_{\text{Mitt}}^{\nu,c}(N) = \sup_{P \in \mathcal{S}} \inf_{\substack{\gamma \in \Gamma(P) \\ \gamma P = P\gamma, \\ \text{tr}(\gamma) = N}} \mathcal{E}_P^{\nu,c}(\gamma).$$

The interpretation is that  $\gamma$  represents the  $N$  electrons, whereas  $P$  is the Dirac sea. Notice that in this interpretation, the real particles are artificially separated from the virtual electrons of the Dirac sea. However, it was believed that (152) could be a simpler problem with interesting practical implications. Remark that compared to (149), we have added in (152) the condition that  $\gamma$  commutes with  $P$ , i.e. it cannot contain off-diagonal terms. Without this requirement, it is known [4] that (152) cannot give any solution of the Dirac-Fock equations.

Mittleman's conjecture consists in saying that the max-min principle (152) is attained by a solution to the Dirac-Fock system. More precisely, we follow [13] and state it as:

**Mittleman's conjecture.** *A solution of (152) is given by a pair*

$$(P, \gamma) = (\chi_{(-\infty, 0)}(D_{c, \Phi}), \gamma_\Phi)$$

where  $\Phi$  is a solution of the Dirac-Fock equation  $D_{c, \Phi} \varphi_i = \varepsilon_i \varphi_i$  with  $\varepsilon_i > 0$ , and where  $D_{c, \Phi}$  is the Dirac-Fock mean-field operator defined in (95).

Mittleman's conjecture for molecules has been investigated in [14, 13]. In [14] Barbaroux, Farkas, Helffer and Siedentop have studied the minimization problem

on  $\Gamma(P)$  in (152) for a fixed  $P$ . Under suitable conditions, they proved that  $\gamma$  is an electronic density matrix  $\gamma = \sum_{i=1}^N |\varphi_i\rangle\langle\varphi_i|$ , where  $(\varphi_1, \dots, \varphi_N)$  is a solution of the projected Dirac-Fock equations  $(1-P)D_{c,\Phi}(1-P)\varphi_i = \varepsilon_i\varphi_i$ . Then they further proved that if the energy is stationary with respect to variations of the projector  $P$ , then the unprojected Dirac-Fock are obtained. But they were unable to prove the existence of such a state.

In [13], Barbaroux, Esteban and Séré investigated the validity of Mittleman's conjecture by a perturbation argument. Namely they added a parameter  $\tau$  in front of the interaction terms as follows:

$$(153) \quad \mathcal{E}_P^{\nu,c,\tau}(\gamma) := \text{tr}(D_c\gamma) - D(\nu, \rho_\gamma) + \frac{\tau}{2}D(\rho_\gamma, \rho_\gamma) - \frac{\tau}{2} \iint_{\mathbb{R}^6} \frac{|\gamma(x,y)|^2}{|x-y|} dx dy.$$

Here  $\nu$  is a fixed positive and smooth radial function with compact support and  $\int_{\mathbb{R}^3} \nu = 1$ . They studied Mittleman's conjecture in the regime  $c \gg 1$  and  $\tau \ll 1$ . We emphasize that by a scaling argument this physically corresponds to assuming  $\alpha \ll 1$  and  $Z \gg 1$  with  $\alpha Z \ll 1$ . We denote by  $\mathcal{E}_{\text{DF}}^{\nu,c,\tau}$  the Dirac-Fock functional which is easily defined with these parameters, and by  $\Phi^{c,\tau,0}$  the DF solution obtained by the corresponding versions of Theorems 23 and 27 for  $j = 0$ .

When  $c > 1$ , it is known that  $D_c - \nu * |\cdot|^{-1}$  is essentially self-adjoint on  $L^2(\mathbb{R}^3, \mathbb{C}^4)$  and that its spectrum is as follows:

$$\sigma(D_c - \nu * |\cdot|^{-1}) = (-\infty, -c^2] \cup \{\mu_1(c) < \mu_2(c) \dots\} \cup [c^2, \infty),$$

where  $\lim_{i \rightarrow \infty} \mu_i(c) = c^2$ . We denote by  $(N_i(c))_{i \geq 1}$  the multiplicities of the  $\mu_i(c)$ 's. The following was proved in [13]:

**Theorem 40** (Validity and non validity of Mittleman's conjecture [13]). *Assume that  $\nu$  is a fixed positive and smooth radial function with compact support and such that  $\int_{\mathbb{R}^3} \nu = 1$ .*

*If  $N = \sum_{i=1}^I N_i(c)$  for some  $I \geq 1$  and some fixed  $c > 1$ , then Mittleman's conjecture is **true** for  $\tau$  small enough: one has*

$$e_{\text{Mitt}}^{\nu,c,\tau}(N) = \mathcal{E}_{\text{DF}}^{\nu,c,\tau}(\Phi^{c,\tau,0}),$$

*where  $\Phi^{c,\tau,0}$  is any of the solutions obtained in Theorem 23 in the case  $j = 0$ . Moreover, the optimal projector for the sup part of (152) is  $P = \chi_{(-\infty,0)}(D_{c,\Phi^{c,\tau,0}})$ .*

*If  $N = \sum_{i=1}^I N_i(c) + 1$  for some  $I \geq 1$  and  $c$  large enough, then Mittleman's conjecture is **wrong** when  $\tau > 0$  is small enough: there is no solution  $\Phi$  of the Dirac-Fock equations with positive multipliers such that the pair  $(\chi_{(-\infty,0)}(D_{c,\Phi}), \gamma_\Phi)$  realizes Mittleman's max-min principle (152).*

One can consider a weaker version of Mittleman's conjecture which consists in only comparing energy levels and not the solutions themselves.

**Weaker Mittleman's conjecture.** *One has*

$$e_{\text{Mitt}}^{\nu,c,\tau}(N) = \mathcal{E}_{\text{DF}}^{\nu,c,\tau}(\Phi^{c,\tau,0}).$$

When  $N = 1$  and for  $c \gg 1$ ,  $\tau \ll 1$ , it is known by Theorem 40 that Mittleman's conjecture is wrong. But Barbaroux, Helffer and Siedentop proved in [15] that the weaker conjecture is indeed true.

**Theorem 41** (Validity of the weaker Mittleman conjecture for  $N = 1$  [15]). *Assume that  $\nu = \delta_0$ . Then for  $c$  large enough and  $\tau$  small enough, the weaker Mittleman conjecture is true:*

$$e_{\text{Mitt}}^{\nu,c,\tau}(N) = \mathcal{E}_{\text{DF}}^{\nu,c,\tau}(\Phi^{c,\tau,0}).$$

The sup-inf (152) is realized by the pair

$$(\bar{P}_c, \gamma) \quad \text{with} \quad \bar{P}_c = \chi_{(-\infty,0)}(D_c - 1/|x|), \quad \gamma = |\varphi^c\rangle\langle\varphi^c|$$

where  $\varphi^c$  is any eigenvector of

$$(154) \quad D_c - 1/|x|$$

with eigenvalue  $e_{\text{Mitt}}^{\nu,c,\tau}(N)$ .

In [15], an explicit condition on  $\tau$  and  $c$  is provided. Notice the equality

$$D_{c,\varphi}\varphi^c = D_{c,0}\varphi^c = (D_c - 1/|x|)\varphi^c = e_{\text{Mitt}}^{\nu,c,\tau}(N)\varphi^c$$

which is very specific to the one-electron case, and means that the electron does not “see” its own Coulomb field. In particular  $\bar{P}_c\varphi^c = 0$ .

For  $N \geq 2$ , the question remains completely open.

#### REFERENCES

- [1] S. Abenda. Solitary waves for Maxwell-Dirac and Coulomb-Dirac models. *Ann. Inst. H. Poincaré, Phys. Théor.* **68(2)** (1998), p. 229–244.
- [2] H. Amann. Saddle points and multiple solutions of differential equations. *Math. Z.* **169** (1979), p. 127–166.
- [3] C.D. Anderson. The Positive Electron. *Phys. Rev.* **43** (1933), p. 491–494.
- [4] W.H. Aschbacher. Lowering the Hartree-Fock minimizer by electron-positron pair correlation. *Lett. Math. Phys.* **70(1)**, (2004), p. 29–41.
- [5] V. Bach, J.M. Barbaroux, B. Helffer, H. Siedentop. Stability of matter for the Hartree-Fock functional of the relativistic electron-positron field. *Doc. Math.* **3** (1998), p. 353–364.
- [6] V. Bach, J.M. Barbaroux, B. Helffer, H. Siedentop. On the stability of the relativistic electron-positron field. *Comm. Math. Phys.* **201(2)** (1999), p. 445–460.
- [7] V. Bach, E.H. Lieb, M. Loss, J.P. Solovej. There are no unfilled shells in unrestricted Hartree-Fock theory. *Phys. Rev. Lett.* **72(19)** (1994), p. 2981–2983.
- [8] A. Bachelot. Problème de Cauchy global pour des systèmes de Dirac-Klein-Gordon. *Ann. Inst. H. Poincaré* **48** (1988), p. 387–422.
- [9] A. Bachelot-Motet. Nonlinear Dirac fields on the Schwarzschild metric. *Class. Quantum Grav.* **15** (1998), p. 1815–1825.
- [10] A. Bahri and J.-M. Coron. On a nonlinear elliptic equation involving the critical Sobolev exponent: the effect of the topology of the domain. *Comm. Pure Appl. Math.* **41(3)** (1988), p. 253–294.
- [11] M. Balabane, T. Cazenave, A. Douady, F. Merle. Existence of excited states for a nonlinear Dirac field. *Comm. Math. Phys.* **119** (1988), p. 153–176.
- [12] M. Balabane, T. Cazenave, L. Vazquez. Existence of standing waves for Dirac fields with singular nonlinearities. *Comm. Math. Phys.* **133** (1990), p. 53–74.
- [13] J.-M. Barbaroux, M. J. Esteban, E. Séré. Some connections between Dirac-Fock and electron-positron Hartree-Fock. *Ann. Henri Poincaré* **6(1)** (2005), p. 85–102.
- [14] J.-M. Barbaroux, W. Farkas, B. Helffer, H. Siedentop. On the Hartree-Fock Equations of the Electron-Positron Field. *Commun. Math. Phys.* **255** (2005), p. 131–159.
- [15] J.-M. Barbaroux, B. Helffer and H. Siedentop. Remarks on the Mittleman max-min variational method for the electron-positron field. *J. Phys. A* **39(1)** (2006), p. 85–98.
- [16] W.E. Bayliss, S. J. Peel. Stable variational calculations with the Dirac Hamiltonian. *Phys. Rev. A*, **28(4)** (1983), p. 2552–2554.
- [17] V. Benci, P.H. Rabinowitz. Critical point theorems for indefinite functionals. *Inv. Math.* **52** (1979), p. 336–352.

- [18] A. Berthier, V. Georgescu. Sur le spectre ponctuel de l'opérateur de Dirac. C. R. Acad. Sci. Paris, Sér. A **297** (1983), p. 335–338.
- [19] A. Berthier, V. Georgescu. On the point spectrum for Dirac operators. J. Func. Anal. **71** (1987), p. 309–338.
- [20] J.D. Bjorken, S.D. Drell. *Relativistic quantum fields*. McGraw-Hill (1965).
- [21] N. Boussaid. Stable directions for small nonlinear Dirac standing waves. Comm. Math. Phys. **268**(3) (2006), p. 757–817.
- [22] N. Bournaveas. Local existence for the Maxwell-Dirac equations in three space dimensions. Commun. Partial Diff. Eqs. **21**(5–6) (1996), p. 693–720.
- [23] N. Bournaveas. Local existence of energy class solutions for the Dirac-Klein-Gordon equations. Comm. Partial Differ. Eqs. **24**(7–8) (1999), p. 1167–1193.
- [24] H. Booth. Electrons with self-field as solutions to nonlinear PDE. Geometric analysis and applications (Canberra, 2000), p. 1–14, Proc. Centre Math. Appl. Austral. Nat. Univ., 39, Austral. Nat. Univ., Canberra, 2001.
- [25] H. Brézis and L. Nirenberg. Positive solutions of nonlinear elliptic equations involving critical Sobolev exponents. Comm. Pure Appl. Math. **36**(4) (1983), p. 437–477.
- [26] B. Buffoni, M.J. Esteban, E. Séré. Normalized solutions to strongly indefinite semilinear equations. Adv. Nonlinear Stud. 6, No. 2 (2006), p. 323–347.
- [27] B. Buffoni, L. Jeanjean. Minimax characterization of solutions for a semi-linear elliptic equation with lack of compactness. Ann. Inst. H. Poincaré **10**(4) (1993), p. 377–404.
- [28] B. Buffoni, L. Jeanjean, C. A. Stuart. Existence of a nontrivial solution to a strongly indefinite semilinear equation. Proc. Amer. Math. Soc. **119**(1) (1993), p. 179–186
- [29] V.I. Burenkov, W.D. Evans. On the evaluation of the norm of an integral operator associated with the stability of one-electron atoms. Proc. Roy. Soc. Edinburgh, sect. A **128**(5) (1998), p. 993–1005.
- [30] É. Cancès, A. Deleurence and M. Lewin. A new approach to the modelling of local defects in crystals: the reduced Hartree-Fock case. Commun. Math. Phys., in press.
- [31] A. Castro, A.C. Lazer. Applications of a min-max principle. Rev. Colomb. Mat. **10** (1976), p. 141–149.
- [32] T. Cazenave. On the existence of stationary states for classical nonlinear Dirac fields. In *Hyperbolic systems and Mathematical Physics*. Textos e Notas **4**, CMAF, Lisbonne (1989).
- [33] T. Cazenave, L. Vazquez. Existence of localized solutions for a classical nonlinear Dirac field. Comm. Math. Phys. **105** (1986), p. 35–47.
- [34] J. Chadam. Global solutions of the Cauchy problem for the (classical) coupled Maxwell-Dirac system in one space dimension. J. Funct. Anal. **13** (1973), p. 173–184.
- [35] J. Chadam, R. Glassey. On the Maxwell-Dirac equations with zero magnetic field and their solutions in two space dimension. J. Math. Anal. Appl. **53** (1976), p. 495–507.
- [36] P. Chaix, D. Iracane. The Bogoliubov-Dirac-Fock formalism. I. J. Phys. At. Mol. Opt. Phys. **22** (1989), p. 3791–3814.
- [37] P. Chaix, D. Iracane, P.-L. Lions. The Bogoliubov-Dirac-Fock formalism. II. J. Phys. At. Mol. Opt. Phys. **22** (1989), p. 3815–3828.
- [38] Y. Choquet-Bruhat. Solutions globales des équations de Maxwell-Dirac-Klein-Gordon (masses nulles). C.R. Acad. Sci. Paris, Série I **292** (1981), p. 153–158.
- [39] F.H. Clarke and I. Ekeland. Hamiltonian trajectories having prescribed minimal period. Comm. Pure Appl. Math. **33**(2) (1980), p. 103–116.
- [40] C.V. Coffman. Uniqueness of the ground state solution for  $\Delta u - u + u^3 = 0$  and a variational characterization of other solutions. Arch. Rat. Mech. Anal. **46** (1972), p. 81–95.
- [41] C. Conley, E. Zehnder. The Birkhoff-Lewis fixed point theorem and a conjecture of V.I. Arnold. Invent. Math. **73** (1983), p. 33–49.
- [42] S.N. Datta, G. Deviah. The minimax technique in relativistic Hartree-Fock calculations. Pramana **30**(5) (1988), p. 393–416.
- [43] J.P. Dias, M. Figueira. Global existence of solutions with small initial data in  $H^s$  for the massive nonlinear Dirac equations in three space dimensions. Boll. Un. Mat. Ital. B (7) **1**(3) (1987), p. 861–874.
- [44] P.A.M. Dirac. The quantum theory of the electron. Proc. Roy. Soc. A **117** (1928), p. 610–624.

- [45] P.A.M. Dirac. A theory of electrons and protons. *Proc. Roy. Soc. A* **126** (1930), p. 360–365.
- [46] P.A.M. Dirac. Théorie du positron. Solvay report (1934), 203–212. Gauthier-Villars, Paris. XXV, 353 S.
- [47] P.A.M. Dirac. Discussion of the infinite distribution of electrons in the theory of the positron. *Proc. Camb. Philos. Soc.* **30** (1934), p. 150–163.
- [48] J. Dolbeault, M.J. Esteban, M. Loss. Relativistic hydrogenic atoms in strong magnetic fields. *Ann. H. Poincaré* **8(4)** (2007), p. 749–779.
- [49] J. Dolbeault, M.J. Esteban, M. Loss, L. Vega. An analytical proof of Hardy-like inequalities related to the Dirac operator. *J. Funct. Anal.* **216** (2004), p. 1–21.
- [50] J. Dolbeault, M.J. Esteban, E. Séré. Variational characterization for eigenvalues of Dirac operators. *Calc. Var. and P.D.E.* **10** (2000), p. 321–347.
- [51] J. Dolbeault, M.J. Esteban, E. Séré. On the eigenvalues of operators with gaps. Application to Dirac operators. *J. Funct. Anal.* **174** (2000), p. 208–226.
- [52] J. Dolbeault, M.J. Esteban, E. Séré. A variational method for relativistic computations in atomic and molecular physics. *Int. J. Quantum. Chemistry* **93** (2003), p. 149–155.
- [53] J. Dolbeault, M. J. Esteban et E. Séré. General results on the eigenvalues of operators with gaps, arising from both ends of the gaps. Application to Dirac operators. *J. Eur. Math. Soc.* **8(2)** (2006), p. 243–251.
- [54] J. Dolbeault, M.J. Esteban, E. Séré, M. Vanbreugel. Minimization methods for the one-particle Dirac equation. *Phys. Rev. Letters* **85(19)** (2000), p. 4020–4023.
- [55] G.W.F. Drake, S.P. Goldman. Application of discrete-basis-set methods to the Dirac equation. *Phys. Rev. A* **23** (1981), p. 2093–2098.
- [56] G.W.F. Drake, S.P. Goldman. Relativistic Sturmian and finite basis set methods in atomic physics. *Adv. Atomic Molecular Phys.* **23** (1988), p. 23–29.
- [57] G.W.F. Drake, S.P. Goldman. *Adv. Atomic Molecular Phys.* **25** (1988), 393.
- [58] Ph. Durand. Transformation du Hamiltonien de Dirac en Hamiltoniens variationnels de type Pauli. Application à des atomes hydrogénéoïdes. *C. R. Acad. Sc. Paris* **303(2)** (1986), p. 119–124.
- [59] Ph. Durand, J.-P. Malrieu. Effective Hamiltonians and pseudo-potentials as tools for rigorous modelling. In *Ab initio methods in Quantum Chemistry I*. K.P. Lawley ed.. J. Wiley & Sons, 1987.
- [60] M. Escobedo, L. Vega. A semilinear Dirac equation in  $H^s(R^3)$  for  $s > 1$ . *SIAM J. Math. Anal.* **28(2)** (1997), p. 338–362.
- [61] M.J. Esteban, V. Georgiev, E. Séré. Stationary solutions of the Maxwell-Dirac and the Klein-Gordon-Dirac equations. *Cal. Var.* **4** (1996), p. 265–281.
- [62] M.J. Esteban, E. Séré. Stationary states of the nonlinear Dirac equation: a variational approach. *Comm. Math. Phys.* **171** (1995), p. 323–350.
- [63] M.J. Esteban, E. Séré. Existence and multiplicity of solutions for linear and nonlinear Dirac problems. *Partial Differential Equations and Their Applications*. CRM Proceedings and Lecture Notes, vol. 12. Eds. P.C. Greiner, V. Ivrii, L.A. Seco and C. Sulem. AMS, 1997.
- [64] M.J. Esteban, E. Séré. Solutions for the Dirac-Fock equations for atoms and molecules. *Comm. Math. Phys.* **203** (1999), p. 499–530.
- [65] M.J. Esteban, E. Séré. Nonrelativistic limit of the Dirac-Fock equations. *Ann. H. Poincaré* **2** (2001), p. 941–961.
- [66] M.J. Esteban, E. Séré. A max-min principle for the ground state of the Dirac-Fock functional. *Contemp. Math.* **307** (2002), p. 135–141.
- [67] W.D. Evans, P. Perry, H. Siedentop. The spectrum of relativistic one-electron atoms according to Bethe and Salpeter. *Comm. Math. Phys.* **178** (1996), p. 733–746.
- [68] G. Fang, N. Ghoussoub. Morse-type information on Palais-Smale sequences obtained by min-max principles. *Comm. Pure Appl. Math.* **47(12)** (1994), p. 1595–1653.
- [69] R. Finkelstein, R. Lelevier, M. Ruderman. Nonlinear Spinor Fields. *Phys. Rev.* **83** (1951), p. 326–332.
- [70] R. Finkelstein, C. Fronsdal, P. Kaus. Nonlinear Spinor Field. *Phys. Rev.* **103** (1956), p. 1571–1579.
- [71] F. Finster, J. Smoller and S.-T. Yau. Particle-like solutions of the Einstein-Dirac-Maxwell equations. *Phys. Lett. A* **259(6)** (1999), p. 431–436.

- [72] F. Finster, J. Smoller and S.-T. Yau. Non-existence of black hole solutions for a spherically symmetric, static Einstein-Dirac-Maxwell system *Comm. Math. Phys.* **205**(2) (1999), p. 249–262.
- [73] M. Flato, J. Simon, E. Taffin. On the global solutions of the Maxwell-Dirac equations. *Comm. Math. Physics* **113** (1987), p. 21–49.
- [74] A. Floer. Morse theory for fixed points of symplectic diffeomorphisms. *Bull. Amer. Math. Soc.* **16**(2) (1987), p. 279–281.
- [75] L.L. Foldy, E. Eriksen. Some Physical Consequences of Vacuum Polarization. *Phys. Rev.* **95**(4) (1954), p. 1048–1051.
- [76] V. Georgiev. Small amplitude solutions of the Maxwell-Dirac equations. *Indiana Univ. Math. J.* **40**(3) (1991), p. 845–883.
- [77] N. Ghoussoub. *Duality and perturbation methods in critical point theory*. Cambridge Univ. Press, 1993.
- [78] R. Glauber, W. Rarita, P. Schwed. Vacuum Polarization Effects on Energy Levels in  $\mu$ -Mesonic Atoms. *Phys. Rev.* **120**(2) (1960), p. 609–613.
- [79] W.T. Grandy, Jr. *Relativistic Quantum Mechanics of Leptons and Fields*. Kluwer Acad. Publisher, Fund. Theories of Physics, Vol. **41**.
- [80] M. Griesemer, R.T. Lewis, H. Siedentop. A minimax principle in spectral gaps: Dirac operators with Coulomb potentials. *Doc. Math.* **4** (1999), p. 275–283 (electronic).
- [81] M. Griesemer, H. Siedentop. A minimax principle for the eigenvalues in spectral gaps. *J. London Math. Soc.* **60**(2) (1999), p. 490–500.
- [82] L. Gross. The Cauchy problem for the coupled Maxwell and Dirac equations. *Comm. Pure Appl. Math.* **19** (1966), p. 1–5.
- [83] C. Hainzl, M. Lewin, É. Séré. Existence of a stable polarized vacuum in the Bogoliubov-Dirac-Fock approximation. *Comm. Math. Phys.* **257**(3) (2005), p. 515–562.
- [84] C. Hainzl, M. Lewin, É. Séré. Self-consistent solution for the polarized vacuum in a no-photon QED model, *J. Phys. A: Math. and Gen.* **38** (2005), p. 4483–4499.
- [85] C. Hainzl, M. Lewin, É. Séré. Existence of atoms and molecules in the mean-field approximation of no-photon Quantum Electrodynamics. Preprint [arXiv:math-ph/0606001](https://arxiv.org/abs/math-ph/0606001).
- [86] C. Hainzl, M. Lewin, J.P. Solovej. The mean-field approximation in Quantum Electrodynamics. The no-photon case. *Comm. Pure Appl. Math.* **60**(4) (2007), p. 546–596.
- [87] C. Hainzl, M. Lewin, É. Séré and J.P. Solovej. A Minimization Method for Relativistic Electrons in a Mean-Field Approximation of Quantum Electrodynamics. *Phys. Rev. A* **76** (2007), 052104.
- [88] C. Hainzl, M. Lewin, C. Sparber. Existence of global-in-time solutions to a generalized Dirac-Fock type evolution equation, *Lett. Math. Phys.* **72** (2005), p. 99–113.
- [89] W. Heisenberg. Bemerkungen zur Diracschen Theorie des Positrons. *Z. Phys.* **90** (1934), p. 209–223.
- [90] W. Heisenberg. Doubts and hopes in quantum electrohydrodynamics. *Physica* **19** (1953), p. 897–908.
- [91] I.W. Herbst. Spectral theory of the operator  $(p^2 + m^2)^{1/2} - Ze^2/r$ . *Comm. Math. Phys.* **53** (1977), p. 285–294.
- [92] R.N. Hill, C. Krauthauser. A solution to the problem of variational collapse for the one-particle Dirac equation. *Phys. Rev. Lett.* **72**(14) (1994), p. 2151–2154.
- [93] H. Hofer, K. Wysocki. First order elliptic systems and the existence of homoclinic orbits in Hamiltonian systems. *Math. Ann.* **288** (1990), p. 483–503.
- [94] M. Huber and H. Siedentop. Solutions of the Dirac-Fock Equations and the Energy of the Electron-Positron Field. *Arch. Rat. Mech. Anal.* **184** (1) (2007), p. 1–22.
- [95] W. Hunziker. On the Spectra of Schrödinger Multiparticle Hamiltonians. *Helv. Phys. Acta* **39** (1966), p. 451–462.
- [96] W. Hunziker, I.M. Sigal. The quantum  $N$ -body problem. *J. Math. Phys.* **41**(6) (2000), p. 3448–3510.
- [97] D. Ivanenko. *Soviet Physics* **13** (1938), p. 141–149.
- [98] T. Kato. *Perturbation theory for linear operators*. Springer, 1966.
- [99] T. Kato. Holomorphic families of Dirac operators. *Math. Z.* **183**(3) (1983), p. 399–406.
- [100] S. Klainerman. Global existence of small amplitude solutions to the nonlinear Klein-Gordon equations in four space dimensions. *Comm. Pure Appl. Math.* **38** (1985), p. 631–641.

- [101] S. Klainerman. The null condition and global existence to nonlinear wave equation. Lect. in Appl. Math. **23** (1986), p. 293–326.
- [102] S. Klainerman. Remarks on the global Sobolev inequalities in the Minkowski space  $\mathbb{R}^{n+1}$ . Comm. Pure Appl. Math. **40** (1986), p. 111–117.
- [103] M. Klaus, R. Wüst. Characterization and uniqueness of distinguished self-adjoint extensions of Dirac operators. Comm. Math. Phys. **64(2)** (1978-79), p. 171–176.
- [104] W. Kutzelnigg. Basis Set Expansion of the Dirac Operator without Variational Collapse. Int. J. Quant. Chem. **25** (1984), p. 107–129.
- [105] W. Kutzelnigg. Relativistic one-electron Hamiltonians “for electrons only” and the variational treatment of the Dirac equation. Chemical Physics **225** (1997), p. 203–222.
- [106] M.K. Kwong, Y. Li. Uniqueness of radial solutions of semilinear elliptic equations. Trans. A.M.S. **333** (1992), 339–363.
- [107] L.D. Landau. On the quantum theory of fields. *Bohr Volume*, Pergamon Press, Oxford, 1955. Reprinted in *Collected papers of L.D. Landau*, edited by D. Ter Haar, Pergamon Press, 1965.
- [108] C. Le Bris, P.-L. Lions. From atoms to crystals: a mathematical journey. Bull. Amer. Math. Soc. (N.S.) **42(3)** (2005), p. 291–363.
- [109] E. van Lenthe, E.J. Baerends, J.G. Snijders. Solving the Dirac equation, using the large component only, in a Dirac-type Slater orbital basis set. Chem. Phys. Lett. **236** (1995), p. 235–241.
- [110] E. van Lenthe, R. van Leeuwen, E.J. Baerends, J.G. Snijders. Relativistic regular two-component Hamiltonians. In *New challenges in computational Quantum Chemistry*. R. Broek et al ed. Publications Dept. Chem. Phys. and Material sciences. University of Groningen, 1994.
- [111] R. van Leeuwen, E. van Lenthe, E.J. Baerends, J.G. Snijders. Exact solutions of regular approximate relativistic wave equations for hydrogen-like atoms. J. Chem. Phys. **101(2)** (1994), p. 1272–1281.
- [112] A. Le Yaouanc, L. Oliver, J.-C. Raynal. The Hamiltonian  $(p^2 + m^2)^{1/2} - \alpha/r$  near the critical value  $\alpha_c = 2/\pi$ . J. Math. Phys. **38(8)** (1997), p. 3397–4012.
- [113] E.H. Lieb. Variational Principle for Many-Fermion Systems. Phys. Rev. Lett. **46** (1981), p. 457–459.
- [114] E.H. Lieb. On the lowest eigenvalue of the Laplacian for the intersection of two domains. Invent. Math. **74(3)** (1983), p. 441–448.
- [115] E.H. Lieb. Bound on the maximum negative ionization of atoms and molecules. Phys. Rev. A. **29** (1984), p. 3018–3028.
- [116] E.H. Lieb and H. Siedentop. Renormalization of the regularized relativistic electron-positron field. Comm. Math. Phys. **213(3)** (2000), p. 673–683.
- [117] E. H. Lieb, B. Simon. The Hartree-Fock theory for Coulomb systems. Comm. Math. Phys., **53** (1977), p. 185–194.
- [118] P.-L. Lions. The concentration-compactness method in the Calculus of Variations. The locally compact case. Part. I: Anal. non-linéaire, Ann. IHP **1** (1984), p. 109–145. Part. II: Anal. non-linéaire, Ann. IHP **1** (1984), p. 223–283.
- [119] P.-L. Lions. Solutions of Hartree-Fock equations for Coulomb systems. Comm. Math. Phys. **109** (1987), p. 33–97.
- [120] A.G. Lisi. A solitary wave solution of the Maxwell-Dirac equations. J. Phys. A **28(18)** (1995), p. 5385–5392.
- [121] S. Machihara, K. Nakanishi, T. Ozawa. Small global solutions and the nonrelativistic limit for the nonlinear Dirac equation. Rev. Mat. Iberoamericana **19(1)** (2003), p. 179–194.
- [122] N. Masmoudi, K. Nakanishi. Nonrelativistic limit from Maxwell-Klein-Gordon and Maxwell-Dirac to Poisson-Schrödinger. In. Math. Res. **13** (2003), p. 697–734.
- [123] N. Masmoudi, K. Nakanishi. Uniqueness of Finite Energy Solutions for Maxwell-Dirac and Maxwell-Klein-Gordon Equations. Commun. Math. Phys. **243** (2003), p. 123–136.
- [124] F. Merle. Existence of stationary states for nonlinear Dirac equations. J. Diff. Eq. **74(1)** (1988), p. 50–68.
- [125] M.H. Mittleman. Theory of relativistic effects on atoms: Configuration-space Hamiltonian. Phys. Rev. A **24(3)** (1981), p. 1167–1175.

- [126] P. J. Mohr, G. Plunien, G. Soff. QED Corrections in Heavy Atoms. *Phy. Rep.* **293(5&6)** (1998), p. 227–372.
- [127] B. Najman. The nonrelativistic limit of the nonlinear Dirac equation. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **9(1)** (1992), p. 3–12.
- [128] G. Nenciu. Self-adjointness and invariance of the essential spectrum for Dirac operators defined as quadratic forms. *Comm. Math. Phys.* **48** (1976), p. 235–247.
- [129] G. Nenciu. Existence of spontaneous pair creation in the external field approximation of Q.E.D., *Commun. Math. Phys.* **109** (1987), p. 303–312.
- [130] H. Ounaies. Perturbation method for a class of nonlinear Dirac equations. *Diff. Int. eqs.* **13(4-6)** (2000), p. 707–720.
- [131] E. Paturel. Solutions of the Dirac equations without projector. *A.H.P.* **1** (2000), p. 1123–1157.
- [132] P. Pickl. Existence of Spontaneous Pair Creation. Dissertation, München (2005), [arXiv:hep-th/0609200](https://arxiv.org/abs/hep-th/0609200).
- [133] I.Ya Pomeranchuk, V.V. Sudakov and K.A. Ter-Martirosyan. Vanishing of renormalized charges in field theories with point interaction. *Phys. Rev.* **103 (3)** (1956), p. 784–802.
- [134] P.H. Rabinowitz. Free vibrations for a semilinear wave equation. *Comm. Pure Applied Math.* **31(1)** (1978), p. 31–68.
- [135] P.H. Rabinowitz. Periodic solutions of Hamiltonian systems. *Comm. Pure Applied Math.* **31(2)** (1978), p. 157–184.
- [136] C.J. Radford. Localized solutions of the Dirac-Maxwell equations. *J. Math. Phys.* **37(9)** (1996), p. 4418–4433.
- [137] C.J. Radford. The stationary Maxwell-Dirac equations. *J. Phys. A* **36(20)** (2003), p. 5663–5681.
- [138] A.F. Rañada. Classical nonlinear Dirac field models of extended particles. In *Quantum theory, groups, fields and particles* (editor A.O. Barut). Reidel, Amsterdam (1982).
- [139] M. Reed. *Abstract non-linear wave equations*. Lecture Notes in Math. **507**, Springer-Verlag, Berlin, 1976.
- [140] M. Reed, B. Simon. *Methods of modern mathematical physics, Vol. 4*. Academic Press, New York, 1978.
- [141] J. Reinhardt, W. Greiner. Quantum Electrodynamics of Strong Fields. *Rep. Prog. Phys.* **40** (1977), p. 219–295.
- [142] P.-G. Reinhard, W. Greiner, H. Arenhövel. Electrons in Strong External Fields. *Nucl. Phys. A* **166** (1971), p. 173–197.
- [143] J. Reinhardt, B. Müller, W. Greiner. Theory of positron production in heavy-ion collision. *Phys. Rev. A*, **24(1)** (1981), p. 103–128.
- [144] M.B. Ruskai. Absence of discrete spectrum in highly negative ions. II. Extension to fermions. *Comm. Math. Phys.* **85(2)** (1982), p. 325–327.
- [145] J. Sacks and K. Uhlenbeck. The existence of minimal immersions of 2-spheres. *Ann. of Math.* **113(1)** (1981), p. 1–24.
- [146] U.W. Schmincke. Distinguished self-adjoint extensions of Dirac operators. *Math. Z.* **129** (1972), p. 335–349.
- [147] J. Schwinger Quantum Electrodynamics I. A Covariant Formulation. *Phys. Rev.* **74(10)** (1948), p. 1439–1461.
- [148] J. Schwinger Quantum Electrodynamics II. Vacuum Polarization and Self-Energy. *Phys. Rev.* **75(4)** (1949), p. 651–679.
- [149] J. Schwinger On Gauge Invariance and Vacuum Polarization. *Phys. Rev., II. Ser.* **82(5)** (1951), p. 664–679.
- [150] E. Séré. Homoclinic orbits on compact hypersurfaces in  $R^{2N}$ , of restricted contact type. *Comm. Math. Phys.* **172(2)** (1995), p. 293–316.
- [151] I.M. Sigal. Geometric methods in the quantum many-body problem. Nonexistence of very negative ions. *Comm. Math. Phys.* **85(2)** (1982), p. 309–324.
- [152] I.M. Sigal. How many electrons can a nucleus bind ? *Annals of Phys.* **157** (1984), p. 307–320.
- [153] S. Smale. An infinite dimensional version of Sard’s Theorem. *Am. J. Math.* **87** (1965), p. 861–866.
- [154] M. Soler. Classical, Stable, Nonlinear Spinor Field with Positive Rest Energy . *Phys. Rev. D* **1** (1970), p. 2766–2769.

- [155] W.A. Strauss, L. Vazquez. Stability under dilations of nonlinear spinor fields. *Phys. Rev. D* **34(2)** (1986), p. 641–643.
- [156] B. Swirles. The relativistic self-consistent field. *Proc. Roy. Soc. A* **152** (1935), p. 625–649.
- [157] W. Kryszewski, A. Szulkin. Generalized linking theorem with an application to a semi-linear Schrödinger equation. *Adv. Differential Equations* **3(3)** (1998), p. 441–472.
- [158] J.D. Talman. Minimax principle for the Dirac equation. *Phys. Rev. Lett.* **57(9)** (1986), p. 1091–1094.
- [159] B. Thaller. *The Dirac Equation*. Springer-Verlag, 1992.
- [160] C. Tix. Lower bound for the ground state energy of the no-pair Hamiltonian. *Phys. Lett. B* **405** (1997), p. 293–296.
- [161] C. Tix. Strict positivity of a relativistic Hamiltonian due to Brown and Ravenhall. *Bull. London Math. Soc.* **30(3)** (1998), p. 283–290.
- [162] XX Troestler, M. Willem. Nontrivial solution of a semilinear Schrödinger equation. *Comm. P.D.E.* **21(9-10)** (1996), p. 1431–1449.
- [163] C. Van Winter. Theory of Finite Systems of Particles. I. The Green function. *Mat.-Fys. Skr. Danske Vid. Selsk.* **2(8)** (1964).
- [164] L. Vázquez. Localised solutions of a non-linear spinor field. *J. Phys. A* **10(8)** (1977), p. 1361–1368.
- [165] S.A. Vugalter and G.M. Zhislin. Finiteness of a discrete spectrum of many-particle Hamiltonians in symmetry spaces (coordinate and momentum representations). *Teoret. Mat. Fiz.* **32(1)** (1977), p. 70–87.
- [166] M. Wakano. Intensely localized solutions of the classical Dirac-Maxwell field equations. *Progr. Theor. Phys.* **35(6)** (1966), p. 1117–1141.
- [167] H. Wallmeier, W. Kutzelnigg. Use of the squared Dirac operator in variational relativistic calculations. *Chem. Phys. Lett.* **78(2)** (1981), p. 341–346.
- [168] M.I. Weinstein. Modulational stability of ground states of nonlinear Schrödinger equations. *Siam J. Math. Anal.* **16(3)** (1985), p. 472–491.
- [169] H. Weyl. A remark on the coupling of gravitation and electron. *Phys. Rev.* **77** (1950), p. 699–701.
- [170] R. Wüst. Dirac operations with strongly singular potentials. Distinguished self-adjoint extensions constructed with a spectral gap theorem and cut-off potentials. *Math. Z.* **152(3)** (1977), p. 259–271.
- [171] G. M. Zhislin. A study of the spectrum of the Schrödinger operator for a system of several particles. (Russian) *Trudy Moskov. Mat. Obšč.* **9** (1960), p. 81–120.
- [172] G. M. Zhislin, A.G. Sigalov. The spectrum of the energy operator for atoms with fixed nuclei on subspaces corresponding to irreducible representations of the group of permutations. (Russian) *Izv. Akad. Nauk SSSR Ser. Mat.* **29** (1965), p. 835–860. *Transl. A.M.S. Ser. 2* **91**, p. 263–296 (English translation).

CNRS AND CEREMADE (UMR 7534), UNIVERSITÉ PARIS-DAUPHINE, PLACE DU MARÉCHAL DE LATTRE DE TASSIGNY, 75775 PARIS CEDEX 16 - FRANCE  
*E-mail address:* [esteban@ceremade.dauphine.fr](mailto:esteban@ceremade.dauphine.fr)

CNRS AND LABORATOIRE DE MATHÉMATIQUES (UMR 8088), UNIVERSITÉ DE CERGY-PONTOISE, 2, AVENUE ADOLPHE CHAUVIN, 95 302 CERGY-PONTOISE CEDEX - FRANCE  
*E-mail address:* [Mathieu.Lewin@math.cnrs.fr](mailto:Mathieu.Lewin@math.cnrs.fr)

CEREMADE (UMR 7534), UNIVERSITÉ PARIS-DAUPHINE, PLACE DU MARÉCHAL DE LATTRE DE TASSIGNY, 75775 PARIS CEDEX 16 - FRANCE  
*E-mail address:* [sere@ceremade.dauphine.fr](mailto:sere@ceremade.dauphine.fr)