Rev. Mat. Iberoam. 40 (2024), no. 2, 649–692
DOI 10.4171/RMI/1443

Keller and Lieb–Thirring estimates of the eigenvalues in the gap of Dirac operators

Jean Dolbeault, David Gontier, Fabio Pizzichillo and Hanne Van Den Bosch

Abstract. We estimate the lowest eigenvalue in the gap of the essential spectrum of a Dirac operator with mass in terms of a Lebesgue norm of the potential. Such a bound is the counterpart for Dirac operators of the Keller estimates for the Schrödinger operator, which are equivalent to some Gagliardo–Nirenberg–Sobolev interpolation inequalities. Domain, self-adjointness, optimality and critical values of the norms are addressed, while the optimal potential is given by a Dirac equation with a Kerr nonlinearity. A new critical bound appears, which is the smallest value of the norm of the potential for which eigenvalues may reach the bottom of the gap in the essential spectrum. The Keller estimate is then extended to a Lieb–Thirring inequality for the eigenvalues in the gap. Most of our result are established in the Birman–Schwinger reformulation.

1. Introduction and main results

In 1961, J. B. Keller established in [45] the expression of the potential which minimizes the lowest eigenvalue, or *ground state*, $\lambda_S(V)$ of the Schrödinger operator $-\Delta - V$ in dimension $d = 1$, under a constraint on the Lebesgue norm $\|V\|_p = (\int_{\mathbb{R}^d} |V|^p \,\mathrm{d}x)^{1/p}$ of exponent p of V. This estimate was later extended in [53] by E. H. Lieb and W. Thirring to higher dimensions and to a sum of the lowest eigenvalues. During the last forty years, various refinements were published. As an example, we quote stability results for $\lambda_S(V)$ proved in [12] by E. A. Carlen, R. L. Frank, and E. H. Lieb. Although Dirac operators inherit many qualitative properties of Schrödinger operators, dealing with Dirac operators turns out to be a delicate issue.

If $\mathcal{D}_m$ denotes the free Dirac operator and V is a nonnegative valued function, then $\mathcal{D}_m - V$ is not bounded from below. One is actually interested in the lowest eigenvalue $\lambda_D(V)$ in the essential gap $(-mc^2, mc^2)$, where m denotes the mass and c the speed of light. We shall speak of $\lambda_D(V)$ as the *ground state* energy of $\mathcal{D}_m - V$. In

2020 Mathematics Subject Classification: Primary 81Q10; Secondary 49R05, 49J35, 47A75, 47B25.
Keywords: Dirac operators, potential, spectral gap, eigenvalues, ground state, min-max principle, Birman–Schwinger operator, domain, self-adjoint operators, Keller estimate, Lieb–Thirring inequality, interpolation, Gagliardo–Nirenberg–Sobolev inequality, Kerr nonlinearity.

the standard setting, it is expected that $\lambda_D(V) - mc^2$ converges to $\lambda_S(V)$ in the non-relativistic limit, i.e., as $c \to +\infty$. It is therefore a natural question to estimate $\lambda_D(V)$ in terms of $\|V\|_p$ and identify the corresponding optimal potential. This question is the main purpose of our paper. A new *critical value* appears, which corresponds to the smallest value of $\|V\|_p$ for which $\lambda_D(V)$ reaches, for some potential $V \geq 0$, the lower end of the essential gap $-mc^2$. In a linear setting, a similar question has been raised in [34, 35], where the authors find a critical value ν_1 so that $\lambda_D(\mu * |\cdot|^{-1}) > -mc^2$ for all positive measures μ with $\mu(\mathbb{R}^3) < \nu_1$, with $2/(\pi/2 + 2/\pi) < \nu_1 \leq 1$. Going back to [21, 27, 28], it is known that *Hardy inequalities* play an essential role in the analysis of the spectrum of Dirac–Coulomb operators. In the present article, except for the case $p = d = 1$, we rather find a nonlinear functional inequality of Gagliardo–Nirenberg–Sobolev nature, instead of a Hardy inequality (see comments in Appendix C.2).

It is possible to characterize the eigenvalues of $\mathcal{D}_m - V$ in the gap by a *min-max principle* according to [28–30], but this raises delicate issues involving the domain of the operator and its self-adjoint extensions addressed in [30, 33, 36, 37, 63]. Applied with a Coulombian potential V, the method gives rise, after the maximizing step in the min-max method, to a lower bounded quadratic form which amounts to a kind of Hardy inequality for the upper component: see [10, 21, 27] for details. The same strategy applies to a general potential V under a constraint on $\|V\|_p$, except that the Keller-type bound on $\lambda_D(V)$ is given by an implicit condition: see Appendix C. The optimal potential solves a nonlinear Dirac equation with Kerr-type nonlinearity. For the two-dimensional case, this equation has been studied in [5–8] by W. Borrelli. In the one-dimensional case, the solution is explicit, which allows us to identify it as in the case of the Schrödinger operator studied in [45]. Alternatively to the min-max principle, the properties of the *Birman–Schwinger operator* corresponding to $\mathcal{D}_m - V$ allows us to characterize $\lambda_D(V)$ and, except in Appendix C, we will adopt this point of view.

The *Keller–Lieb–Thirring inequality for a Schrödinger operator* goes as follows. Let us assume that $q > 2$, with $q < 2^* := 2d/(d-2)$ if $d \geq 3$, and let $\vartheta = d\,(q-2)/(2q)$. For any function $u \in \mathrm{H}^1(\mathbb{R}^d)$, the Gagliardo–Nirenberg–Sobolev inequality

$$\|\nabla u\|_2^{\vartheta}\,\|u\|_2^{1-\vartheta} \geq \mathscr{C}_q\,\|u\|_q$$

can be rewritten in the non-scale invariant form as

$$\forall\,(\lambda, u) \in (0, +\infty) \times \mathrm{H}^1(\mathbb{R}^d), \quad \|\nabla u\|_2^2 + \lambda\,\|u\|_2^2 \geq \mathsf{C}_q\,\lambda^{1-\vartheta}\,\|u\|_q^2, \tag{1.1}$$

with an optimal constant C_q such that $\mathscr{C}_q^2 = \vartheta^{\vartheta}(1-\vartheta)^{1-\vartheta}\,\mathsf{C}_q$. The equivalence of the two forms can be recovered by optimizing on λ in (1.1). There is also an inequality which is dual of (1.1) and goes as follows. Consider a potential $V \in \mathrm{L}^p(\mathbb{R}^d)$. Using Hölder's inequality with exponents p and q such that $1/p + 2/q = 1$ and $p > d/2$, and taking λ so that $\mathsf{C}_q\,\lambda^{1-\vartheta} = \|V\|_p$, we deduce from (1.1) that

$$\int_{\mathbb{R}^d} |\nabla u|^2\,\mathrm{d}x - \int_{\mathbb{R}^d} V\,|u|^2\,\mathrm{d}x \geq \|\nabla u\|_2^2 - \|V\|_p\,\|u\|_q^2 \geq -\left(\mathsf{C}_q^{-1}\,\|V\|_p\right)^{1/(1-\vartheta)}\,\|u\|_2^2.$$

This is the Keller–Lieb–Thirring estimate for $-\Delta - V$, i.e.,

$$\forall\,V \in \mathrm{L}^p(\mathbb{R}^d), \quad 0 \leq \lambda_S^-(V) \leq \mathsf{K}_p\,\|V\|_p^{\eta}, \tag{1.2}$$

where $\eta := 1/(1-\vartheta) = 2p/(2p-d)$ and $\lambda^- := \max(0, -\lambda)$ denotes the negative part of λ. See [22–24] for details. An optimization on V shows that (1.1) and (1.2) are equivalent. The optimal constant in (1.2) is $\mathsf{K}_p = \mathsf{C}_q^{-\eta}$. In addition, for all $\lambda > 0$, if u is a radial positive solution of

$$-\Delta u - u^{(p+1)/(p-1)} = -\lambda u, \tag{1.3}$$

then (u, λ) is an optimal pair for (1.1), and $V := u^{q-2} = u^{2/(p-1)}$ is an optimal potential for (1.2), which moreover satisfies $\lambda_S(V) = -\lambda$. It turns out that the solution of (1.3) is unique up to translations according to [16, 49, 55], and that it can be explicitly computed if $d = 1$: see [45] or [23], and references therein, for additional related results.

In order to state a *Keller–Lieb–Thirring inequality for the Dirac operator*, we need some definitions and preliminary properties. Let us start with the free Dirac operator on $\mathbb{R}^d$. We refer to [68] for a comprehensive list of results and properties. For simplicity, we choose units in which $c = 1$, except in Appendix C, in which we consider the non-relativistic limit as $c \to +\infty$. Let $d \ge 1$ and set $N := 2^{\lfloor (d+1)/2 \rfloor}$, where $\lfloor x \rfloor = \max\{n \in \mathbb{Z} : n \le x\}$ denotes the integer part of x. Let $\alpha_1, \dots, \alpha_d$ and β be $N \times N$ Hermitian matrices satisfying the following anti-commutation rules:

$$\forall\, j,k = 1,\dots,d, \quad \begin{cases} \alpha_j\,\alpha_k + \alpha_k\,\alpha_j = 2\,\delta_{jk}\,\mathbb{I}_N, \\ \alpha_j\,\beta + \beta\,\alpha_j = 0, \\ \beta^2 = \mathbb{I}_N, \end{cases} \tag{1.4}$$

where δ_{jk} denotes the Kronecker symbol and $\mathbb{I}_N$ is the $N \times N$ identity matrix. See, e.g., [41] for an existence result for such matrices. The *free Dirac operator* in dimension d is defined by

$$\mathcal{D}_m := \sum_{j=1}^{d} \alpha_j\,(-\mathrm{i}\,\partial_j) + m\,\beta = \boldsymbol{\alpha} \cdot (-\mathrm{i}\nabla) + m\,\beta,$$

where we consider Cartesian coordinates $(x_1, \dots, x_d)$, $\partial_j := \partial/\partial x_j$ and $\boldsymbol{\alpha} = (\alpha_k)_{k=1,\dots,d}$. With the Pauli matrices

$$\sigma_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 := \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} \quad \text{and} \quad \sigma_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

explicit expressions of $\mathcal{D}_m$ are given

(i) in dimension $d = 1$, by $\boldsymbol{\alpha} = \sigma_2$ and $\beta = \sigma_3$ so that

$$\mathcal{D}_m := \sigma_2\,(-\mathrm{i}\,\partial_1) + m\sigma_3,$$

(ii) in dimension $d = 2$, by $\boldsymbol{\alpha} = (\sigma_j)_{j=1,2}$ and $\beta = \sigma_3$ so that

$$\mathcal{D}_m := \sum_{j=1}^{2} \sigma_j(-\mathrm{i}\,\partial_j) + m\sigma_3,$$

(iii) in dimension $d = 3$, by $\boldsymbol{\alpha} = (\alpha_k)_{k=1,2,3}$ and β such that

$$\alpha_k := \begin{pmatrix} 0 & \sigma_k \\ \sigma_k & 0 \end{pmatrix} \quad \text{and} \quad \beta := \begin{pmatrix} \mathbb{I}_2 & 0 \\ 0 & -\mathbb{I}_2 \end{pmatrix}.$$

The free Dirac operator satisfies $\mathcal{D}_m^2 = -\Delta + m^2$. It is self-adjoint on $\mathrm{L}^2(\mathbb{R}^d, \mathbb{C}^N)$, with domain

$$\mathrm{Dom}(\mathcal{D}_m) = \mathrm{H}^1(\mathbb{R}^d, \mathbb{C}^N)$$

and spectrum

$$\sigma(\mathcal{D}_m) = \sigma_{\mathrm{ess}}(\mathcal{D}_m) = (-\infty, -m] \cup [m, +\infty).$$

Next we consider *Dirac operators* $\mathcal{D}_m - V$ *with potentials* $V \in \mathrm{L}^p(\mathbb{R}^d, \mathbb{R}^+)$, where the notation $\mathcal{D}_m - V$ is a shorthand for $\mathcal{D}_m - V\,\mathbb{I}_N$. When switching on a potential V, we expect that some eigenvalues of $\mathcal{D}_m - V$ emerge from the upper essential spectrum $[m, +\infty)$. We shall prove in Section 2 that $\mathcal{D}_m - V$ can be defined as a self-adjoint operator with *essential spectrum* $\sigma_{\mathrm{ess}}(\mathcal{D}_m - V) = \sigma_{\mathrm{ess}}(\mathcal{D}_m)$. This allows us to define the *ground state* $\lambda_D(V)$ as the lowest eigenvalue in the gap $(-m, m)$.

Our first result states that the *ground state* is bounded by a function of $\|V\|_p$. Let

$$\Lambda_D(\alpha, p) := \inf\left\{\lambda_D(V) \,:\, V \in \mathrm{L}^p(\mathbb{R}^d, \mathbb{R}^+) \text{ and } \|V\|_p = \alpha\right\}. \tag{1.5}$$

Theorem 1.1. *Assume that* $p \geq d \geq 1$. *Then there exists* $\alpha_\star(p) > 0$ *such that the map* $\alpha \mapsto \Lambda_D(\alpha, p)$ *defined on* $[0, \alpha_\star(p))$ *is continuous, strictly decreasing, takes values in* $(-m, m]$, *and is such that*

$$\lim_{\alpha \to 0_+} \Lambda_D(\alpha, p) = m \quad \textit{and} \quad \lim_{\alpha \to \alpha_\star(p)} \Lambda_D(\alpha, p) = -m.$$

Moreover, if $(p, d) \neq (1, 1)$, *the infimum* (1.5) *is attained on* $(0, \alpha_\star(p))$, *and*

$$\forall \alpha \in (0, \alpha_\star(p)), \quad \Lambda_D(\alpha, p) = \lambda_D(V_{\alpha,p}),$$

where $V_{\alpha,p} = |\Psi|^{2/(p-1)}$, *and* $\Psi \in \mathrm{L}^2(\mathbb{R}^d, \mathbb{C}^N)$ *solves the* nonlinear Dirac equation

$$\mathcal{D}_m \Psi - |\Psi|^{2/(p-1)}\,\Psi = \Lambda_D(\alpha, p)\,\Psi \tag{1.6}$$

and satisfies the constraint $\int_{\mathbb{R}^d} |\Psi|^{2p/(p-1)}\,\mathrm{d}x = \|V_{\alpha,p}\|_p^p = \alpha^p$.

The proof of Theorem 1.1 is given in Section 3, and relies on the properties of the inverse map of $\alpha \mapsto \Lambda_D(\alpha, p)$ defined by

$$\alpha_D(\lambda, p) := \inf\left\{\|V\|_p \,:\, V \in \mathrm{L}^p(\mathbb{R}^d, \mathbb{R}^+) \text{ and } \lambda_D(V) = \lambda\right\}. \tag{1.7}$$

The critical value is $\alpha_\star(p) = \lim_{\lambda \to (-m)_+} \alpha_D(\lambda, p)$. It is such that

$$\lim_{\alpha \to \alpha_\star(p)_-} \lambda_D(V_{\alpha,p}) = -m,$$

and this limit is the upper bound of the lower essential spectrum $(-\infty, -m]$ or, equivalently, the lower end of the gap. For sake of simplicity, we adopt the convention that $\alpha_\star(p) = \alpha_D(-m, p)$. In the subcritical range of potentials, a simple consequence of Theorem 1.1 is the following *Keller–Lieb–Thirring estimate* for the Dirac operator $\mathcal{D}_m - V$.

Corollary 1.2. *Assume $p \geq d \geq 1$. For all $V \in \mathrm{L}^p(\mathbb{R}^d, \mathbb{R}^+)$ with $\|V\|_p < \alpha_\star(p)$, we have the optimal bound*

$$-m \leq \Lambda_D(\|V\|_p, p) \leq \lambda_D(V) \leq m. \tag{1.8}$$

If $(p, d) \neq (1, 1)$, then $V_{\alpha,p}$ as in Theorem 1.1 realizes the equality case, i.e., $\lambda_D(V_{\alpha,p}) = \Lambda_D(\alpha, p)$.

Some plots of $\alpha \mapsto \Lambda_D(\alpha, p)$ are displayed in Figure 1 (right).

The nonlinear Dirac equation (1.6) plays for the Dirac operator $\mathcal{D}_m - V$ the same role as (1.3) for the Schrödinger operator $-\Delta + V$. However, $\Lambda_D(\alpha, V)$ is not obtained as the infimum but as a critical point of a Rayleigh quotient with infinitely many negative directions corresponding to a min-max principle (see [28]), and for this reason there is no simple interpolation inequality such as (1.1) in the case the Dirac operator. A more involved functional inequality holds: see Appendix C.

Nonlinear Dirac equations have been introduced to model extended fermions, as effective operators for nonlinear effects in graphene-like materials or Bose–Einstein condensates: see Section 1.6 of [32] and the introduction of [5] for an overview of the literature. Since the spinors in the Dirac equation have at least two components, many types of nonlinearities can be considered (see, e.g., [59] and references therein) and give rise to various phenomena. For instance, localized solutions to a nonlinear equation of the form

$$\mathcal{D}_m \Psi - G(\Psi) = \lambda \Psi,$$

for some function $G\colon \mathbb{C}^N \mapsto \mathbb{C}^N$, correspond to solitary wave solutions to the time-dependent nonlinear Dirac equation, and have attracted considerable attention; see, for instance, [4, 15, 38, 56].

It is a common assumption to consider a nonlinearity that preserves Lorentz, or particle-hole, symmetry. Such a nonlinearity takes the form

$$\mathcal{D}_m \Psi - F(\langle \Psi, \beta \Psi \rangle_{\mathbb{C}^N})\, \beta \Psi = \lambda \Psi, \tag{1.9}$$

and is called the *Soler-type nonlinearity*. The Soler nonlinearity formally appears when minimizing the first positive eigenvalue of $\mathcal{D}_m - \beta V$, but will not be studied in this paper. In contrast, the nonlinearity that appears in (1.6) is of the form $F(\langle \Psi, \Psi \rangle_{\mathbb{C}^N})\, \Psi$, which is sometimes called a *Kerr-type nonlinearity* as in [5], apparently by extension of the cubic nonlinearity used in optics. Existence of localized solutions for (1.6) is studied in [5] in the critical exponent case $p = d = 2$, and in [8] in the critical exponent case $p = d$ for all dimensions $d \in \mathbb{N}$ with $m = 0$. Our results give an independent proof of the existence of a localized solution.

In [4, 38], the authors proved that equations of the form (1.9) have many solutions if $d \geq 2$ by looking for solutions of (1.9) in subspaces of fixed angular momentum. It seems that similar techniques could also be applied to (1.6). While it is reasonable to expect that the optimal potential is radially symmetric and the corresponding *ground state* Ψ is the solution with lowest positive angular momentum and smallest number of oscillations, this is so far an open question: see Appendix A. In Appendix B, we also give numerical results that point in this direction.

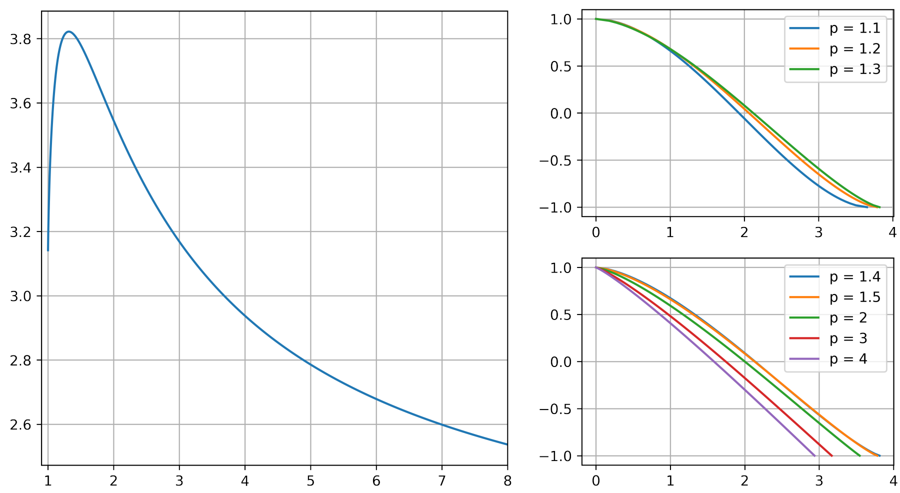

Figure 1. Let $d = 1$ and $m = 1$. (Left) The function $p \mapsto \alpha_\star(p)$, with maximum at $p \approx 1.32$, satisfies $\lim_{p\to 1_+} \alpha_\star(p) = \pi$ and $\lim_{p\to+\infty} \alpha_\star(p) = 2$. (Right) For various values of p, the maps $\alpha \mapsto \Lambda_D(\alpha, p)$ take value -1 at $\alpha = \alpha_\star(p)$. Upper (respectively, lower) right plots correspond to $p < 1.32$ (respectively, $p > 1.32$).

We now focus on the dimension $d = 1$. It turns out that one can completely solve (1.6) using special functions. Explicit formulae are given below, where B and ${}_2F_1$ respectively denote the Euler *Beta* function and the hypergeometric function.

Theorem 1.3. *Let $d = 1$ and $p \in (1, +\infty)$. For all $\lambda \in [-m, m]$, the equation*

$$\mathcal{D}_m \Psi - |\Psi|^{2/(p-1)}\, \Psi = \lambda \Psi, \quad \textit{with} \quad \mathcal{D}_m := \begin{pmatrix} m & \partial_x \\ -\partial_x & -m \end{pmatrix},$$

has a unique solution $\Psi \in L^2(\mathbb{R}, \mathbb{C}^2) \setminus \{0\}$, up to a phase factor and a translation. Up to a translation, $V = |\Psi|^{2/(p-1)}$ is even, decreasing on $\mathbb{R}^+$ and such that $\alpha_D(\lambda, p) = \|V\|_p$.

- Subcritical regime $\lambda > -m$. *With* $\mathsf{A} := \frac{p}{p-1}\,(m^2 - \lambda^2)$, $\mathsf{B} := \frac{2}{p-1}\,\sqrt{m^2 - \lambda^2}$ *and* $z_0 := \frac{m-\lambda}{m+\lambda}$, *we have*

$$\forall x \in \mathbb{R}, \quad V(x) = \frac{\mathsf{A}}{m \cosh(\mathsf{B}\, x) + \lambda} \tag{1.10}$$

  and

$$(\alpha_D(\lambda, p))^p = p^p \left(\tfrac{m+\lambda}{p-1}\right)^{p-1} z_0^{p-1/2}\, B\left(\tfrac{1}{2}, p\right)\, {}_2F_1\left(\tfrac{1}{2}, p; p + \tfrac{1}{2}; -z_0\right).$$

- Critical case $\lambda = -m$. *With $\zeta = 2m/(p-1)$, we have*

$$\forall x \in \mathbb{R}, \quad V(x) := \frac{\zeta\, p}{1 + \zeta^2 x^2} \tag{1.11}$$

  and

$$(\alpha_\star(p))^p = p^p \left(\tfrac{2m}{p-1}\right)^{p-1} B\left(\tfrac{1}{2}, p - \tfrac{1}{2}\right).$$

If $p = d = 1$*, then* $\alpha_D(\lambda, 1) = \arccos(\lambda/m)$ *and*

$$\lim_{\lambda \to (-m)_+} \alpha_D(\lambda, 1) = \lim_{p \to 1_+} \alpha_\star(p) = \pi.$$

See Figure 1. With the notations of Theorem 1.1 and $\alpha = \alpha_D(\lambda, p)$, up to translations, we know that $V = V_{\alpha,p}$ in (1.10) and (1.11). For the proof of Theorem 1.3 and some additional details, see Section 5.1. Formally, as $p \to 1_+$, the potential given by (1.10) converges to a *delta* Dirac distribution at $x = 0$ of mass $\arccos(\lambda/m)$ (see [65] for the study of self-adjoint extensions of $\not{D}_m - \alpha\delta_0$). A remarkable consequence of the estimate in the case $p = d = 1$ is the Keller–Lieb–Thirring inequality,

$$m \cos(\|V\|_1) \le \lambda_D(V) \le m \tag{1.12}$$

for any nonnegative potential $V \in \mathrm{L}^1(\mathbb{R})$ with $\|V\|_1 \le \pi$. See Appendix D for a result on optimality cases in the case $p = d = 1$, with a proof.

The case of Theorem 1.3 presents some similarities with the results of [35]: in the case $p = 1$, it is expected that optimality is achieved only by singular measures. Our goals differ from those of [35] as we adopt the point of view of functional interpolation inequalities with Keller-type estimates as a subproduct, while [35] is concerned with the issue of the optimal charge distribution for a Dirac–Coulomb equation. In terms of methods, there are many similarities since we use Birman–Schwinger reformulations as well as classical tools of the concentration-compactness method. However, there are also significant differences because requesting that the potential is in $\mathrm{L}^p(\mathbb{R}^d)$ means that the optimal V is obtained through a nonlinear Dirac equation which is not measure-valued as soon as $p > 1$.

Our results are not limited to estimates for the ground state and we also have a *Lieb–Thirring inequality* for the sum of eigenvalues *in the gap* $(-m, m)$ of Dirac operators of the form $\not{D}_m - V$ with $V \in \mathrm{L}^p(\mathbb{R}^d, \mathbb{R}^+)$. We denote by $-m < \lambda_1 \le \lambda_2 \le \cdots < m$ the possibly infinite sequence of eigenvalues in the gap $(-m, m)$, and write

$$e_k = e_k(m, V) := (m - \lambda_k) > 0,$$

so that $2m > e_1 \ge e_2 \ge \cdots > 0$. The quantity e_k is the distance between the eigenvalue λ_k and the bottom of the upper essential spectrum $+m$.

Theorem 1.4. *For all* $\gamma > d/2$ *and* $p \in (d, \gamma + d/2]$*, there is a constant* $L_{\gamma,d,p} > 0$ *so that, for all* $V \in \mathrm{L}^p(\mathbb{R}^d, \mathbb{R}^+)$*, and all* $m > 0$*, we have*

$$\sum_{k \ge 1} e_k^\gamma(m, V) \le L_{\gamma,d,p}\, m^{d/2} \int_{\mathbb{R}^d} V_m^{\gamma+d/2-p}\, V^p \,\mathrm{d}x, \quad \textit{with} \quad V_m := \min\{m, V\}. \tag{1.13}$$

If $V \in \mathrm{L}^p(\mathbb{R}^d, \mathbb{R}^+) \cap \mathrm{L}^{\gamma+d/2}(\mathbb{R}^d, \mathbb{R}^+)$, using the inequalities $V_m \le m$ and $V_m \le V$ gives, respectively,

$$\sum_{k \ge 1} e_k^\gamma \le L_{\gamma,d,p}\, m^{\gamma+d-p} \int_{\mathbb{R}^d} V^p \,\mathrm{d}x \quad \text{and} \quad \sum_{k \ge 1} e_k^\gamma \le L_{\gamma,d,p}\, m^{d/2} \int_{\mathbb{R}^d} V^{\gamma+d/2} \,\mathrm{d}x.$$

The inequality (1.13) is, in some sense, an interpolation between these two critical cases. In the proof, we use rough estimates: the method is constructive, but there is a lot of space for improving on the constant $L_{\gamma,d,p}$.

Structure of the paper. This paper is organized as follows. In Section 2, we establish some properties of the operator $\mathcal{D}_m - V$ with $V \in \mathrm{L}^p(\mathbb{R}^d)$: domain, associated Birman–Schwinger operator and self-adjointness. Section 3 is devoted to the variational problem associated with (1.5), after reformulation in the Birman–Schwinger framework. Theorem 3.1 is devoted to the existence of an optimal potential V by concentration-compactness methods (Section 3.2). The regularity of the optimizers is studied in Section 3.3. Section 4 is devoted to the proof of Theorem 1.4. Explicit and numerical computations are performed in Section 5 in dimensions $d = 1, 2$ and 3. Open questions, numerical observations, remarks on the non-relativistic limit and Gagliardo–Nirenberg–Sobolev inequalities, and a result in the case $p = d = 1$ are collected in Appendices A, B, C and D respectively.

2. Properties of Dirac operators

2.1. A self-adjoint realization

We assume that $V \in \mathrm{L}^p(\mathbb{R}^d, \mathbb{R}^+)$ is positive valued and deal with the self-adjoint extensions of $\mathcal{D}_m - V$.

Proposition 2.1. *Let $p \geq d \geq 1$ and $V \in \mathrm{L}^p(\mathbb{R}^d, \mathbb{R}^+)$. Then the operator $\mathcal{D}_m - V$ is self-adjoint with domain*

$$\mathrm{Dom}(\mathcal{D}_m - V) := \left\{\psi \in \mathrm{L}^2(\mathbb{R}^d, \mathbb{C}^N) : \sqrt{V}\psi,\ (\mathcal{D}_m - V)\,\psi \in \mathrm{L}^2(\mathbb{R}^d, \mathbb{C}^N)\right\}.$$

This is the unique self-adjoint realization satisfying

$$\mathrm{H}^1(\mathbb{R}^d, \mathbb{C}^N) \subseteq \mathrm{Dom}(\mathcal{D}_m - V) \subseteq \mathrm{H}^{1/2}(\mathbb{R}^d, \mathbb{C}^N).$$

Moreover, we have the following properties.

(i) *If p satisfies*

$$\begin{cases} p \geq 2 & \text{if } d = 1, \\ p > 2 & \text{if } d = 2, \\ p \geq d & \text{if } d \geq 3, \end{cases} \tag{2.1}$$

then $\mathrm{Dom}(\mathcal{D}_m - V) = \mathrm{H}^1(\mathbb{R}^d, \mathbb{C}^N)$.

(ii) *If $1 < p \leq 2$ and $d = 1$, then* $\mathrm{Dom}(\mathcal{D}_m - V)$ *is also included in* $\mathrm{H}^{3/2-1/p}(\mathbb{R}, \mathbb{C}^2)$*, hence in* $\mathrm{L}^\infty(\mathbb{R}, \mathbb{C}^2)$.

We call the extension of Proposition 2.1 the *distinguished* extension, because it is the unique one whose domain is included in the unperturbed form domain $\mathrm{Dom}(|\mathcal{D}_m|^{1/2}) = \mathrm{H}^{1/2}(\mathbb{R}^d, \mathbb{C}^N)$. We will consider only this extension in what follows, so that the operator $\mathcal{D}_m - V$ is self-adjoint under the condition $p \geq d \geq 1$. The proof of the first part of Proposition 2.1 follows from [57]. For completeness, we provide a short proof using the associated Birman–Schwinger operator. Under condition (2.1), the point (i) comes from the usual Kato–Rellich theorem [43, 62]. The result in (ii) is derived by bootstrapping the Sobolev embedding theorem. See Section 2.3 for the proof of Proposition 2.1.

Remark 2.2. For comparison, it is interesting to consider limit cases. The Coulomb potential $V(x) = 1/|x|$ in dimension $d = 3$ is in the weak Sobolev space $\mathrm{L}^3_w(\mathbb{R}^d)$. The operator $\mathcal{D}_m - \kappa V$ is essentially self-adjoint if $0 \le \kappa \le \sqrt{3}/2$, it has a distinguished extension if $\sqrt{3}/2 < \kappa \le 1$, and no distinguished self-adjoint extension if $\kappa > 1$; see [13] and references therein. Also see Remark 2.5.

2.2. The Birman–Schwinger operator

The Birman–Schwinger operator is a powerful tool for analyzing the spectral properties of $\mathcal{D}_m - V$ when V belongs to a large class of perturbations. In the relativistic case, Klaus in [46] used it extensively to characterize and study the first eigenvalue of Dirac operators when proving the existence of a distinguished self-adjoint extension. For non-Hermitian potentials V, it can be employed to locate the eigenvalues of $\mathcal{D}_m - V$, as shown for example by Cuenin, Laptev and Tretter in [18], and by Fanelli and Krejčiřík in [39]. Furthermore, it can be applied to discuss properties of the ground state of $\mathcal{D}_m - V$ when V is a generalized Coulomb-type potential, see, e.g., [14,34,35,46]. Throughout this paper, following the approach by Kato [44] and by Konno and Kuroda [48], the Birman–Schwinger operator is used to define the self-adjoint extension of the operator $\mathcal{D}_m - V$. Then, with this rigorous definition at hand, we prove the existence of the optimization problem which defines the *ground state* by applying variational methods directly on the Birman–Schwinger reformulation of the problem.

For $z \notin \sigma(\mathcal{D}_m)$, let

$$R_0(z) := (\mathcal{D}_m - z)^{-1} \tag{2.2}$$

denote the resolvent operator. Recall that we assume $V \ge 0$. We introduce the *Birman–Schwinger operator*

$$K_V(z) := \sqrt{V}\, R_0(z)\, \sqrt{V} = \sqrt{V}\, \frac{1}{\mathcal{D}_m - z}\, \sqrt{V}. \tag{2.3}$$

Lemma 2.3. *For all $p \ge d \ge 1$, all $V \in \mathrm{L}^p(\mathbb{R}^d, \mathbb{R}^+)$ and all $z \notin \sigma(\mathcal{D}_m)$, the operator $K_V(z)$ is compact (hence bounded). In addition,*

$$\lim_{s \to \pm\infty} \|K_V(\mathrm{i}s)\|_{\mathrm{op}} = 0.$$

This result follows from the *Kato–Seiler–Simon inequality*: see proof in Section 2.3. A consequence of Lemma 2.3 is the following result (also see Section 2.3 for its proof).

Proposition 2.4. *Let $\mathcal{D}_m - V$ be the distinguished self-adjoint extension defined as in Proposition 2.1. Then*

$$\sigma_{\mathrm{ess}}(\mathcal{D}_m - V) = \sigma_{\mathrm{ess}}(\mathcal{D}_m) = (-\infty, -m] \cup [m, +\infty).$$

Moreover, the Birman–Schwinger principle *holds: for all $\lambda \in (-m, m)$, λ is an eigenvalue of $\mathcal{D}_m - V$ if and only if 1 is an eigenvalue of $K_V(\lambda)$.*

Let us point out some differences with Birman–Schwinger operators associated with Schrödinger operators (see Figure 2). In the Schrödinger case, the Birman–Schwinger operator is of the form

$$\widetilde{K}_V(\lambda) = \sqrt{V}\,\frac{1}{-\Delta-\lambda}\,\sqrt{V}.$$

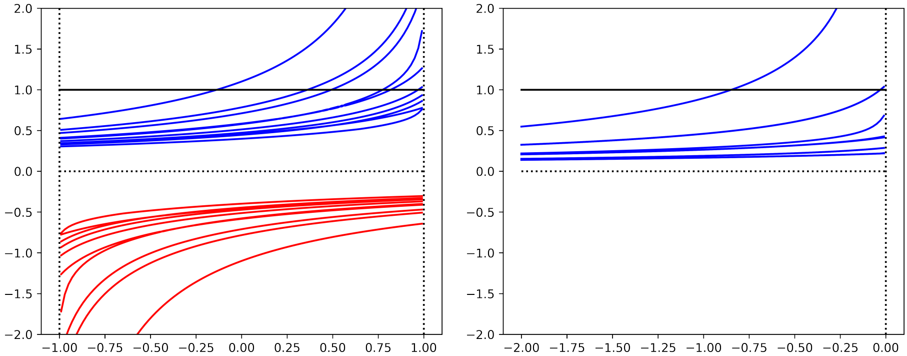

Figure 2. *The Birman–Schwinger principle*. (Left). The spectrum of $\lambda \mapsto K_V(\lambda)$ (Dirac case) for $\lambda \in (-1,1)$, and $V(x) = 2\exp(-|x|^2/4)$ in dimension $d = 2$. We only plotted the 10 largest (blue) and the 10 lowest (red) eigenvalues. An energy λ is an eigenvalue of $\mathcal{D}_m - V$ if one eigenvalue of $K_V(\lambda)$ crosses the black line 1. (Right) Same for $\lambda \mapsto \widetilde{K}_V(\lambda)$ (Schrödinger case) with $\lambda \in (-2,0)$.

For any $\lambda < 0$, the operator $\widetilde{K}_V(\lambda)$ is a positive compact operator and the map $\lambda \mapsto \widetilde{K}_V(\lambda)$ is operator increasing on $\mathbb{R}^-$. In particular, if $\widetilde{\mu}_1(\lambda) > \widetilde{\mu}_2(\lambda) \geq \cdots \geq 0$ denote the eigenvalues of $\widetilde{K}_V(\lambda)$, ranked in decreasing order and counted with multiplicities, all functions $\lambda \mapsto \widetilde{\mu}_j(\lambda)$ are increasing on $\mathbb{R}^-$. In addition, the first eigenvalue $\widetilde{\mu}_1$ is simple because the kernel $\widetilde{K}_V(x,y)$ is pointwise positive, together with the Krein–Rutman theorem; see Theorem 6.13 in [11] for a statement, and also Section XIII.12 of [61].

In the Dirac case, the operator $K_V(\lambda)$ with $\lambda \in \mathbb{R}$ is defined only in the gap $(-m,m)$ of the essential spectrum. It is compact by Lemma 2.3, and symmetric because λ is real, but it is *not* a positive operator. Its eigenvalues are real valued, and can be ranked as $\mu_1(\lambda) \geq \mu_2(\lambda) \geq \cdots \geq 0$ for the positive eigenvalues, and $\nu_1(\lambda) \leq \nu_2(\lambda) \leq \cdots \leq 0$ for the negative ones. As the map $\lambda \mapsto (\mathcal{D}_m - \lambda)^{-1}$ is operator increasing on $(-m,m) \ni \lambda$, all maps $\lambda \mapsto \mu_j(\lambda)$ and $\lambda \mapsto \nu_j(\lambda)$ are increasing. This explains in particular why we expect eigenvalues to emerge from the upper essential spectrum in this setting. We do not know whether $\mu_1(\lambda)$ is always a simple eigenvalue or not (see Appendix A for more details on open questions).

For $\lambda \in (-m,m)$, $p \geq d \geq 1$, and $V \in \mathrm{L}^p(\mathbb{R}^d,\mathbb{R}^+)$, let $\mu_1(K_V(\lambda))$ denote the largest (positive) eigenvalue of $K_V(\lambda)$. We rephrase the optimization problem (1.7) as

$$\alpha_D(\lambda,p) := \inf\big\{\|V\|_p \,:\, V \in \mathrm{L}^p(\mathbb{R}^d,\mathbb{R}^+) \text{ and } \mu_1(K_V(\lambda)) = 1\big\}. \tag{2.4}$$

2.3. Proofs of Proposition 2.1, Lemma 2.3 and Proposition 2.4

We start by establishing that K_V defined by (2.3) is a compact operator (Lemma 2.3) before proving Propositions 2.1 and 2.4.

Proof of Lemma 2.3. Assume first that $p > d$. We claim that, for $z \notin \sigma(\not{D}_m)$, the operator $K_V(z)$ is compact. We have $R_0(z) = g_z(-\mathrm{i}\nabla)$, with

$$g_z(k) := \frac{1}{|k|^2 + m^2 - z^2}\Big(\sum_{j=1}^{d} \alpha_j\, k_j + m\,\beta + z\,\mathbb{I}_N\Big) =: \sum_{j=1}^{d} g_z^j(k) + g_z^m(k) + g_z^z(k),$$

with obvious notation. Let us focus on the $g_z^1(k)$ term. We write $g_z^1(k) = g_z^A(k)\, g_z^B(k)$, with

$$g_z^A(k) := \frac{\sqrt{|k_1|}\,\mathrm{sgn}(k_1)}{\sqrt{|k|^2 + m^2 - z^2}}\,\alpha_1 \quad \text{and} \quad g_z^B(k) := \frac{\sqrt{|k_1|}}{\sqrt{|k|^2 + m^2 - z^2}}\,\mathbb{I}_d.$$

All components of the functions g_z^A and g_z^B are in $\mathrm{L}^q(\mathbb{R}^d)$ for all $q > 2d$. Since $\sqrt{V}$ is in $\mathrm{L}^{2p}(\mathbb{R}^d)$ with $2p > 2d$, we can use the Kato–Seiler–Simon inequality (see Theorem 4.1 in Chapter 4 of [66]) and conclude that the operator

$$K_V^1(z) := \sqrt{V}\, g_z^1(-\mathrm{i}\nabla)\,\sqrt{V}$$

is in the Schatten class $\mathfrak{S}_p(\mathrm{L}^2(\mathbb{R}^d, \mathbb{C}^N))$, with

$$\|K_V^1(z)\|_{\mathfrak{S}_p} \le \big\|\sqrt{V(x)}\, g_z^A(-\mathrm{i}\nabla)\big\|_{\mathfrak{S}_{2p}}\, \big\|g_z^B(-\mathrm{i}\nabla)\,\sqrt{V(x)}\,\big\|_{\mathfrak{S}_{2p}} \le C\,\|V\|_p\,\|g_z^1\|_p.$$

In addition, we have $\|g_{z=\mathrm{i}s}^1\|_p \to 0$ as $s \to \pm\infty$. Similar computations for the other terms show that K_V is in the Schatten class $\mathfrak{S}_p$, and that $\lim_{s\to\pm\infty}\|K_V(\mathrm{i}s)\|_{\mathrm{op}} = 0$.

In the case $p = d$ with $d \ge 2$, we use that all components of the functions g_z^A and g_z^B are in the weak-Sobolev space $\mathrm{L}_w^{2d}(\mathbb{R}^d)$. According to [66], Chapter 4, $g_z^A(-\mathrm{i}\nabla)\sqrt{V}$ and $\sqrt{V}\, g_z^B(-\mathrm{i}\nabla)$ are in the weak Schatten class $\mathfrak{S}_{2d,w}$. In particular, they are compact operators. This already proves that K_V is compact as well. Note that $\|g_{z=\mathrm{i}s}^1\|_{d,w}$ does not converge to 0 as $s \to \pm\infty$.

For any $R > 0$, $\sqrt{V_R} := \min(\sqrt{V}, R)$ belongs to $\mathrm{L}^d(\mathbb{R}^d) \cap \mathrm{L}^\infty(\mathbb{R}^d)$. We have

$$\big\|g_z^A(-\mathrm{i}\nabla)\,\sqrt{V}\big\|_{\mathrm{op}} \le \big\|g_z^A(-\mathrm{i}\nabla)\,\sqrt{V_R}\big\|_{\mathrm{op}} + \big\|g_z^A(-\mathrm{i}\nabla)\,(\sqrt{V} - \sqrt{V_R})\big\|_{\mathrm{op}}.$$

For R large enough, the second term is small in the Schatten space $\mathfrak{S}_{2d,w}$, and for $z = \mathrm{i}s$ with $|s|$ large, the first term is small in $\mathfrak{S}_q$ with $q > p$. Hence $\lim_{s\to\pm\infty}\|K_V(\mathrm{i}s)\|_{\mathrm{op}} = 0$.

Let us finally assume $p = d = 1$. In this case, with explicit computations, the kernel of the Birman–Schwinger operator $K_V(z)$ is given by

$$\sqrt{V(x)}\cdot\frac{1}{2}\begin{pmatrix}(z+m)/k & \mathrm{sign}(x-y)\\ \mathrm{sign}(x-y) & (z-m)/k\end{pmatrix}\mathrm{e}^{-k\,|x-y|}\cdot\sqrt{V(y)} \in \mathrm{L}^2(\mathbb{R}\times\mathbb{R}, \mathbb{C}^2)$$

where $k = \sqrt{m^2 - z^2}$ is chosen with a positive real part. Thus, K_V is a Hilbert–Schmidt operator (hence it is compact), and by the dominated convergence theorem we can conclude that $\lim_{s\to\pm\infty}\|K_V(\mathrm{i}s)\|_{\mathrm{op}} = 0$. ∎

Proof of Proposition 2.1. We divide the proof in several steps.

(a) *Distinguished self-adjoint extension.*

We define the domain of self-adjointness for the operator $\mathcal{D}_m - V$ as a perturbation of $\mathcal{D}_m$ by applying the method of G. Nenciu in [57]. Using similar techniques as in the proof of Lemma 2.3, one can show that the operators $R_0(z)\sqrt{V}$ and $\sqrt{V}R_0(z)$ can be extended into bounded linear operators on $\mathrm{L}^2(\mathbb{R}^d, \mathbb{C}^N)$. These operators are compact operators, in the Schatten class $\mathfrak{S}_{2p}$. We are now in the setting of [57]. Let $\Omega := \{z \in \mathbb{C} : 1 \notin \sigma(K_V(z))\}$, where K_V is defined by (2.3). The set Ω is non-empty by Lemma 2.3. For $z \in \Omega$, define

$$R(z) := R_0(z) + R_0(z)\sqrt{V}\,(1 - K_V(z))^{-1}\sqrt{V}\,R_0(z).$$

According to [57], the operator $\mathcal{D}_m - V$ has a unique self-adjoint extension whose resolvent is the operator $R(z)$ defined in (2.2). Its domain is $\mathrm{Dom}(\mathcal{D}_m - V) := \mathrm{Ran}\, R(z)$, which is independent of $z \in \Omega$. This is the unique extension which is included in the unperturbed form domain $\mathrm{Dom}(|\mathcal{D}_m|^{1/2}) = \mathrm{H}^{1/2}(\mathbb{R}^d, \mathbb{C}^N)$.

(b) *Domain of the distinguished extension.*

Define the *maximal domain* as

$$\mathrm{Dom}_{\max}(\mathcal{D}_m - V) := \big\{\psi \in \mathrm{L}^2(\mathbb{R}^d, \mathbb{C}^N) : (\mathcal{D}_m - V)\,\psi \in \mathrm{L}^2(\mathbb{R}^d, \mathbb{C}^N)\big\}.$$

Then, the set

$$\big\{\psi \in \mathrm{L}^2(\mathbb{R}^d, \mathbb{C}^N) : \sqrt{V}\,\psi,\ (\mathcal{D}_m - V)\psi \in \mathrm{L}^2(\mathbb{R}^d, \mathbb{C}^N)\big\}$$

is also $\mathrm{Dom}_{\max}(\mathcal{D}_m - V) \cap \mathrm{Dom}(\sqrt{V})$. We write $\psi \in \mathrm{Dom}(\mathcal{D}_m - V)$ as $R(z)f$ for some $f \in \mathrm{L}^2(\mathbb{R}^d, \mathbb{C}^N)$. Then

$$\sqrt{V}\,\psi = \sqrt{V}\,R_0(z)f + K_V\,(1 - K_V(z))^{-1}\sqrt{V}\,R_0(z)f \in \mathrm{L}^2(\mathbb{R}^d, \mathbb{C}^N).$$

This proves that $\mathrm{Dom}(\mathcal{D}_m - V) \subset \mathrm{Dom}_{\max}(\mathcal{D}_m - V) \cap \mathrm{Dom}(\sqrt{V})$. For the opposite inclusion, consider $\psi \in \mathrm{Dom}_{\max}(\mathcal{D}_m - V) \cap \mathrm{Dom}(\sqrt{V})$. We set $f := (\mathcal{D}_m + V - z)\,\psi \in \mathrm{L}^2(\mathbb{R}^d, \mathbb{C}^N)$ and $\psi_0 = R(z)f \in \mathrm{Dom}(\mathcal{D}_m - V)$. Note that $\sqrt{V}\,\psi_0 \in \mathrm{L}^2(\mathbb{R}^d, \mathbb{C}^N)$, since

$$\sqrt{V}\,R(z) = \sqrt{V}\,R_0(z) + K_V(z)\,(1 - K_V(z))^{-1}\sqrt{V}\,R_0(z)$$

is a bounded operator on $\mathrm{L}^2(\mathbb{R}^d, \mathbb{C}^N)$. So $\phi := \psi - \psi_0$ is such that

$$(\mathcal{D}_m - V - z)\,\phi = 0 \quad \text{and} \quad \phi \in \mathrm{Dom}(\sqrt{V}).$$

From the relation

$$(1 - R_0(z))^{-1}\sqrt{V}\,R_0(z)\,(\mathcal{D}_m - V - z) = \sqrt{V},$$

we obtain $\sqrt{V}\,\phi = 0$, and from the relation

$$R_0(z)\,(\mathcal{D}_m - V - z) = 1 - R_0(z)\,V = 1 - (R_0(z)\sqrt{V}\,)\sqrt{V},$$

we finally get $\phi = 0$, hence $\psi = \psi_0 \in \mathrm{Dom}(\mathcal{D}_m - V)$.

Finally, if $p \geq d \geq 1$ by the Hölder inequality and the Sobolev embedding theorem we get that $\mathrm{H}^1(\mathbb{R}^d, \mathbb{C}^N) \subseteq \mathrm{Dom}(\mathcal{D}_m - V)$, and this concludes the first part of the proof.

(c) *Self-adjointness on* $\mathrm{H}^1(\mathbb{R}^d, \mathbb{C}^N)$.

Let us prove (i). Assume that p satisfies (2.1). Thanks to (1.4), we have

$$\forall \psi \in \mathrm{H}^2(\mathbb{R}^d, \mathbb{C}^N), \quad \|\mathcal{D}_m \psi\|_2^2 = \|\nabla \psi\|_2^2 + m^2 \|\psi\|_2^2.$$

This shows that the graph norm of $\mathcal{D}_m$ is equivalent to the usual $\mathrm{H}^1(\mathbb{R}^d, \mathbb{C}^N)$ norm. Set

$$q := \begin{cases} 2p/(p-2) & \text{if } p > 2, \\ +\infty & \text{if } d = 1 \text{ and } p = 2, \end{cases} \quad \text{so that} \quad \frac{1}{p} + \frac{1}{q} = \frac{1}{2}.$$

We write $V = V_1 + V_2$ with $V_1 := V \mathbb{1}_{V \geq R}$ and $V_2 := V \mathbb{1}_{V \leq R}$. We have

$$\begin{aligned} \|V \psi\|_2 &\leq \|V_1 \psi\|_2 + \|V_2 \psi\|_2 \leq \|V_1\|_p \|\psi\|_q + \|V_2\|_\infty \|\psi\|_2 \\ &\leq C_S \|V_1\|_p \|\psi\|_{\mathrm{H}^1} + \|V_2\|_\infty \|\psi\|_2, \end{aligned}$$

where, in the last inequality, we used Sobolev's embedding $\mathrm{H}^1(\mathbb{R}^d) \hookrightarrow \mathrm{L}^q(\mathbb{R}^d)$ and that, according to (2.1), q satisfies

$$\begin{cases} 2 \leq q \leq +\infty & \text{if } d = 1, \\ 2 \leq q < +\infty & \text{if } d = 2, \\ 2 \leq q \leq 2d/(d-2) & \text{if } d \geq 3. \end{cases}$$

We choose R large enough so that $C_S \|V_1\|_p < 1$ and conclude with the *Kato–Rellich theorem* (see Theorem X.12 in [60]) that $\mathcal{D}_m - V$ is self-adjoint with domain $\mathrm{H}^1(\mathbb{R}^d, \mathbb{C}^N)$. Since any self-adjoint operator only admits trivial self-adjoint extensions, we can conclude that $\mathrm{H}^1(\mathbb{R}^d, \mathbb{C}^N) = \mathrm{Dom}(\mathcal{D}_m - V)$.

(d) *Regularity for* $d = 1$.

We now focus on (ii) and assume that $d = 1$ and $1 < p \leq 2$. Let us prove that $\mathrm{Dom}(\mathcal{D}_m - V)$ is also included in $\mathrm{H}^{3/2-1/p}(\mathbb{R}, \mathbb{C}^2)$. For any $\psi \in \mathrm{Dom}(\mathcal{D}_m - V)$, we have

$$(\mathcal{D}_m - V)\psi =: f \in \mathrm{L}^2(\mathbb{R}, \mathbb{C}^2),$$

hence $\mathcal{D}_m \psi = f + V\psi$. We recall the following negative Sobolev embeddings: for all $1 < r \leq 2$, we have $\mathrm{L}^r(\mathbb{R}) \hookrightarrow \mathrm{H}^{-s}(\mathbb{R})$ for all $s \geq (2-r)/(2r)$ and $\mathrm{L}^2(\mathbb{R}) \hookrightarrow \mathrm{H}^{-s}(\mathbb{R})$ for all $s \geq 0$, while

$$\forall s \geq 1/(2p), \quad V\psi = \sqrt{V}\, \underbrace{(\sqrt{V}\psi)}_{\in \mathrm{L}^2(\mathbb{R}, \mathbb{C}^2)} \in \mathrm{L}^{2p/(p+1)}(\mathbb{R}, \mathbb{C}^2) \hookrightarrow \mathrm{H}^{-s}(\mathbb{R}, \mathbb{C}^2).$$

We deduce that $\mathcal{D}_m \psi \in \mathrm{H}^{-1/(2p)}(\mathbb{R}, \mathbb{C}^2)$, hence that $\psi \in \mathrm{H}^{1-1/(2p)}(\mathbb{R}, \mathbb{C}^2)$. We now bootstrap the argument. For $p > 1$, we have $1 - 1/(2p) > 1/2$, so $\psi \in \mathrm{L}^\infty(\mathbb{R}, \mathbb{C}^2)$ by Sobolev's embedding. This gives $V\psi \in \mathrm{L}^p(\mathbb{R}, \mathbb{C}^2) \hookrightarrow \mathrm{H}^{(p-2)/(2p)}(\mathbb{R}, \mathbb{C}^2)$. So $\mathcal{D}_m \psi = f - V\psi \in \mathrm{H}^{(p-2)/(2p)}(\mathbb{R}, \mathbb{C}^2)$ as well, and we obtain $\psi \in \mathrm{H}^{1+(p-2)/(2p)}(\mathbb{R}, \mathbb{C}^2)$, with $1 + (p-2)/(2p) = 3/2 - 1/p$, as wanted. ∎

Proof of Proposition 2.4. Since $R_0(z)\sqrt{V}$ is compact, then $R(z)$ is a compact perturbation of the free resolvent $R_0(z)$. The result on $\sigma_{\text{ess}}(\mathcal{D}_m - V)$ follows from Theorem 4.5 in [68] (also see Theorem XIII.14 and Corollary 1 in [61]). Such a result is known in the literature as *Weyl's theorem*.

By construction, the Birman–Schwinger principle holds for the distinguished self-adjoint extension defined as in Proposition 2.1: $\lambda \in (-m, m)$ is an eigenvalue of $\mathcal{D}_m - V$ if and only if 1 is an eigenvalue of $K_V(\lambda)$. See Theorem 1.3 in [2] for a similar application of the Birman–Schwinger principle in a non-relativistic setting. ∎

Remark 2.5. We point out that the self-adjointness of Dirac operators involving potentials with one Coulomb singularity or several Coulomb singularities has been intensively studied in respectively [3, 44, 57, 64, 69–71] (with additional references therein) and [46, 58]. In the alternative strategy of [36, 37] based on [28], a distinguished self-adjoint extension is built using the underlying Hardy inequality, which was related with the other constructions for Dirac–Coulomb operators in [33, 34]. Also see [31, 63] for further considerations on min-max principles, Hardy inequalities and self-adjointness issues. Optimal Hardy inequalities have been repeatedly use to establish optimal conditions for the existence of a ground state. For instance, in presence of a magnetic field as in [20, 25, 26], a critical magnetic field is obtained as the ground state energy approaches $-mc^2$, which determines the optimal constant of the corresponding Hardy inequality. In the approach of [34, 35], as well as in our paper, the Birman–Schwinger formula is essential as it was in [46, 47, 57]. Notice that we do not rely on Nenciu's method, see Corollary 2.1 in [57], but instead use the method of Konno and Kuroda [48] and Kato's approach [44].

3. The variational problem

In this section, we consider the minimization problem (2.4) and prove Theorem 1.1 in a reformulation which relies on the Birman–Schwinger operator associated to $\mathcal{D}_m - V$, as introduced in Section 2.2. The proof of Theorem 1.1 is given below, right after the statement of Corollary 3.2, as a simple consequence of previous results in the Birman–Schwinger framework.

3.1. An auxiliary maximization problem

First, we notice that, for all $t > 0$, we have $K_{tV}(\lambda) = tK_V(\lambda)$, hence $\mu_1(K_{tV}(\lambda)) = t\mu_1(K_V(\lambda))$. So, introducing the auxiliary problem

$$\mathcal{N}(\lambda, p) := \sup\big\{\mu_1(K_W(\lambda)) : W \in \mathrm{L}^p(\mathbb{R}^d, \mathbb{R}^+),\ \|W\|_p = 1\big\}, \tag{3.1}$$

we deduce that

$$\alpha_D(\lambda, p) = \frac{1}{\mathcal{N}(\lambda, p)}. \tag{3.2}$$

If W is a maximizer for $\mathcal{N}(\lambda, p)$, then $V = W/\mathcal{N}(\lambda, p)$ is a minimizer for $\alpha_D(\lambda, p)$. In what follows, we study the maximization problem (3.1). We perform several changes of variables to study this problem. First, the min-max principle shows that $\mathcal{N}(\lambda, p)$ equals

$$\mathcal{N}(\lambda,p)=\sup_{\substack{W\in \mathrm{L}^p(\mathbb{R}^d,\mathbb{R}^+)\\ \|W\|_p=1}}\ \sup_{\substack{\phi\in \mathrm{L}^2(\mathbb{R}^d,\mathbb{C}^N)\\ \|\phi\|_2=1}} \langle \phi, \sqrt{W}\, R_0(\lambda)\sqrt{W}\phi\rangle.$$

We make the change of variable

$$w := \sqrt{W}\,\phi$$

so that, by Hölder's inequality, $w\in \mathrm{L}^{2p/(p+1)}(\mathbb{R}^d,\mathbb{C}^N)$, and with the convention that $\|w\|_r=\||w|_{\mathbb{C}^N}\|_r$, we have

$$\|w\|_{2p/(p+1)} \le \|W\|_p^{1/2}\,\|\phi\|_2 = 1.$$

In addition, there is equality if and only if W^p is proportional to $|\phi|^2$, both proportional to $|w|^{2p/(p+1)}$. With

$$q := \frac{2p}{p+1}\in(1,2),$$

this shows that $\mathcal{N}(\lambda, p)$ is also solution to the optimization problem

$$\mathcal{N}(\lambda,p)=\sup\big\{\langle w, R_0(\lambda)\, w\rangle \,:\, w\in \mathrm{L}^q(\mathbb{R}^d,\mathbb{C}^N),\ \|w\|_q=1\big\}. \tag{3.3}$$

In addition, if $w\in \mathrm{L}^q(\mathbb{R}^d,\mathbb{R}^+)$ is an optimizer of (3.3), then the corresponding optimal W and ϕ are given by

$$W=|w|^{q/p}=|w|^{2/(p+1)} \quad\text{and}\quad \phi=|w|^{q/2-1}\,w=|w|^{-1/(p+1)}\,w.$$

Thus, by showing the existence of an optimizer for (3.3), we solve problem (3.1), and by definition of the Birman–Schwinger operator, find an optimal potential and eigenfunction for our original problem (1.7).

Since $\alpha\mapsto\Lambda_D(\alpha,p)$ is the inverse map of $\lambda\mapsto\alpha_D(\lambda,p)$ according to (3.2), and since $\alpha_D(\lambda,p)=1/\mathcal{N}(\lambda,p)$, it is enough to focus on the properties of $\mathcal{N}(\cdot,p)$.

Theorem 3.1. *Let us consider $\mathcal{N}$ defined by (3.3). For all $\lambda\in(-m,m)$ and all $p>d$, we have $\mathcal{N}(\lambda,p)>0$. All maximizing sequences for (3.3) are precompact up to translations, hence (3.3) has maximizers. If w is such an optimizer, then w satisfies the Euler–Lagrange equation*

$$R_0(\lambda)\,w=\tau\,|w|^{-2/(p+1)}\,w, \quad \text{with } \tau=\mathcal{N}(\lambda,p). \tag{3.4}$$

Finally, the map $\lambda\mapsto\mathcal{N}(\lambda,p)$ is continuous, strictly increasing, and satisfies

$$\lim_{\lambda\to-m}\mathcal{N}(\lambda,p)=:\mathcal{N}_c(p)>0 \quad\text{and}\quad \lim_{\lambda\to+m}\mathcal{N}(\lambda,p)=\infty.$$

The proof of the first part relies on the profile decomposition method (concentration-compactness) used by Lions [54], and is given in the next section. Theorem 3.1 implies the existence of an optimal potential and an optimal spinor.

Corollary 3.2. *Under the assumptions of Theorem 1.1, the infimum (2.4) is attained for any $\lambda \in (-m, m)$ by a potential $V = |\Psi|^{2/(p-1)}$, where $\Psi \in L^2(\mathbb{R}^d, \mathbb{R}^N)$ solves the* nonlinear Dirac equation

$$\mathcal{D}_m \Psi - |\Psi|^{2/(p-1)}\, \Psi = \lambda\, \Psi, \tag{3.5}$$

such that $\lambda_D(V) = \lambda$ and

$$\Big(\int_{\mathbb{R}^d} |\Psi|^{2p/(p-1)}\, \mathrm{d}x \Big)^{1/p} = \|V\|_p = \alpha_D(\lambda, p) = \frac{1}{\mathcal{N}(\lambda, p)}.$$

Proof of Theorem 1.1. Our main Theorem 1.1 is a direct consequence of Theorem 3.1 and Corollary 3.2. Since $\mathcal{N}_c(p) > 0$, we have indeed $\alpha_c(p) := 1/\mathcal{N}_c(p) < \infty$. ∎

Proof of Corollary 3.2. First, we translate the Euler–Lagrange equation for w into an equation for the potential V and an eigenfunction (not normalized) Ψ. We set

$$\Psi = \tau^{(1-p)/2}\, |w|^{-2/(p+1)}\, w, \quad \text{so that} \quad w = \tau^{(p+1)/2}\, |\Psi|^{2/(p-1)}\, \Psi.$$

Applying $\mathcal{D}_m - \lambda$ to (3.4) shows that Ψ satisfies the nonlinear Dirac equation (3.5). The optimal potential W for the $\mathcal{N}(\lambda, p)$ problem in (3.1) is $W = |w|^{2/(p+1)} = \tau\, |\Psi|^{2/(p-1)}$, and finally, the optimal potential V for the $\alpha_D(\lambda, p)$ problem is, as wanted,

$$V = \frac{W}{\mathcal{N}(\lambda, p)} = |\Psi|^{2/(p-1)}.$$

We recover the value of $\mathcal{N}(\lambda, p)$ and $\alpha_D(\lambda, p)$ from the solution Ψ because

$$\int_{\mathbb{R}^d} |\Psi|^{2p/(p-1)}\, \mathrm{d}x = \tau^{-p} \int_{\mathbb{R}^d} |w|^{2p/(p+1)}\, \mathrm{d}x = \tau^{-p} = \mathcal{N}(\lambda, p)^{-p} = \alpha_D(\lambda, p)^p.$$

Among all solutions of (3.5), Ψ is the one with the smallest $L^{2p/(p-1)}(\mathbb{R}^d, \mathbb{C}^N)$ norm so that $\lambda = \Lambda_D(\alpha, p)$, and Ψ actually solves (1.6). ∎

3.2. Proof of Theorem 3.1

We now prove Theorem 3.1. We consider a more general case, and study a general optimization problem. In what follows, we use the notation

$$\langle w, K * w\rangle := \iint_{\mathbb{R}^d \times \mathbb{R}^d} \langle w(x), K(x-y)\, w(y)\rangle_{\mathbb{C}^N}\, \mathrm{d}x\, \mathrm{d}y,$$

and define for any $s > 0$ the maximization problem

$$J(s) := \sup \Big\{ \langle w, K * w\rangle \,:\, w \in L^q(\mathbb{R}^d, \mathbb{C}^N),\ \int_{\mathbb{R}^d} |w|^q\, \mathrm{d}x = s \Big\}. \tag{3.6}$$

Here, K is a convolution operator, or equivalently a multiplication operator in Fourier space. In our case, $K(x-y) = R_0(\lambda)(x-y)$ is the kernel of the Dirac resolvent, but we state a more general result.

Lemma 3.3. *Let* $q \in (1,2)$, *set* $q' := q/(q-1) \in (2,+\infty)$ *and* $r := q'/2 \in (1,+\infty)$. *Let* $K\colon \mathbb{R}^d \to \mathcal{M}_N(\mathbb{C})$ *be a matrix-valued function satisfying* $K(x) = K(-x)^*$, *and such that one of the two properties holds*:

(i) *either* $K \in \mathrm{L}^r(\mathbb{R}^d, \mathcal{M}_N(\mathbb{C}))$,

(ii) *or* $K = R_0(\lambda)$ *is a Dirac resolvent for some* $\lambda \in (-m, m)$.

Then the map $w \mapsto \langle w, K * w\rangle$ *is well-defined on* $\mathrm{L}^q(\mathbb{R}^d, \mathbb{C}^N)$ *and real valued. Moreover, if* $J(1) > 0$, *then* (3.6) *admits maximizers.*

Before proving this result, we make several remarks.

Remark 3.4. Lemma 3.3 fails at the endpoint $q = 2$. Indeed, by applying the Fourier transform, we have

$$\langle w, K * w\rangle = \int_{\mathbb{R}^d} \langle \widehat{w}(k), \widehat{K}(k)\, \widehat{w}(k)\rangle_{\mathbb{C}^N}\, \mathrm{d}k.$$

This means that all optimizing sequences must concentrate on Dirac masses in Fourier space at locations where $k \mapsto \sup\mathrm{spec}(\widehat{K}(k))$ has maxima. Since the Fourier transform is an isometry on $\mathrm{L}^2(\mathbb{R}^d)$, we deduce that the maximization problem has no maximum in general. The same argument shows that the existence of optimizers is closely related to the fact that the Fourier transform is not a bijection between $\mathrm{L}^q(\mathbb{R}^d)$ and $\mathrm{L}^{q'}(\mathbb{R}^d)$ if $1 < q < 2$.

Remark 3.5. In the case of the Dirac operator, one has an explicit expression for $K = R_0(\lambda)$, the fundamental solution of $\mathcal{D}_m - \lambda$. Using that

$$(\mathcal{D}_m - \lambda)^{-1} = (\mathcal{D}_m + \lambda)\, \frac{1}{-\Delta + m^2 - \lambda^2},$$

we first deduce that $R_0(\lambda)(\cdot)$ is the Fourier transform of

$$g_\lambda(k) = \Big(\sum_{j=1}^{d} \alpha_j\, k_j + m\,\beta + \lambda\,\mathbb{I}_N\Big)\, \frac{1}{k^2 + m^2 - \lambda^2}.$$

The function $k \mapsto g_\lambda(k)$ is analytic on $\mathbb{R}^d$ because there is no singularity in the denominator since $|\lambda| < m$, so its Fourier transform is exponentially decaying in x. Actually, we have

$$R_0(\lambda)(x) = \frac{c_{d,\lambda}}{|x|^{d/2-1}} \Big(\mathrm{i} \sum_{j=1}^{d} \alpha_j\, \frac{x_j}{|x|}\, \sqrt{m^2-\lambda^2}\, K_{d/2}\big(\sqrt{m^2-\lambda^2}\,|x|\big) + (m\,\beta + \lambda\,\mathbb{I}_N)\, K_{d/2-1}\big(\sqrt{m^2-\lambda^2}\,|x|\big) \Big),$$

where $c_{d,\lambda} = \frac{1}{2\pi}\big(\frac{\sqrt{m^2-\lambda^2}}{2\pi}\big)^{d/2-1}$ and K_ν is the modified Bessel function of the second kind. In particular, there is $C \geq 0$ so that

$$\big|R_0(\lambda)(x)\big| \leq \begin{cases} C\,|x|^{1-d} & \text{as } |x| \to 0, \\ C\,\mathrm{e}^{-\sqrt{m^2-\lambda^2}\,|x|} & \text{as } |x| \to +\infty. \end{cases}$$

Therefore, in the Dirac case, we have that $R_0(\lambda) \in \mathrm{L}^r(\mathbb{R}^d)$ for all $r < d/(d-1)$ and that $R_0(\lambda) \in \mathrm{L}^{d/(d-1)}_w(\mathbb{R}^d)$. In particular, the case (ii) is not covered by (i) only in the case where $r = d/(d-1)$, which corresponds to the critical exponent case $p = d \geq 2$, that is, $q = 2d/(d+2)$ in (3.3).

Remark 3.6. Let us consider the case $s = 1$ in (3.6). In order to see that $J(1) > 0$ in the Dirac case with $\lambda \in (-m, m)$, let $f \in \mathrm{L}^q(\mathbb{R}^d, \mathbb{C})$ be a normalized function and let $\phi_+ \in \mathbb{C}^N$ be a normalized vector such that $\beta\,\phi_+ = \phi_+$. We find that

$$(\mathcal{D}_m + \lambda)\, f\, \phi_+ = (m + \lambda)\, f\, \phi_+ + (-\mathrm{i}\nabla f) \cdot \alpha\, \phi_+.$$

Moreover, by (1.4), we have that $\langle \phi_+, \alpha_j\, \phi_+ \rangle_{\mathbb{C}^N} = 0$. Thus,

$$\begin{aligned} J(1) \geq \langle f\,\phi_+, R_0(\lambda)\, f\,\phi_+ \rangle &= \langle f\,\phi_+, (-\Delta + m^2 - \lambda^2)^{-1}\, (\mathcal{D}_m + \lambda)\, f\,\phi_+ \rangle \\ &= (m+\lambda)\, \langle f, (-\Delta + m^2 - \lambda^2)^{-1}\, f \rangle_{\mathrm{L}^2(\mathbb{R}^d,\mathbb{C})} > 0. \end{aligned} \tag{3.7}$$

Proof of Lemma 3.3. Note first that the condition $K(x) = K(-x)^*$ reads $\widehat{K}(k) = \widehat{K}(k)^*$, so the operator K is symmetric.

In the first part of the proof, we cover both cases (i) and (ii) by assuming

$$K \in \mathrm{L}^r_w(\mathbb{R}^d, \mathcal{M}_N(\mathbb{C})) \cap \mathrm{L}^r(\mathcal{B}_1^c, \mathcal{M}_N(\mathbb{C})) \quad \text{with } \mathcal{B}_R := \{x \in \mathbb{R}^d \,:\, |x| < R\} \tag{3.8}$$

with $2/q + 1/r = 2$. From the Hardy–Littlewood–Sobolev inequality, and since $K \in \mathrm{L}^r_w(\mathbb{R}^d)$, we have

$$\forall\, w_1, w_2 \in \mathrm{L}^q(\mathbb{R}^d), \quad |\langle w_1, K * w_2 \rangle| \leq C\, \|w_1\|_q\, \|w_2\|_q\, \|K\|_{r,w}. \tag{3.9}$$

In particular, $w \mapsto \langle w, K * w \rangle$ is well-defined and real valued on $\mathrm{L}^q(\mathbb{R}^d)$.

Using the scaling $w_s = s^{1/q}\, w_1$, we obtain that

$$J(s) = s^{2/q}\, J(1). \tag{3.10}$$

Since $J(1) > 0$, we deduce first that $J(s)$ is increasing. Also, since $2/q > 1$, $J(s)$ is convex and so we have the strong binding inequality

$$\forall\, s, s' > 0, \quad J(s + s') > J(s) + J(s'). \tag{3.11}$$

Let $(w_n)_{n\in\mathbb{N}}$ be a maximizing sequence for $J(1)$. Our argument relies on the concentration-compactness method for the sequence $(w_n)_{n\in\mathbb{N}}$, by following the approach of Lions [54] and using Levy's functional. It differs from the concentration-compactness method used in [34], as we work directly with the Birman–Schwinger operator instead of the min-max quadratic form. We set

$$Q(\rho) := \liminf_{n\to+\infty} Q_n(\rho), \quad \text{with } Q_n(\rho) := \sup_{x\in\mathbb{R}^d} \int_{\mathcal{B}(x,\rho)} |w_n|^q\, \mathrm{d}x.$$

It is clear from the definition that $\rho \mapsto Q(\rho)$ is non-decreasing, and that $Q(\rho) \leq 1$ for all $\rho > 0$. We set

$$\mu := \lim_{\rho\to+\infty} Q(\rho) \in [0, 1].$$

We divide the proof in the classical steps of the concentration-compactness method and start by discarding the cases $\mu = 0$ (vanishing) and $\mu < 1$ (dichotomy).

(a) *Vanishing*.

Fix $\varepsilon := J(1)/4 > 0$. Since $K \in \mathrm{L}^r(\mathcal{B}_1^c)$, there is $R > 1$ large enough so that

$$\|K\|_{\mathrm{L}^r(\mathcal{B}_R^c)} \le \varepsilon.$$

By Young's inequality, since $2/q + 1/r = 2$, we obtain that for all $w \in \mathrm{L}^q(\mathbb{R}^d)$ with $\|w\|_q = 1$,

$$\langle w, (\mathbb{1}_{\mathcal{B}_R^c} K) * w\rangle \le \varepsilon \|w\|_q^2 \le \varepsilon.$$

We now estimate the contribution of $\mathbb{1}_{\mathcal{B}_R} K$. For $z \in \mathbb{Z}^d$, let C_z be the cube $z + [0,1]^d$, so that $\{C_z\}_{z\in\mathbb{Z}^d}$ covers $\mathbb{R}^d$. For a function $w\colon \mathbb{R}^d \to \mathbb{C}^N$, we have

$$\begin{aligned}\langle w, (\mathbb{1}_{\mathcal{B}_R} K) * w\,\rangle &= \sum_{z,z'\in\mathbb{Z}^d} \iint_{C_z\times C_{z'}} \big\langle w(x), (\mathbb{1}_{\mathcal{B}_R} K)(x-y)\, w(y)\big\rangle_{\mathbb{C}^N}\, \mathrm{d}x\, \mathrm{d}y \\ &\le \|K\|_{\mathrm{L}^r_w(\mathbb{R}^d)} \sum_{z,z'\in\mathbb{Z}^d} \|w\|_{\mathrm{L}^q(C_z)}\, \|w\|_{\mathrm{L}^q(C_{z'})}\, \mathbb{1}_{\{|z-z'|\le R+2\sqrt{d}\,\}},\end{aligned}$$

using again the Hardy–Littlewood–Sobolev inequality. The double sum can be seen as a discrete convolution, and we apply Young's inequality with $z \mapsto \|w\|_{\mathrm{L}^q(C_z)} \in \ell^2(\mathbb{Z}^d)$ and $z \mapsto \mathbb{1}_{\{|z|\le R+2\sqrt{d}\,\}} \in \ell^1(\mathbb{Z}^d)$ to bound

$$\langle w, (\mathbb{1}_{\mathcal{B}_R} K) * w\rangle \le C_R \sum_{z\in\mathbb{Z}^d} \|w\|^2_{\mathrm{L}^q(C_z)} \le C_R \sup_{z\in\mathbb{Z}^d} \|w\|^{2-q}_{\mathrm{L}^q(C_z)}\, \|w\|_q^q,$$

where C_R is a positive constant which is independent of w: for all $w \in \mathrm{L}^q(\mathbb{R}^d)$ with $\|w\|_q = 1$, we have

$$\langle w, K * w\rangle \le \varepsilon + C_R \sup_{z\in\mathbb{Z}^d} \|w\|^{2-q}_{\mathrm{L}^q(C_z)}.$$

Applying this estimate to a maximizing sequence $(w_n)_{n\in\mathbb{N}}$ for $J(1) = 4\varepsilon$, we obtain that, up to a subsequence,

$$\frac{1}{2}\, J(1) \le \langle w_n, K * w_n\rangle \le \frac{1}{4}\, J(1) + C_R \sup_{z\in\mathbb{Z}^d} \|w_n\|^{2-q}_{\mathrm{L}^q(C_z)}.$$

This implies

$$Q_n(\sqrt{d}\,) \ge \sup_{z\in\mathbb{Z}^d} \|w_n\|^q_{\mathrm{L}^q(C_z)} \ge \frac{J(1)}{(4C_R))^{q/(2-q)}} > 0,$$

and finally $\mu > 0$, which discards the *vanishing* case of the concentration-compactness method.

(b) *Dichotomy*.

By definition of Q_n, there are sequences of centers $x_n \in \mathbb{R}^d$ and radii $\rho_n > 0$ going to infinity so that

$$\lim_{n\to+\infty} \int_{\mathcal{B}(x_n,\rho_n)} |w_n|^q\, \mathrm{d}x = \mu.$$

Without loss of generality, by translating the functions w_n, we may assume $x_n = 0$. In addition, up to a non-displayed subsequence, we have that for all $\varepsilon > 0$, there is n_0 large enough so that, for all $n \geq n_0$, we have

$$\int_{\rho_n<|x|<2\rho_n} |w_n|^q \,\mathrm{d}x < \varepsilon \quad \text{and} \quad \Big|1-\mu-\int_{|x|>2\rho_n} |w_n|^q \,\mathrm{d}x\Big| < \varepsilon.$$

We set

$$\begin{cases} w_n^{(1)} := w_n\,\mathbb{1}_{\{|x|\leq\rho_n\}}, \\ w_n^{(2)} := w_n\,\mathbb{1}_{\{\rho_n\leq|x|\leq 2\rho_n\}}, \\ w_n^{(3)} := w_n\,\mathbb{1}_{\{|x|>2\rho_n\}}. \end{cases}$$

Introducing

$$E(w_1,w_2) := \langle w_1, K * w_2\rangle \quad \text{and} \quad \mathcal{E}(w) := E(w,w),$$

we have

$$\begin{aligned} \mathcal{E}(w_n) &= \mathcal{E}(w_n^{(1)}) + \mathcal{E}(w_n^{(2)}) + \mathcal{E}(w_n^{(3)}) \\ &\qquad + 2\,\mathrm{Re}\big(E(w_n^{(1)},w_n^{(2)}) + E(w_n^{(1)},w_n^{(3)}) + E(w_n^{(2)},w_n^{(3)})\big) \\ &\leq J(\mu) + J(\varepsilon) + J(1-\mu+\varepsilon) \\ &\qquad + 2\,\mathrm{Re}\big(E(w_n^{(1)},w_n^{(2)}) + E(w_n^{(1)},w_n^{(3)}) + E(w_n^{(2)},w_n^{(3)})\big). \end{aligned}$$

From (3.9), and the fact that $\|w_n^{(2)}\|_q \leq \varepsilon^{1/q}$, we get that

$$E(w_n^{(1)},w_n^{(2)}) \leq C\,\mu\,\varepsilon^{1/q} \quad \text{and} \quad E(w_n^{(2)},w_n^{(3)}) \leq C\,(1-\mu)\,\varepsilon^{1/q}.$$

Finally, we have

$$\begin{aligned} |E(w_n^{(1)},w_n^{(3)})| &\leq \int_{|x|\leq\rho_n}\int_{|y|\geq 2\rho_n} \big|\langle w_n(x), K(x-y)\,w_n(y)\rangle_{\mathbb{C}^N}\big|\,\mathrm{d}x\,\mathrm{d}y \\ &\leq \iint_{\mathbb{R}^d\times\mathbb{R}^d} |w_n(x)|\,|w_n(y)|\,(K\,\mathbb{1}_{\mathcal{B}_{\rho_n}^c})(x-y)\,\mathrm{d}x\,\mathrm{d}y \\ &\leq \|K\,\mathbb{1}_{\mathcal{B}_{\rho_n}^c}\|_r \leq C\,\varepsilon \end{aligned}$$

for n large enough, where in the last line we used Young's inequality, and the fact that $\rho_n \to +\infty$. Thanks to these facts, we can conclude that

$$J(1) \leq J(\mu) + J(1-\mu+\varepsilon) + J(\varepsilon) + C\,\varepsilon^{1/q}.$$

In the limit as $\varepsilon \to 0$, we obtain $J(1) \leq J(\mu) + J(1-\mu)$, which contradicts (3.11) if $\mu \neq 1$. So $\mu = 1$, which discards the *dichotomy* case of the concentration-compactness method.

(c) *Convergence for tight sequences.*

At this point, we proved that for all $\varepsilon > 0$ there is $\rho > 0$ and n_0 large enough so that, for all $n > n_0$, and after appropriate translations and subsequences,

$$\|\mathbb{1}_{\mathcal{B}_\rho^c}\,w_n\|_q \leq \varepsilon. \tag{3.12}$$

In other words, the sequence $(w_n)_{n\in\mathbb{N}}$ is tight in $\mathrm{L}^q(\mathbb{R}^d,\mathbb{C}^N)$. The sequence $(w_n)_{n\in\mathbb{N}}$ is bounded in the reflexive Banach space $\mathrm{L}^q(\mathbb{R}^d,\mathbb{C}^N)$. Hence, up to a non-displayed subsequence, $(w_n)_{n\in\mathbb{N}}$ converges weakly to some $w \in \mathrm{L}^q(\mathbb{R}^d,\mathbb{C}^N)$, and we have $\|w\|_q \leq 1$.

Let us prove that $\mathcal{E}(w) = J(1)$. Let $\varepsilon > 0$, and let $\rho > 0$ be large enough so that (3.12) holds. In particular, by the Hardy–Littlewood–Sobolev inequality, we have

$$|\langle w_n \mathbb{1}_{\mathcal{B}_\rho^c}, K * w_n\rangle| \le C \|w_n\|_q \|w_n \mathbb{1}_{\mathcal{B}_\rho^c}\|_q \le C\varepsilon,$$

and we have a similar inequality with w instead of w_n. On the other hand, we have

$$\langle w_n \mathbb{1}_{\mathcal{B}_\rho}, K * w_n\rangle = \langle w_n, T w_n\rangle_{\mathrm{L}^q,\mathrm{L}^{q'}},$$

where T is the operator from $\mathrm{L}^q(\mathbb{R}^d)$ to $\mathrm{L}^{q'}(\mathbb{R}^d)$ with kernel $T(x,y) = \mathbb{1}_{\mathcal{B}_\rho}(x)\,K(x-y)$. The operator $T\colon \mathrm{L}^q(\mathbb{R}^d,\mathbb{C}^N) \to \mathrm{L}^{q'}(\mathbb{R}^d,\mathbb{C}^N)$ is bounded. We claim that T is a compact operator. In the Dirac case (ii), this comes from the fact that $K * w_n \in \mathrm{W}^{1,q}$ with $\|K * w_n\|_{\mathrm{W}^{1,q}(\mathbb{R}^d,\mathbb{C}^N)} \le C\,\|w_n\|_q$ together with the Rellich–Kondrachov compact embedding theorem. In the case (i), where $K \in \mathrm{L}^r(\mathbb{R}^d)$, setting $\tau_h f(x) := f(x-h)$, we have

$$\|\tau_h(K * w) - K * w\|_{\mathrm{L}^{q'}(B_\rho)} = \|(\tau_h K - K) * w\|_{q'} \le \|\tau_h K - K\|_r \,\|w\|_q.$$

Since $K \in \mathrm{L}^r(\mathbb{R}^d)$, we have $\|\tau_h K - K\|_r \to 0$ as $h \to 0$, and we conclude with the Kolmogorov–Riesz–Fréchet theorem (see, for instance, Theorem 4.26 in [11]).

As a consequence, $(T w_n)_{n\in\mathbb{N}}$ converges strongly to Tw in $\mathrm{L}^{q'}(\mathbb{R}^d,\mathbb{C}^N)$. In particular, we obtain that

$$\lim_{n\to+\infty} \langle w_n \mathbb{1}_{\mathcal{B}_\rho}, K * w_n\rangle = \langle w \mathbb{1}_{\mathcal{B}_\rho}, K * w\rangle.$$

Gathering the two inequalities gives

$$\begin{aligned}
&|\langle w_n, K * w_n\rangle - \langle w, K * w\rangle| \\
&\quad\le |\langle w_n \mathbb{1}_{\mathcal{B}_\rho}, K * w_n\rangle - \langle w \mathbb{1}_{\mathcal{B}_\rho}, K * w\rangle| + |\langle w_n \mathbb{1}^c_{\mathcal{B}_\rho}, K * w_n\rangle| + |\langle w \mathbb{1}^c_{\mathcal{B}_\rho}, K * w\rangle| \\
&\quad\le |\langle w_n \mathbb{1}_{\mathcal{B}_\rho}, K * w_n\rangle - \langle w \mathbb{1}_{\mathcal{B}_\rho}, K * w\rangle| + 2C\varepsilon.
\end{aligned}$$

Sending first n to $+\infty$, and then ε to 0 shows that $\langle w, K * w\rangle = J(1)$. Finally, since $\|w\|_q \le 1$, by (3.10) we deduce that $\|w\|_q = 1$. This proves that $(w_n)_{n\in\mathbb{N}}$ converges strongly to w in $\mathrm{L}^q(\mathbb{R}^d,\mathbb{C}^N)$ and that w is an optimizer. ∎

It is an open question to decide whether T is compact or not under the condition (3.8).

Proof of Theorem 3.1. In the setting of Theorem 3.1, we take $K = R_0(\lambda)$. In this case, we have $J(1) > 0$, as noticed in Remark 3.6, so by Lemma 3.3, the problem (3.3) admits maximizers. By standard arguments, optimizers satisfy the Euler–Lagrange equation (3.4).

The fact that $\lambda \mapsto \mathcal{N}(\lambda,p)$ is strictly increasing comes from the fact that $\lambda \mapsto R_0(\lambda)$ is operator strictly increasing: for instance, we have $\partial_\lambda R_0(\lambda) = (R_0(\lambda))^2 > 0$. Let us prove the continuity. Let $-m < \lambda' < \lambda < m$, and let w_λ be the optimizer for $\mathcal{N}(\lambda,p)$. Using that $\mathcal{N}(\cdot)$ is strictly increasing and the resolvent identity

$$R_0(\lambda') = R_0(\lambda) - (\lambda - \lambda')\,R_0(\lambda')\,R_0(\lambda),$$

we obtain

$$0 < \mathcal{N}(\lambda,p) - \mathcal{N}(\lambda',p) \le (\lambda - \lambda')\,\langle w_\lambda, R_0(\lambda')\,R_0(\lambda)\,w_\lambda\rangle.$$

Using that R_0 is a bounded operator from $\mathrm{L}^q(\mathbb{R}^d)$ to $\mathrm{L}^{q'}(\mathbb{R}^d)$, and from $\mathrm{L}^{q'}(\mathbb{R}^d)$ into itself, with uniform bounds in a neighborhood of λ, we deduce that there is $C > 0$ so that

$$|\langle w_\lambda, R_0(\lambda')\, R_0(\lambda)\, w_\lambda\rangle| \le C\, \|w_\lambda\|_q^2 = C,$$

This proves that $\mathcal{N}(\cdot, p)$ is locally Lipschitz, hence continuous.

We now prove the bounds on $\lim_{\lambda\to\pm m} \mathcal{N}(\lambda, p)$.

To prove that $\lim_{\lambda\to m} \mathcal{N}(\lambda, p) = +\infty$, we go back to (3.7) and take a function $f = L^{-d/q}\, g(\cdot/L)$, where g is an arbitrary test function that is normalized in $\mathrm{L}^q(\mathbb{R}^d)$. This gives

$$\mathcal{N}(\lambda, p) \ge L^{-2d/q}\, (m+\lambda) \big\langle g(\cdot/L), (-\Delta + m^2 - \lambda^2)^{-1}\, g(\cdot/L)\big\rangle.$$

We bound the resolvent as

$$(-\Delta + m^2 - \lambda^2)^{-1} \ge (m^2 - \lambda^2)^{-1}\, (1 + (m^2-\lambda^2)^{-1}\,\Delta),$$

and change variables to obtain

$$\mathcal{N}(\lambda, p) \ge L^{d(1-2/q)}\, (m-\lambda)^{-1} \big(\|g\|_2^2 - L^{-2}\, (m^2-\lambda^2)^{-1}\, \|\nabla g\|_2^2\big).$$

Since $1 - 2/q = 1/p$, we may take $L = (m-\lambda)^{-\alpha}$ for any $\alpha \in (1/2, p/d)$ and conclude that $\lim_{\lambda\to m} \mathcal{N}(\lambda, p) = +\infty$.

Finally, to prove that $\lim_{\lambda\to -m} \mathcal{N}(\lambda, p) > 0$, we claim that

(3.13) *there exists a function* $w \in \mathrm{L}^2 \cap \mathrm{L}^q(\mathbb{R}^d, \mathbb{C}^N)$ *such that* $\|w\|_q = 1$ *and* $Pw = w$,

where $P := \mathbb{1}_{m < \mathcal{D}_m < 2m}$ is the spectral projection of the free Dirac operator onto $(m, 2m)$. This would give

$$\mathcal{N}(\lambda, p) \ge \langle w, R_0(\lambda)\, w\rangle = \Big\langle w, \frac{P}{\mathcal{D}_m - \lambda}\, w\Big\rangle + \Big\langle w, \frac{P^\perp}{\mathcal{D}_m - \lambda}\, w\Big\rangle.$$

The second term is null since $P^\perp\, w = 0$. For the first term, we have $m < \mathcal{D}_m < 2m$ on the range of P, and in particular $P\, (\mathcal{D}_m - \lambda)^{-1}\, P \ge P\, (2m-\lambda)^{-1}\, P$, hence $\mathcal{N}(\lambda, p) \ge (2m-\lambda)^{-1}\, \|w\|_2$. Taking $\lambda \to -m$ shows that $\lim_{\lambda\to -m} \mathcal{N}(\lambda, p) \ge (3\,m)^{-1}\, \|w\|_2 > 0$.

It remains to prove (3.13). Recall that $\mathcal{D}_m = \mathcal{F}\, M(k)\, \mathcal{F}^*$, where $\mathcal{F}$ denotes the Fourier transform and $M(k)$ is the $d \times d$ matrix $M(k) := \boldsymbol{\alpha}\cdot k + m\,\beta$, which satisfies $M(k) = M(k)^*$, $M(k)^2 = (|k|^2 + m^2)\,\mathbb{I}_d$, and $\sigma(M(k)) = \{\pm(|k|^2+m^2)^{1/2}\}$. Let $v \mapsto v(k)$ be a smooth family of spinors from some open ball $\mathcal{B}(k=0,\varepsilon)$ to $\mathbb{C}^d$, with $0 < \varepsilon < m$, so that $M(k)\, v(k) = (|k|^2+m^2)^{1/2}\, v(k)$. To construct such a local family of spinors, one can consider v_0 a normalized eigenfunction of $M(k=0)$, and set

$$v(k) := \frac{P(k)\, v_0}{\|P(k)\, v_0\|_2}, \qquad P(k) := \mathbb{1}_{M(k)>0}.$$

Since $k \mapsto P(k)$ is smooth locally around 0 ($P(k)$ can be written as a Cauchy integral $P(k) = (2\,\mathrm{i}\,\pi)^{-1} \oint_{\mathcal{C}} (z - M(k))^{-1}\, \mathrm{d}z$ with a contour enclosing m), so is $k \mapsto v(k)$. Let also $\chi(k)\colon \mathbb{R}^d \to \mathbb{R}^+$ be a non-null smooth compactly supported function, with $\chi(k) = 0$ for $|k| > \varepsilon$. We consider the function

$$w := \frac{\widetilde{w}}{\|\widetilde{w}\|_q}, \quad \text{with } \widetilde{w} := \mathcal{F}(\chi(k)\, v(k)).$$

By construction, we have $\widetilde{w} \neq 0$, and since $\widetilde{w}$ has a Fourier transform which is smooth and compactly supported, it belongs to the Schwartz class $\mathcal{S}(\mathbb{R}^d, \mathbb{C}^N)$. Finally, since on the support of χ, we have $M(k)\, v(k) = (|k|^2 + m^2)^{1/2}\, v(k)$ with $m < (|k|^2 + m^2)^{1/2} < \sqrt{2}\, m^2$, we deduce that

$$P\, \widetilde{w} = \mathcal{F}\big(\mathbb{1}_{m < M(k) < 2m}\, \chi(k)\, v(k)\big) = \mathcal{F}(\chi(k)\, v(k)) = \widetilde{w},$$

which concludes the proof of (3.13). ∎

3.3. Regularity of the solutions of the nonlinear Dirac equation

Under condition (2.1), solutions of (3.5) with $\Psi \in \mathrm{L}^2(\mathbb{R}^d, \mathbb{C}^N)$ are in $\mathrm{Dom}(\mathcal{D}_m - V) = \mathrm{H}^1(\mathbb{R}^d, \mathbb{C}^N)$. Let us consider the other cases of Proposition 2.1. If $d = 1$ and $1 < p \le 2$, any optimal function for (3.3) obtained in Theorem 3.1 gives rise to a solution $\Psi \in \mathrm{W}^{1,q}(\mathbb{R}, \mathbb{C}^2)$ of (3.5) with $q = 2p/(p+1)$. We conclude that Ψ is continuous. Now, if $p = d = 2$ and $q = 4/3$, the corresponding solution Ψ of (3.5) is in $\mathrm{W}^{1,q}(\mathbb{R}, \mathbb{C}^2) \hookrightarrow \mathrm{H}^{1/2}(\mathbb{R}, \mathbb{C}^2)$, hence $V\, |\Psi|^2 = |\Psi|^{2p/(p-1)}$ is integrable and $\Psi \in \mathrm{Dom}(\mathcal{D}_m - V)$ of the distinguished extension of Proposition 2.1, but we do not know whether $\Psi \in \mathrm{H}^1(\mathbb{R}^2, \mathbb{C}^2)$ or not.

In dimension $d = 1$, an explicit expression of the solutions of (3.5) such that

$$\lim_{x \to \pm\infty} \Psi(x) = (0, 0)^\intercal$$

is given in Theorem 1.3. In the case $p = d = 2$, it is unclear how to obtain $\Psi \in \mathrm{H}^1(\mathbb{R}^d, \mathbb{C}^N)$ by general arguments, as pointed out in [6]. However, any solution to (3.5) (and not only the ones found in Theorem 3.1) have additional regularity properties under Condition (2.1).

Proposition 3.7. *Let $\lambda \in [-m, m)$ and either $p \ge d$ if $d \ge 3$, or $p > d$ in dimensions $d = 1, 2$. If $\Psi \in \mathrm{H}^1(\mathbb{R}^d, \mathbb{C}^N)$ solves (3.5), then $\Psi \in C^\infty(\mathbb{R}^d, \mathbb{C}^N)$.*

Proof. Let us first prove that $\Psi \in \mathrm{L}^\infty(\mathbb{R}^d, \mathbb{C}^N)$ with a usual bootstrap argument. If $\Psi \in \mathrm{L}^q(\mathbb{R}^d, \mathbb{C}^N)$, then $|\Psi|^{\frac{2}{p-1}}\, \Psi \in \mathrm{L}^{\frac{p-1}{p+1} q}(\mathbb{R}^d, \mathbb{C}^N)$. Also, if $q > 2\, \frac{p+1}{p-1}$, then $2 < \frac{p-1}{p+1}\, q < q$, so if $\Psi \in \mathrm{L}^2(\mathbb{R}^d, \mathbb{C}^N) \cap \mathrm{L}^q(\mathbb{R}^d, \mathbb{C}^N)$, then $\lambda\, \Psi + |\Psi|^{\frac{2}{p-1}}\, \Psi \in \mathrm{L}^{\frac{p-1}{p+1} q}(\mathbb{R}^d, \mathbb{C}^N)$. In particular, $\mathcal{D}_m \Psi \in \mathrm{L}^{\frac{p-1}{p+1} q}(\mathbb{R}^d, \mathbb{C}^N)$, hence $\Psi \in \mathrm{W}^{1, \frac{p-1}{p+1} q}(\mathbb{R}^d, \mathbb{C}^N) \hookrightarrow \mathrm{L}^{\tilde{q}}(\mathbb{R}^d, \mathbb{C}^N)$, with $\tilde{q} = +\infty$ if $\frac{p-1}{p+1}\, q > d$, and

$$\frac{1}{\tilde{q}} = \frac{p+1}{p-1}\, \frac{1}{q} - \frac{1}{d}$$

otherwise. As a first step of an iteration scheme, we proved that if $\Psi \in \mathrm{L}^q(\mathbb{R}^d, \mathbb{C}^N)$, then $\Psi \in \mathrm{L}^{\tilde{q}}(\mathbb{R}^d, \mathbb{C}^N)$ as well. For the initialization, we note that $\mathrm{H}^1(\mathbb{R}^d, \mathbb{C}^N) \hookrightarrow \mathrm{L}^q(\mathbb{R}^d, \mathbb{C}^N)$

for all q such that $2 \le q \le \frac{2d}{d-2} =: 2^*$ if $d \ge 3$ and $2 \le q < +\infty =: 2^*$ if $d = 2$. Hence with $2\,\frac{p+1}{p-1} < 2^*$, there is $q_0 > 2\,\frac{p+1}{p-1}$ so that $\Psi \in \mathrm{L}^2(\mathbb{R}^d, \mathbb{C}^N) \cap \mathrm{L}^{q_0}(\mathbb{R}^d, \mathbb{C}^N)$. The map $F\colon x \mapsto \frac{p+1}{p-1}\,x - \frac{1}{d}$ satisfies $F(x) < x$ for $x \in [0, x^*]$ with $x^* = \frac{p-1}{2d} < \frac{1}{2}\,\frac{p-1}{p+1}$. We easily deduce that there is $n \in \mathbb{N}$ so that $F^{(n)}(1/q_0) < 0$, which proves $\Psi \in \mathrm{L}^\infty(\mathbb{R}^d, \mathbb{C}^N)$ as wanted.

Since $\not{D}_m \Psi \in \mathrm{L}^\infty(\mathbb{R}^d, \mathbb{C}^N)$, we have $\Psi \in \mathrm{W}^{1,\infty}(\mathbb{R}^d, \mathbb{C}^N) \hookrightarrow C^{0,\alpha}(\mathbb{R}^d, \mathbb{C}^N)$ for all $0 \le \alpha < 1$; by bootstrapping again, we obtain $\Psi \in C^\infty(\mathbb{R}^d, \mathbb{C}^N)$. ∎

4. Lieb–Thirring inequality

This section contains the proof of Theorem 1.4. We closely follow the original proof by Lieb and Thirring [52, 53] (see also [51]). This is possible since we are assuming that $V \ge 0$. In the general case where V has no sign, some results can be found in the works of Cuenin [17], and Frank–Simon [40], where the authors control the Riesz-mean

$$\sum_k \operatorname{dist}(\lambda_k, \sigma(\not{D}_m - V))^\gamma,$$

that is, the distance to the whole spectrum. Actually, without assuming a sign on V, one cannot expect to control the sums in (1.13), since, for $V \le 0$ small, the eigenvalues of $\not{D}_m - V$ emerge from the bottom essential spectrum (hence have a distance of order $2m > 0$ to the upper essential spectrum). Here, since V is nonnegative, the eigenvalues emerge from the upper essential spectrum as the strength of the potential increases.

Proof of Theorem 1.4. It is sufficient to prove the result for V bounded and compactly supported. By the *Birman–Schwinger principle* introduced in Section 2.2, we know that λ is an eigenvalue for $\not{D}_m - V$ acting on $\mathbb{C}^N$ valued spinors if and only if 1 is an eigenvalue of $K_V(\lambda)$ defined by (2.3); see Proposition 2.4. We also proved that $\lambda \mapsto K_V(\lambda)$ is operator increasing. In particular, if we set

$$N_e(V) := \text{number of eigenvalues of } \not{D}_m - V \text{ in } [-m, m-e]$$

and

$$B_e(V) := \text{number of eigenvalues of } K_V(m-e) \text{ which are greater or equal than } 1,$$

then we have $N_e(V) \le B_e(V)$. We have equality if the highest eigenvalues of $K_V(\lambda)$ gets strictly smaller than 1 as $\lambda \to -m$. This happens for instance if $\|V\|_p \le \alpha_*(p)$.

With R_0 defined by (2.2), using the operator inequality

$$R_0(\lambda) \le \mathbb{1}_{\mathbb{C}^N}\big(\sqrt{-\Delta + m^2} - \lambda\big)^{-1},$$

we can estimate $B_e(V)$ by $N\,B_e^{\mathrm{pr}}(V)$, where $B_e^{\mathrm{pr}}(V)$ is the number of eigenvalues above 1 of the pseudo-relativistic Birman–Schwinger operator

$$K_V^{\mathrm{pr}}(m-e) := \sqrt{V}\big(\sqrt{-\Delta + m^2} - m + e\big)^{-1}\sqrt{V}.$$

In addition, with the definition

$$N_e^{\mathrm{pr}}(V) := \text{number of eigenvalues of } \big(\sqrt{-\Delta + m^2} - m\big) - V \text{ less or equal than } -e,$$

the usual Birman–Schwinger principle shows that $B_e^{\rm pr}(V) = N_e^{\rm pr}(V)$. To sum up, we have

$$N_e(V) \leq B_e(V) \leq N B_e^{\rm pr}(V) = N N_e^{\rm pr}(V). \tag{4.1}$$

The operator $\sqrt{-\Delta + m^2} - m$ is sometimes called the Chandrasekhar (or pseudo-relativistic) kinetic energy operator. It is a positive operator, $\sqrt{-\Delta + m^2} - m - V$ is bounded from below, and the min-max formula applies. We can now repeat the usual arguments of Lieb and Thirring for the pseudo-relativistic operator.

First, for $\gamma > 0$, the cake-layer representation gives

$$\sum_{k\geq 1} e_k^\gamma = \gamma \int_0^{2m} e^{\gamma-1}\, N_e(V)\, \mathrm{d}e \leq \gamma\, N \int_0^{2m} e^{\gamma-1}\, B_e^{\rm pr}(V)\, \mathrm{d}e. \tag{4.2}$$

Note that for the pseudo-relativistic model, if $-e_1^{\rm pr} \leq -e_2^{\rm pr} \leq \cdots < 0$ are the negative eigenvalues of $(\sqrt{-\Delta + m^2} - m) - V$, we have

$$\sum_{k\geq 1} (e_k^{\rm pr})^\gamma = \gamma \int_0^{\infty} e^{\gamma-1}\, N_e^{\rm pr}(V)\, \mathrm{d}e = \gamma \int_0^{\infty} e^{\gamma-1}\, B_e^{\rm pr}(V)\, \mathrm{d}e,$$

and the integral runs over $e \in \mathbb{R}^+$ instead of $e \in (0, 2m)$. Actually, the previous two inequalities together with (4.1) show that

$$\sum_{k\geq 1} e_k^\gamma \leq N \sum_{k\geq 1} (e_k^{\rm pr})^\gamma.$$

In other words, the Riesz-mean of the eigenvalues increases when one replaces the Dirac operator by the pseudo–relativistic one (up to the N factor). Lieb–Thirring inequalities for the last sum have been derived by Daubechies in [19] (and used, e.g., in [50]). In what follows, we derive another inequality specifically for the Dirac operator. We use in particular the fact that the integral in (4.2) only runs for e in the bounded interval $(0, 2m)$ instead of $\mathbb{R}^+$.

• *Bound for* $B_e^{\rm pr}(V)$.

Assume $V \in \mathrm{L}^p(\mathbb{R}^d)$ with $d < p$. The number of eigenvalues above 1 of $K_V^{\rm pr}(m-e)$ is bounded from above by $\|K_V^{\rm pr}(m-e)\|_{\mathfrak{S}^p}^p$. We estimate this norm using the Kato–Simon–Seiler inequality (see Theorem 4.2 in [66]). Using a decomposition similar to the one in the proof of Lemma 2.3, we obtain

$$B_e^{\rm pr}(V) \leq \|K_V^{\rm pr}(m-e)\|_{\mathfrak{S}_p}^p \leq C_p\, \|g_{m,e}\|_p^p\, \|V\|_p^p,$$

where we introduced the function

$$g_{m,e}(k) := \big(\sqrt{k^2 + m^2} - (m-e)\big)^{-1}.$$

Note that $g_{m,e} \in \mathrm{L}^p(\mathbb{R}^d)$ since $p > d$, and

$$\|g_{m,e}\|_p^p = \int_{\mathbb{R}^d} \frac{\mathrm{d}k}{\big(\sqrt{k^2+m^2} - m + e\big)^p} = |\mathbb{S}^{d-1}| \int_0^\infty \frac{r^{d-1}\,\mathrm{d}r}{\big(\sqrt{r^2+m^2} - m + e\big)^p}.$$

To estimate this norm, we make the change of variable $X = \frac{1}{e}(\sqrt{r^2+m^2}-m)$, so that $r = \sqrt{(eX+m)^2-m^2} = \sqrt{eX\,(eX+2m)}$. We obtain

$$\|g_{m,e}\|_p^p = \frac{|\mathbb{S}^{d-1}|}{e^{p-d/2}} \int_0^\infty \frac{\big(X(eX+2m)\big)^{d/2-1}(eX+m)\,\mathrm{d}X}{(X+1)^p}.$$

The last integral is an increasing function of e (and has a finite value as $e \to 0$ by the monotone convergence theorem). Since $e \in (0, 2m)$, we can bound this integral by its value at $e = 2m$. We deduce that there is a constant $C_{p,d}$ such that

$$B_e^{\mathrm{pr}}(V) \le C_{p,d}\, \|V\|_p^p \, \frac{m^{d/2}}{e^{p-d/2}}. \tag{4.3}$$

• *Proof of the Lieb–Thirring estimate.*

We now follow [51–53]. The min–max principle for the pseudo-relativistic operator shows that its eigenvalues are decreasing when V increases. Since $V \le [V-e/2]_+ + e/2$, we may bound

$$\begin{aligned} B_e^{\mathrm{pr}}(V) = N_e^{\mathrm{pr}}(V) &\le N_e^{\mathrm{pr}}([V-e/2]_+ + e/2) \\ &= N_{e/2}^{\mathrm{pr}}([V-e/2]_+) = B_{e/2}^{\mathrm{pr}}([V-e/2]_+). \end{aligned}$$

For any $p > d$, we can apply the bound in (4.3) to estimate $B_{e/2}^{\mathrm{pr}}([V-e_2]_+)$. Inserting this estimate into (4.2), we get

$$\begin{aligned} \sum_{k\ge1} e_k^\gamma &\le N\, C_{p,d}\, \gamma m^{d/2} \int_0^{2m} (e/2)^{\gamma-1+d/2-p}\, \|[V-e/2]_+\|_p^p \,\mathrm{d}e \\ &= C_{\gamma,d,p}\, m^{d/2} \int_{\mathbb{R}^d} \int_0^{2m} e^{\gamma-1+d/2-p}\, [V(x)-e/2]_+^p \,\mathrm{d}e\,\mathrm{d}x \\ &\le C_{\gamma,d,p}\, m^{d/2} \int_{\mathbb{R}^d} V^{\gamma+d/2}(x) \int_0^{s^*(x)} s^{\gamma-1+d/2-p}\, (1-s)_+^p \,\mathrm{d}s\,\mathrm{d}x, \end{aligned}$$

where $s^*(x) := \min\{m/V(x), 1\}$, with the convention that $s^*(x) = 1$ if $V(x) = 0$. The second integral converges whenever $p < \gamma + d/2$. Using simply the bound $(1-s)^p \le 1$ in the last integral, we finally obtain

$$\sum_{k\ge1} e_k^\gamma \le L_{\gamma,d,p}\, m^{d/2} \int_{\mathbb{R}^d} V_m^{\gamma+d/2-p}\, V^p \,\mathrm{d}x, \quad \text{with } V_m := \min\{m, V\}. \tag{4.4}$$

This inequality is valid for all $d < p < \gamma + d/2$. Note that $C_{\gamma,d,p}$ stays bounded in the limit as $p \to \gamma + d/2$, so a similar inequality also holds if $p = \gamma + d/2$. ■

Remark 4.1. The result of Theorem 1.4 can be extended to the case of a potential $V \in \mathrm{L}^p(\mathbb{R}^d, \mathbb{R}^+) + \mathrm{L}^{\gamma+d/2}(\mathbb{R}^d, \mathbb{R}^+)$ by noticing that the right-hand side of (4.4) is continuous for V in this space.

5. Explicit computations

5.1. The case $d = 1$: proof of Theorem 1.3

In this section, we prove the uniqueness and the symmetry up to translations of the solution of the nonlinear Dirac equation (3.5). We also compute the map $\alpha_D(\lambda, p)$.

Proof of Theorem 1.3. In the one-dimensional case, equation (3.5) can be rewritten for the components of $\Psi =: (\varphi, \chi)^\mathsf{T}$ as

$$\begin{cases} \varphi' = -\big(\lambda + m + (|\chi|^2 + |\varphi|^2)^{1/(p-1)}\big)\,\chi, \\ \chi' = \big(\lambda - m + (|\chi|^2 + |\varphi|^2)^{1/(p-1)}\big)\,\varphi. \end{cases} \tag{5.1}$$

The corresponding potential is $V = (|\chi|^2 + |\varphi|^2)^{1/(p-1)}$. This system conserves

$$H(\varphi, \chi) := m\,(|\chi|^2 - |\varphi|^2) + \lambda\,(|\chi|^2 + |\varphi|^2) + \tfrac{p-1}{p}\,(|\chi|^2 + |\varphi|^2)^{p/(p-1)},$$
$$G(\varphi, \chi) := \bar\chi\,\varphi - \bar\varphi\,\chi.$$

Since we are looking for solutions vanishing at $\pm\infty$, they satisfy $H(\varphi(x), \chi(x)) = 0$ and $G(\varphi(x), \chi(x)) = 0$ for all $x \in \mathbb{R}$. This second condition shows that solutions can be chosen real valued. For real valued variables in the (φ, χ)-plane, the level set $H(\varphi, \chi) = 0$ has the shape of an infinity sign. Among real valued functions, uniqueness up to translations follows from the phase plane analysis. We can choose the unique solution with $\chi(0) = 0$ and $\varphi(0) > 0$ given. For this solution, φ is even and χ is odd and positive on $\mathbb{R}^+$. Hence symmetry and uniqueness, up to translations and multiplication by a phase, are granted by elementary considerations. Next, we have

$$V' = \frac{1}{p-1}\,\frac{(\chi^2 + \varphi^2)'}{(\chi^2 + \varphi^2)^{p/(p-1)}} \quad \text{with} \;\; (\chi^2 + \varphi^2)' = -4m\chi\varphi,$$

which proves that V is increasing in the quadrant $\{\chi < 0, \varphi > 0\}$ and decreasing in the quadrant $\{\chi > 0, \varphi > 0\}$. Hence V is even and decreasing on $\mathbb{R}^+$, while on $\mathbb{R}^+$ both χ and φ are positive valued.

Now let us compute $\|V\|_p$. It is enough to do the computation on $\mathbb{R}^+$. First, the equation $H(\varphi, \chi) = 0$ can be rewritten as

$$2m\,\varphi^2 = (m + \lambda)\,V^{p-1} + \frac{p-1}{p}\,V^p,$$

and so

$$\varphi = \sqrt{\tfrac{1}{2m}\,V^{p-1}\big(m + \lambda + \tfrac{p-1}{p}\,V\big)}.$$

Next, from the equation $V^{p-1} = \chi^2 + \varphi^2$, we deduce that

$$\chi = \sqrt{V^{p-1} - \varphi^2} = \sqrt{\tfrac{1}{2m}\,V^{p-1}\big(m - \lambda - \tfrac{p-1}{p}\,V\big)}.$$

Finally, we have

$$(p-1)\,V^{p-2}\,V' = (V^{p-1})' = (\chi^2 + \varphi^2)' = -4m\chi\varphi.$$

Collecting the three last equalities shows that V solves the autonomous differential equation

$$V' = -\frac{2}{p-1}\,V\,\sqrt{\left(m-\lambda-\tfrac{p-1}{p}\,V\right)\left(m+\lambda+\tfrac{p-1}{p}\,V\right)}.$$

At $x=0$, we have $V'(0)=0$, which implies

$$V(0) = \frac{p}{p-1}\,(m-\lambda).$$

- *Subcritical regime* $\lambda > -m$. The function

$$Z(x) := \frac{p-1}{p\,(m+\lambda)}\,V\Big(\frac{p-1}{2\,(m+\lambda)}\,x\Big)$$

satisfies

$$Z' = -Z\,\sqrt{(z_0-Z)\,(1+Z)},\quad Z(0)=z_0=(m-\lambda)/(m+\lambda). \tag{5.2}$$

One can directly check that the solution of (5.2) is

$$Z(x) = \frac{2\,z_0}{(1+z_0)\cosh(\sqrt{z_0}\,x)+1-z_0}.$$

This gives (1.10). The $\mathrm{L}^p(\mathbb{R})$ norm of V is computed as

$$\|V\|_p^p = \frac{p^p\,(m+\lambda)^{p-1}}{2\,(p-1)^{p-1}}\,\|Z\|_p^p.$$

Using that Z is even, monotone decreasing on $\mathbb{R}^+$, with the change of variable $z=Z(x)$ and $t=z/z_0$, we obtain, using (5.2),

$$\begin{aligned}\|Z\|_p^p &= 2\int_0^{+\infty} Z^p(x)\,\mathrm{d}x\\ &= 2\int_0^{z_0}\frac{z^{p-1}}{\sqrt{(z_0-z)\,(1+z)}}\,\mathrm{d}z = 2\,z_0^{p-1/2}\int_0^1\frac{t^{p-1}}{\sqrt{(1-t)\,(1-(-z_0)\,t)}}\,\mathrm{d}t\\ &= 2\,z_0^{p-1/2}\,B\left(\tfrac12,p\right)\,{}_2F_1\left(\tfrac12,p;p+\tfrac12;-z_0\right).\end{aligned}$$

See formula 15.3.1 in p. 558 of [1] for the last equality. This completes the computation of $\alpha_D(\lambda,p)$. By taking the limit as $p\to 1_+$, we obtain $\alpha_D(\lambda,1)=\arccos(\lambda/m)$.

- *Critical case* $\lambda=-m$. The function

$$Z(x) := \frac{p-1}{2m\,p}\,V\Big(\frac{p-1}{2m}\,x\Big)$$

solves

$$Z' = -2\,Z^{3/2}\,\sqrt{1-Z},\quad Z(0)=1$$

on $\mathbb{R}^+$. The solution is

$$\forall\,x\in\mathbb{R},\quad Z(x)=\frac{1}{1+x^2}.$$

This gives (1.11), and the expression of $\alpha_\star(p)$ follows from

$$\|V\|_p^p = \frac{p^p\,(2m)^{p-1}}{(p-1)^{p-1}}\,\|Z\|_p^p$$

with

$$\|Z\|_p^p = \int_{\mathbb{R}} \frac{\mathrm{d}x}{(1+x^2)^p} = B\left(\tfrac{1}{2}, p-\tfrac{1}{2}\right)$$

according to formula 8.380.3 in p. 917 of [42]. This concludes the proof of Theorem 1.3. ∎

Notice that

$$\lim_{z_0\to+\infty} \sqrt{z_0}\, B\left(\tfrac{1}{2}, p\right)\, {}_2F_1\left(\tfrac{1}{2}, p; p+\tfrac{1}{2}; -z_0\right) = B\left(\tfrac{1}{2}, p-\tfrac{1}{2}\right),$$

so that $\lim_{\lambda\to(-1)_+} \alpha_D(\lambda, p) = \alpha_\star(p)$.

5.2. The radial case in dimension $d = 2$

We now provide some numerical simulations to get upper bounds for the maps $\Lambda_D(\alpha, p)$.

First, we restrict the minimization problem (1.5) to radial potentials, that is, we compute

$$\Lambda_D^{\mathrm{rad}}(\alpha, p) := \inf\{\lambda_D(V) \,:\, V \in \mathrm{L}^p(\mathbb{R}^d, \mathbb{R}^+),\; V \text{ radial and } \|V\|_p = \alpha\}.$$

Below in Appendix B, we provide some numerical evidences that the optimal potentials are radial. We abusively write $V(x) = V(r)$ with $r = |x|$, $x \in \mathbb{R}^2$, use polar coordinates $(x, y) = (r\cos\theta, r\sin\theta)$, and write

$$\partial_x = \cos\theta\,\partial_r - \frac{1}{r}\sin\theta\,\partial_\theta \quad\text{and}\quad \partial_y = \sin\theta\,\partial_r + \frac{1}{r}\cos\theta\,\partial_\theta.$$

In these coordinates, the Dirac operator becomes

$$\mathcal{D}_m = \begin{pmatrix} m & -\mathrm{i}\,\partial_x - \partial_y \\ -\mathrm{i}\,\partial_x + \partial_y & -m \end{pmatrix} = \begin{pmatrix} m & \mathrm{e}^{-\mathrm{i}\theta}(-\mathrm{i}\,\partial_r - \frac{1}{r}\,\partial_\theta) \\ \mathrm{e}^{\mathrm{i}\theta}\,(-\mathrm{i}\,\partial_r + \frac{1}{r}\,\partial_\theta) & -m \end{pmatrix}.$$

This suggests to decompose a spinor Ψ in Fourier modes with the convention

$$\Psi(r,\theta) = \sum_{n\in\mathbb{Z}} \begin{pmatrix} \varphi_n(r)\,\mathrm{e}^{\mathrm{i}n\theta} \\ \mathrm{i}\,\chi_n(r)\,\mathrm{e}^{\mathrm{i}(n+1)\theta} \end{pmatrix}.$$

If $\Phi := (\mathcal{D}_m - V)\,\Psi$ with corresponding Fourier modes $\big((\widetilde{\varphi}_n, \widetilde{\chi}_n)^\top\big)_{n\in\mathbb{Z}}$, then we have

$$\begin{pmatrix} \widetilde{\varphi}_n \\ \widetilde{\chi}_n \end{pmatrix} = (\mathcal{D}_m^{(n)} - V)\begin{pmatrix} \varphi_n \\ \chi_n \end{pmatrix}, \quad \text{with } \mathcal{D}_m^{(n)} = \begin{pmatrix} m & \partial_r + (n+1)/r \\ -\partial_r + n/r & -m \end{pmatrix}.$$

The operator $\mathscr{D}_m^{(n)}$ is self-adjoint in the Hilbert space $\mathrm{L}^2(\mathbb{R}^+ \times (0, 2\pi), r\,\mathrm{d}r\,\mathrm{d}\theta)$ because $(\partial_r)^* = -\partial_r - 1/r$. Let $\lambda_D^{(n)}(V)$ denote the lowest eigenvalue of $\mathscr{D}_m^{(n)} - V$ in the gap $(-m, m)$, and let

$$\Lambda_D^{\mathrm{rad},(n)}(\alpha, p) := \inf \left\{\lambda_D^{(n)}(V) \,:\, V \in \mathrm{L}^p(\mathbb{R}^2, \mathbb{R}^+),\ V \text{ radial and } \|V\|_p = \alpha\right\}$$

and

$$\Lambda_D^{\mathrm{rad}}(\alpha, p) := \inf_{n\in\mathbb{Z}} \Lambda_D^{\mathrm{rad},(n)}(\alpha, p).$$

We have the estimates

$$\Lambda_D(\alpha, p) \le \Lambda_D^{\mathrm{rad}}(\alpha, p) \le \Lambda_D^{\mathrm{rad},(0)}(\alpha, p). \tag{5.3}$$

A wavefunction $\Psi(r, \theta) = (\varphi(r)\,\mathrm{e}^{\mathrm{i}n\theta}, \mathrm{i}\,\chi(r)\,\mathrm{e}^{\mathrm{i}(n+1)\theta})^\intercal$ solves the nonlinear Dirac equation (3.5) if and only if

$$\begin{cases} \varphi' - \frac{n}{r}\,\varphi = -\left(\lambda + m + (|\chi|^2 + |\varphi|^2)^{1/(p-1)}\right)\chi, \\ \chi' + \frac{n+1}{r}\,\chi = \left(\lambda - m + (|\chi|^2 + |\varphi|^2)^{1/(p-1)}\right)\varphi. \end{cases} \tag{5.4}$$

This system with $n = 0$ is studied by W. Borrelli in [5]. It is an open question to decide whether $\Lambda_D^{\mathrm{rad}}(\alpha, p)$ is attained by $\Lambda_D^{\mathrm{rad},(n)}(\alpha, p)$ with $n = 0$ or not, and if equality holds in (5.3) so that $\Lambda_D(\alpha, p) = \Lambda_D^{\mathrm{rad},(0)}(\alpha, p)$. See Figure 3 for some numerical results.

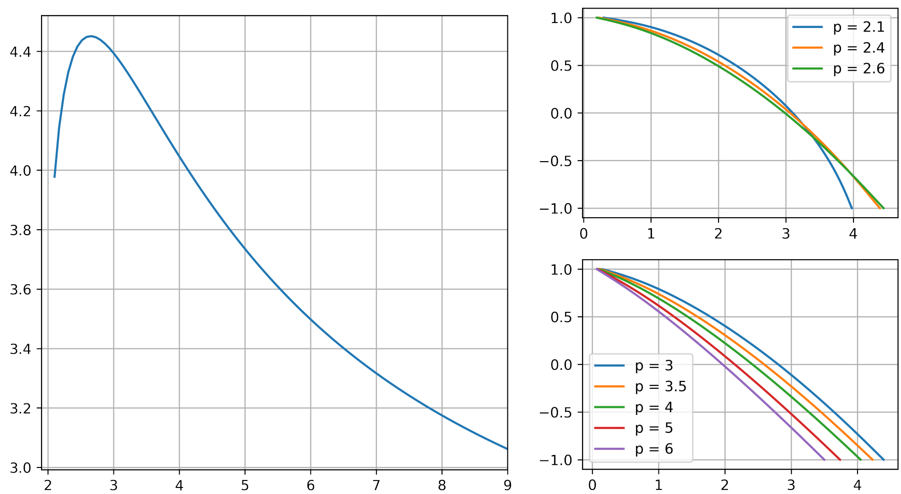

Figure 3. Radial case with $d = 2$ and $m = 1$. (Left) The function $p \mapsto \alpha_\star^{\mathrm{rad},(n=0)}(p)$ is an upper bound for $\alpha_\star(p)$ and reaches its maximum for $p \approx 2.66$. (Right) The maps $\alpha \mapsto \Lambda_D^{\mathrm{rad},(n=0)}(\alpha, p)$ for values of p corresponding either to $p < 2.66$ (top) or $p > 2.66$ (bottom). Numerically, the case $n = -1$ gives worse estimates.

5.3. The radial case in dimension $d = 3$

As in the two-dimensional case, we restrict the minimization problem (1.5) to radially symmetric decreasing potentials. The corresponding Dirac operator decomposes as a direct sum in eigenspaces of the *spin-orbit operator*

$$K = \beta\,(2\,S \cdot L + 1) = \beta\,(J^2 - \mathrm{L}^2 + 1/4), \quad \mathrm{spec}(K) = \pm 1, \pm 2, \ldots$$

and the *total angular momentum* in the z-direction J_3, with $\mathrm{spec}(J_3) = \frac{1}{2}\{1, 2, 3, \ldots\}$. See Section 4.6.4 of [68] for details. For any $\kappa \in \mathrm{spec}(K)$, we introduce the operator

$$\mathcal{D}_m^{(\kappa)} := \begin{pmatrix} m - V & \partial_r + \frac{\kappa+1}{r} \\ -\partial_r + \frac{\kappa-1}{r} & -m - V \end{pmatrix}$$

as a self-adjoint operator acting on $\mathrm{L}^2(\mathbb{R}^+, r^2\,\mathrm{d}r)$. If $\lambda_D^{(\kappa)}(V)$ denotes the lowest eigenvalue of $\mathcal{D}_m^{(\kappa)} - V$ in the gap $(-m, m)$, let us define

$$\Lambda_D^{\mathrm{rad},(\kappa)}(\alpha, p) := \inf\left\{\lambda_D^{(\kappa)}(V) \,:\, V \in \mathrm{L}^p(\mathbb{R}^3, \mathbb{R}^+),\ V \text{ radial and } \|V\|_p = \alpha\right\}.$$

We have $\Lambda_D^{\mathrm{rad}}(\alpha, p) = \inf_{\kappa \in \mathbb{Z}\setminus\{0\}} \Lambda_D^{\mathrm{rad},(\kappa)}(\alpha, p)$ and

$$\Lambda_D(\alpha, p) \le \Lambda_D^{\mathrm{rad}}(\alpha, p) \le \Lambda_D^{\mathrm{rad},(\kappa=1)}(\alpha, p).$$

It is an open question to decide whether the above inequalities are in fact equalities or not. If $\kappa = 1$, we look for an eigenstate of $\mathcal{D}_m - V$ in the Wakano form of [67], that is,

$$\Psi(r, \theta, \phi) = \begin{pmatrix} \varphi(r) \\ 0 \\ \mathrm{i}\,\chi(r)\cos\theta \\ \mathrm{i}\,e^{\mathrm{i}\phi}\,\chi(r)\sin\theta \end{pmatrix}$$

so that the nonlinear equation becomes

$$\begin{cases} \varphi' = -\left((\lambda + m) + (|\chi|^2 + |\varphi|^2)^{1/(p-1)}\right)\chi, \\ \chi' + \frac{2}{r}\,\chi = \left((\lambda - m) + (|\chi|^2 + |\varphi|^2)^{1/(p-1)}\right)\varphi. \end{cases} \tag{5.5}$$

The system (5.5) provides us with numerical upper estimates of $\Lambda_D(\alpha, p)$: see Figure 4.

5.4. An explicit bound in the radial case in dimensions $d = 2$ or $d = 3$

Let us assume that $m = 1$ and consider at $\lambda = -1$ (lower end of the gap) the system

$$\varphi' = -W\,\chi, \quad \chi' + \frac{\delta}{r}\,\chi = (W - 2)\,\varphi, \quad W^{p-1} = |\varphi|^2 + |\chi|^2. \tag{5.6}$$

According to the previous section, the radial case $d = 2$ corresponds to $\delta = 1$ (that is, $n = 0$ in (5.4)), and the radial case $d = 3$ to $\delta = 2$ (that is, $\kappa = 1$ in (5.5)). Writing $\chi(r) = f(r)\,\varphi(r)$, the system becomes

$$\varphi' = -W\,f\,\varphi, \quad f' = W\,(f^2 + 1) - \frac{\delta}{r}\,f - 2, \quad W^{p-1} = |\varphi|^2\,(1 + |f|^2).$$

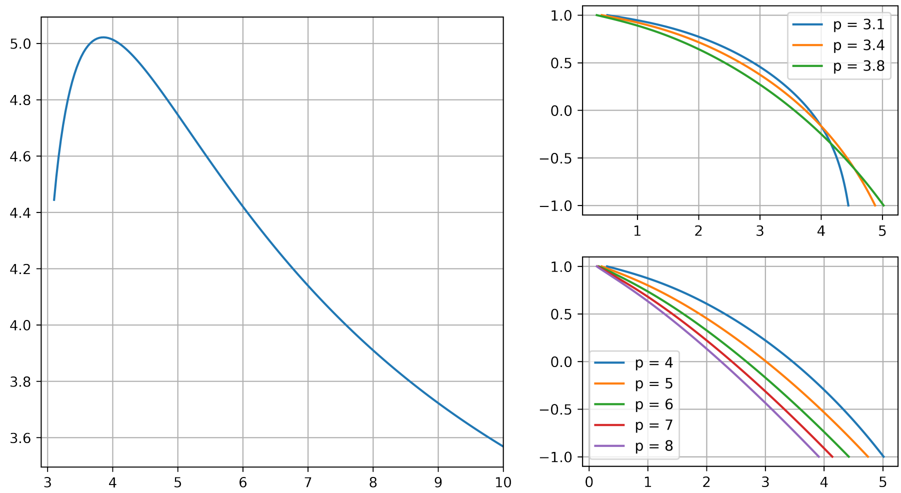

Figure 4. Radial case with $d = 3$ and $m = 1$. (Left) The function $p \mapsto \alpha_\star^{\mathrm{rad},(\kappa=1)}(p)$ reaches its maximum at $p \approx 3.86$. (Right) The maps $\alpha \mapsto \Lambda_D^{\mathrm{rad},(\kappa=1)}(\alpha, p)$ for values of p corresponding either to $p < 3.86$ (top) or $p > 3.86$ (bottom).

We now notice that this system admits a solution with $f(r) = r/\mu$ (so that all functions in the middle equality are constant functions). Explicitly, assuming $\delta < p - 1$, with $\mu := \frac{1}{2}(p - 1 - \delta)$, we find a solution of (5.6) given by

$$\varphi_p(r) = \frac{(p\mu)^{(p-1)/2}\,\mu}{(\mu^2 + r^2)^{p/2}}, \quad \chi_p(r) = \frac{(p\mu)^{(p-1)/2}\,r}{(\mu^2 + r^2)^{p/2}} \quad \text{and} \quad W_p(r) = \frac{p\mu}{\mu^2 + r^2}. \tag{5.7}$$

This solution is reminiscent of the solution of Corollary 1.4 in [9]. Up to a slight abuse of notations, we can consider W_p as a function of $x \in \mathbb{R}^d$ with $r = |x|$.

Lemma 5.1. *For all $p \geq d \geq 2$ or $p > 1$ if $d = 1$, and all $\delta < p - 1$, the potential W_p in (5.7), seen as a radial function in $\mathrm{L}^p(\mathbb{R}^d)$, satisfies*

$$\|W_p\|_p^p = p^p\,\pi^{d/2} \left(\frac{2}{p - 1 - \delta}\right)^{p-d} \frac{\Gamma(p - d/2)}{\Gamma(p)}, \tag{5.8}$$

so that in particular $\lim_{p \to d_+} \|W_p\|_p = d\,\sqrt{\pi}\,\left(\frac{\Gamma(d/2)}{\Gamma(d)}\right)^{1/d}$.

Applied either with $d = 2$ and $\delta = 1$, or $d = 3$ and $\delta = 2$, the expression (5.8) gives an upper bound for $\alpha_*(p)$ in dimension $d = 2$ and $d = 3$. We find that

$$\alpha_\star(p)^p \leq \frac{p^p}{p - 1}\left(\frac{2}{p - 2}\right)^{p-2} \pi \quad \text{if } d = 2,$$

$$\alpha_\star(p)^p \leq p^p \left(\frac{2}{p - 3}\right)^{p-3} \frac{\Gamma(p - 3/2)}{\Gamma(p)}\,\pi^{3/2} \quad \text{if } d = 3.$$

In particular,

$$\alpha_\star^{\mathrm{rad},(n=0)}(2) \le 2\sqrt{\pi} \approx 3.54491 \quad \text{if } d = 2,$$
$$\alpha_\star^{\mathrm{rad},(\kappa=1)}(3) \le 3(\pi/2)^{2/3} \approx 4.05385 \quad \text{if } d = 3.$$

The upper bound given by this expression for $d = 1$ (with $\delta = 0$) coincides with the expression found in Theorem 1.3, and we conjecture that we actually have equality in $d = 2$ and $d = 3$ as well. Numerically, the curve $p \mapsto \|W_p\|_p$ coincides with the numerical solution $p \mapsto \alpha_\star^{\mathrm{rad},(n=0)}(p)$ if $d = 2$ and $p \mapsto \alpha_\star^{\mathrm{rad},(\kappa=1)}(p)$ if $d = 3$ of Figures 3 and 4. It is however an open question to decide whether φ_p, χ_p and W_p is the unique solution of (5.6) and if it is optimal among radial optimal functions, and also among non-radial optimal functions (see Appendix B).

A. Open questions

In this article, we study the *ground state* defined the *lowest eigenvalue in the gap* $\lambda_D(V)$ of a general Dirac operator $\not{D}_m - V$ with $V \in \mathrm{L}^p(\mathbb{R}^d, \mathbb{R}^+)$ using Birman–Schwinger techniques, and prove that this quantity always makes sense if the $\mathrm{L}^p(\mathbb{R}^d)$ norm of V is small enough. To our knowledge, there are several open questions concerning this lowest eigenvalue, which we recall here.

- Is the map $V \mapsto \lambda_D(V)$ concave?
- Is $\lambda_D(V)$ always a simple eigenvalue, or equivalently, is $\mu_1(K_V)$ always simple?

Assuming that the answer of the last question is positive, we denote by Ψ the corresponding eigenfunction for the Dirac operator. We decompose it as $\Psi = \Psi_+ + \Psi_-$, with $\beta\Psi_+ = \Psi_+$ (upper component) and $\beta\Psi_- = -\Psi_-$ (lower component).

- If V is radial (decreasing), is Ψ_+ also radial (decreasing)?

Concerning the variational problem associated with (1.5), we recall two questions that were already raised earlier:

- Is the optimal potential V radial (decreasing) if $d \ge 2$?
- If so, is the corresponding *ground state* Ψ the solution with lowest angular momentum and smallest number of oscillations, as it is suggested in Sections 5.2 and 5.3?

B. Is the optimal potential radial? A numerical answer

In dimension $d = 2$, we investigate numerically whether the optimal potential V for (1.5) is radial, or equivalently whether the optimal potential W for (3.1) is radial. In order to do so, we run the following self-consistent algorithm[1]. Recall that $K_W := \sqrt{W}\, R_0(\lambda) \sqrt{W}$, where R_0 denotes the resolvent of the free Dirac operator. For $p > d = 2$ and $\lambda \in [-m, m)$, we choose an initial potential W_0 at random, and set

$$\phi_k := \text{normalized eigenvector corresponding to } \mu_1(K_{W_k}) \quad \text{and} \quad W_{k+1} := |\phi_k|^{2/p}.$$

[1]The code is available upon request to the authors.

In practice, the potential W_{k+1} is also translated so that its maximum is at the origin. We can check that the quantity $\mu_1(K_{W_k})$ is increasing, and that the sequence $(W_k)_{k\in\mathbb{N}}$ converges to some limit potential W_* in $\mathrm{L}^p(\mathbb{R}^2)$. A typical run of the algorithm is displayed in Figure 5.

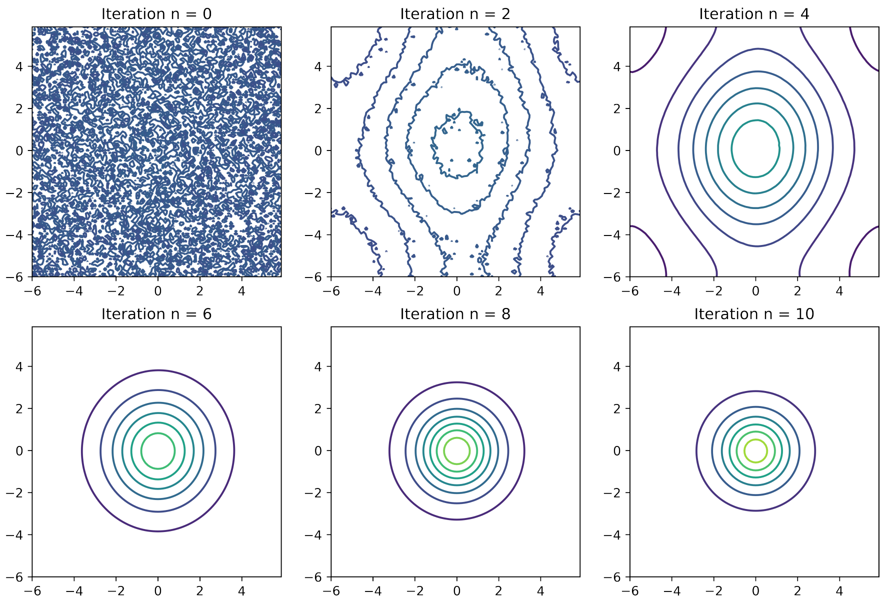

Figure 5. Contour lines of the potential W_k during the iterations, for $p = 3$ and $\lambda = 1/2$, for some W_0 chosen at random. The quantities W_k and ϕ_k are computed on a square $[-a, a]^2$ with $a = 6$, $L = 100$ discretization points per direction and periodic boundary conditions. The Dirac operator and its inverse are computed in Fourier space and the $\mathrm{L}^p(\mathbb{R}^2)$ integrals in direct space.

In order to check whether W_* is radial or not, we compute the $\mathrm{L}^p(\mathbb{R}^2)$ norm of its angular derivative. For $\lambda \in [-0.9, 0.9]$, $m = 1$ and $p \in (2, 8)$, this norm is always much smaller than 1 and usually of the order of 10^{-2} or 10^{-3}, after less than 100 iterations, depending on the parameters we chose. These numerical results suggest that the optimal potentials might be radial, up to translations.

C. A nonlinear interpolation inequality for the Dirac operator

C.1. Non-relativistic limit and Keller–Lieb–Thirring inequalities

In order to consider the *non-relativistic limit* $c \to +\infty$, it is interesting to reintroduce the parameters $\hbar$, m and c. The eigenvalue problem

$$(\mathcal{D}_m^{\hbar,c} - W)\psi = \mu\,\psi, \quad \text{where } \mathcal{D}_m^{\hbar,c} := -\mathrm{i}\,\hbar\,c\,\boldsymbol{\alpha}\cdot\nabla + mc^2\beta,$$

is reduced to the eigenvalue problem corresponding to $\hbar = c = m = 1$ by the change of variables

$$\psi(x) = \Psi\Big(\frac{mc}{\hbar}\,x\Big), \quad W(x) = mc^2\,V\Big(\frac{mc}{\hbar}\,x\Big) \quad \text{and} \quad \mu = mc^2\lambda.$$

As a consequence, the *ground state* $\lambda_D^{\hbar,c}(W)$ of $\mathcal{D}_m^{\hbar,c} - W$, defined as its lowest eigenvalue in the gap $(-mc^2, mc^2)$, and estimated by $\lambda_D(V) \ge \Lambda_D^{(m=1)}(\|V\|_p, p)$ according to the *Keller–Lieb–Thirring inequality for the Dirac operator* (1.8), becomes

$$\text{(C.1)} \qquad \lambda_D^{\hbar,c}(W) \ge mc^2\,\Lambda_D^{(m=1)}\big(\hbar^{-d/p}\,m^{d/p-1}\,c^{d/p-2}\,\|W\|_p, p\big)$$

using the above change of variables. Here $\Lambda_D^{(m=1)}$ stands for Λ_D when we assume $m = 1$ in notations of Theorem 1.1.

Proposition C.1. *Let either $d \ge 1$ and $p > 1$ if $d = 1$, or $p \ge d$ if $d \ge 2$. We have, with $\Lambda_D^{(m=1)}(\alpha, p)$ defined by (1.5), $\eta = 2p/(2p-d)$ and K_p as in (1.2),*

$$1 - \Lambda_D^{(m=1)}(\alpha, p) = 2^{d/(2p-d)}\,\mathsf{K}_p\,\alpha^\eta\,(1 + o(1)) \quad \text{as } \alpha \to 0_+.$$

If $d = 1$, we obtain that

$$\mathsf{K}_p = \big(p^p\,(p-1)^{-(p-1)}\,B(\tfrac{1}{2}, p)\big)^{-2/(2p-1)}$$

by expanding the expression of $\alpha_D(\lambda, p)$ given in Theorem 1.3 as $\lambda \to 1_-$. This is consistent with $\mathsf{K}_p = \mathsf{C}_q^{-\eta}$ and the expression of the explicit, optimal value of the constant C_q in (1.1) if the dimension is $d = 1$; we refer to [23] and references therein for details.

Proof. Let us consider the general case $d \ge 1$. The non-relativistic limit of the ground state $\lambda_D^{\hbar,c}(W)$ of the Dirac operator $\mathcal{D}_m^{\hbar,c} - W$ is, up to the mass energy mc^2, given by the ground state of the Schrödinger operator

$$-\frac{\hbar^2}{2m}\,\Delta - W$$

by standard results: see for instance [32], Section 2.4. Hence

$$\lim_{c\to+\infty}(mc^2 - \lambda_D^{\hbar,c}(W)) = \lambda_S^-(W_\mu), \quad \text{where } W_\mu(x) := W(\mu\,x) \text{ and } \mu = \frac{\hbar}{\sqrt{2m}}\cdot$$

Here $-\lambda_S^-(W_\mu)$ denotes, if it exists, the negative ground state of the Schrödinger operator $-\Delta - W_\mu$. The factor $\mu = \hbar/\sqrt{2m}$ arises from a scaling argument. By the definition in (1.5), we obtain

$$\lim_{c\to+\infty} mc^2\big(1 - \Lambda_D^{(m=1)}\,(\hbar^{-d/p}\,m^{d/p-1}\,c^{d/p-2}\,\|W\|_p, p)\big) \le \mathsf{K}_p\,\|W_\mu\|_p^\eta$$
$$= \mathsf{K}_p\,\mu^{-d\eta/p}\,\|W\|_p^\eta,$$

but there is in fact equality if we use as test function an optimal function W for (1.2). Taking $\alpha = \hbar^{-d/p}\,m^{d/p-1}\,c^{d/p-2}\,\|W\|_p$ in the limit as $c \to +\infty$ concludes the proof. ∎

Proposition C.1 is in fact equivalent to

$$\lim_{c\to+\infty}\big(mc^2-\lambda_D^{\hbar,c}(W)\big)\le \mathsf{K}_p\Big(\frac{2m}{\hbar^2}\Big)^{d/(2p-d)}\,\|W\|_p^\eta, \tag{C.2}$$

written with the physical constants. In other words, we recover a standard *Keller–Lieb–Thirring inequality for the Schrödinger operator* (1.2) in the non-relativistic limit. In dimension $d=1$, a tedious but elementary computation directly shows that the constant obtained by taking the non-relativistic limit in the Keller–Lieb–Thirring inequality for the Dirac operator written with optimal constant is the optimal constant in the Keller–Lieb–Thirring inequality for the Schrödinger operator, as it can be deduced for instance from [23, 45].

The definition (1.7) can be generalized to the case $(\hbar,c)\neq(1,1)$ using the monotonicity of $\alpha\mapsto\Lambda_D^{(m=1)}(\alpha,p)$ stated in Theorem 1.1 and (C.1). If $\alpha_D^{(m=1)}$ denotes the inverse of $\alpha\mapsto\Lambda_D^{(m=1)}(\alpha,p)$, the condition

$$\|W\|_p\le\alpha_D^{\hbar,c}(\lambda,p):=\hbar^{d/p}\,m^{1-d/p}\,c^{2-d/p}\,\alpha_D^{(m=1)}\Big(\frac{\lambda}{mc^2},p\Big) \tag{C.3}$$

guarantees that $\lambda_D^{\hbar,c}(W)\ge\lambda$. Notice that $p\ge d$ implies that

$$\lim_{c\to\infty}\|W\|_p\le\alpha_D^{\hbar,c}(\lambda,p)=\infty.$$

C.2. An interpolation inequality for the Dirac operator

Using a min-max principle as in [28], it is possible to write an optimal interpolation inequality of Gagliardo–Nirenberg–Sobolev type which plays for the free Dirac operator the same role as (1.1). The inequality is somewhat involved, but inequality (1.1) is recovered in the non-relativistic limit as $c\to+\infty$. For sake of simplicity, we consider only the case $d=1$.

Let us start by a short and formal summary of the *min-max principle* applied to the determination of the ground state of the Dirac operator. If $(\varphi,\chi)^\mathsf{T}$ is an eigenspinor of the operator $\mathcal{D}_m^{\hbar,c}-V$ with eigenvalue $\lambda\in(-mc^2,mc^2)$, then, as in (5.1) we have

$$\begin{cases}\hbar c\varphi'=-(\lambda+mc^2+V)\,\chi,\\ \hbar c\chi'=(\lambda-mc^2+V)\,\varphi.\end{cases}$$

The first line gives

$$\chi=-\hbar c\,\frac{\varphi'}{\lambda+mc^2+V},$$

so that the problem amounts to solving

$$-(\hbar c)^2\Big(\frac{\varphi'}{\lambda+mc^2+V}\Big)'+(mc^2-\lambda-V)\,\varphi=0.$$

Multiplying by φ and integrating suggests to introduce the functional

$$\mathcal{E}[\mu,V,\phi]:=(\hbar c)^2\int_{\mathbb{R}}\frac{|\phi'|^2}{\mu+mc^2+V}\,\mathrm{d}x+\int_{\mathbb{R}}(mc^2-\mu-V)\,|\phi|^2\,\mathrm{d}x.$$

Clearly, we have $\mathcal{E}[\lambda, V, \varphi] = 0$. In addition, for all fixed V and ϕ, the map $\mu \mapsto \mathcal{E}[\mu, V, \phi]$ is decreasing. It is proved in Lemma 2.4 of [63] that for all $-m < \mu < \lambda_D^{\hbar,c}(V)$, the quadratic map $\phi \mapsto \mathcal{E}[\mu, V, \phi]$ is positive definite, and that for $\mu = \lambda_D(V)$, we have $\mathcal{E}[\mu, V, \phi] = 0$ if and only if $\phi = \varphi$, up to a multiplicative constant. In particular, we have

$$\forall \phi \in C_0^\infty(\mathbb{R}),\ \forall V \in \mathrm{L}^p(\mathbb{R}), \quad \|V\|_p \le \alpha_D^{\hbar,c}(\lambda, p) \implies \mathcal{E}[\lambda, V, \phi] \ge 0.$$

Minimizing $\mathcal{E}[\lambda, W, \phi]$ in W such that $\|W\|_p = \alpha \le \alpha_D^{\hbar,c}(\lambda, p)$ shows that the optimal W solves the Euler–Lagrange equation of the implicit form

$$\nu\, W^{p-1} = |\phi|^2 + \frac{(\hbar c)^2\, |\phi'|^2}{(\lambda + mc^2 + W)^2}, \tag{C.4}$$

where $\nu \ge 0$ is now the Lagrange multiplier for the constraint $\|W\|_p = \alpha$. Note that for all fixed $a, b, c \ge 0$, the equation

$$\nu\, X^{p-1} = a + \frac{b}{(c + X)^2}$$

has a unique solution in $X_\nu \ge 0$, as the left-hand side is an increasing function of X, while the right-hand side is decreasing. We also learn that $\nu \mapsto X_\nu$ is increasing. So for fixed $\nu \ge 0$, there is a unique $W = V_\nu[\phi]$ satisfying (C.4) and the map $\nu \mapsto V_\nu$ is pointwise decreasing, hence so is the map $\nu \mapsto \|V_\nu\|_p$. With $\alpha_D^{\hbar,c}(\lambda, p)$ given by (C.3), we define

$$\nu_*(\lambda, p, \phi) := \inf\left\{\nu > 0 \,:\, \|V_\nu[\phi]\|_p \le \alpha_D^{\hbar,c}(\lambda, p)\right\}.$$

Summarizing, we proved that for all $\phi \in C_0^\infty(\mathbb{R})$ and all $\nu \ge \nu_*(\lambda, p, \phi)$,

$$(\hbar c)^2 \int_{\mathbb{R}} \frac{|\phi'|^2}{\lambda + mc^2 + V_\nu[\phi]}\, \mathrm{d}x + \int_{\mathbb{R}} (mc^2 - \lambda - V_\nu[\phi])\, |\phi|^2\, \mathrm{d}x \ge 0, \tag{C.5}$$

which can be interpreted as a *Gagliardo–Nirenberg type inequality* for ϕ alone. Such an inequality is known for a fixed, given potential V from [21, 27, 28] and it is then of Hardy-type, as for instance the new Hardy inequality in [35], but the novelty in this paper is that we take $V = V_\nu[\phi]$ thus making it a nonlinear interpolation inequality. While the form (C.5) is non-explicit, it allows to recover the usual Gagliardo–Nirenberg inequality in the non-relativistic limit as $c \to \infty$. By writing $\lambda = mc^2 + E$ for some $E < 0$, (C.5) becomes

$$(\hbar c)^2 \int_{\mathbb{R}} \frac{|\phi'|^2}{2mc^2 + E + V_\nu[\phi]}\, \mathrm{d}x - \int_{\mathbb{R}} (E + V_\nu[\phi])\, |\phi|^2\, \mathrm{d}x \ge 0.$$

Let us choose $\nu = \|\phi\|_{2p/(p-1)}^2$. As $c \to \infty$, we get from (C.4) that $V_\nu[\phi]$ converges to $\|\phi\|_{2p/(p-1)}^{-2/(p-1)}\, |\phi|^{2/(p-1)}$. Together with (C.2), we get that $\nu \ge \nu_*(\lambda, p, \phi)$ in the limit $c \to \infty$ whenever $|E| \ge \mathsf{K}_p\, (2m/\hbar^2)^{d/(2p-d)}$. We obtain

$$\frac{\hbar^2}{2m} \int_{\mathbb{R}} |\phi'|^2\, \mathrm{d}x - \|\phi\|_{2p/(p-1)}^2 \ge E \int_{\mathbb{R}} |\phi|^2\, \mathrm{d}x.$$

This inequality is the Gagliardo–Nirenberg inequality (1.1) written in non-scale invariant form, for an appropriate choice of the parameter λ in (1.1).

D. The case $p = d = 1$

This appendix deals with the limit case $p = 1$ of Theorem 1.3 devoted to the one-dimensional Keller estimates. We give a computation of $\alpha_D(\lambda, 1)$ which is not based on the limit as $p \to 1_+$ of the nonlinear estimates and prove that any sequence of optimizing potentials concentrates into a Dirac δ distribution.

Proposition D.1. *If $d = 1$, then $\alpha_D(\lambda, 1) = \arccos(\lambda/m)$. More specifically, for all $\alpha \in (0, \pi)$ and all $V \in L^1(\mathbb{R}, \mathbb{R}^+)$ with $\|V\|_1 = \alpha$, if $\lambda \in (-m, m)$ is an eigenvalue of $\mathcal{D}_m - V$, then we have the strict inequality*

$$m \cos\alpha < \lambda.$$

In addition, any sequence of nonnegative potentials $(V_n)_{n\in\mathbb{N}}$ with $\|V_n\|_1 = \alpha$, and eigenvalues λ_n approaching $m \cos\alpha$, converges as $n \to +\infty$ to a Dirac δ distribution.

According to [65], "the method of directly solving the Dirac equation with a δ-function potential and the method of obtaining the solution by first solving the Dirac equation with a short-range potential and afterward taking the δ-function limit, lead to different results" [concerning the spectrum]. This issue is known as Klein's paradox. Although the Keller–Lieb–Thirring inequality (1.12) makes sense for any nonnegative potential $V \in L^1(\mathbb{R}^d)$, it is a natural question to investigate by direct methods whether the bound is achieved in the larger set of bounded nonnegative measures and consider sequences of optimizing potentials.

Proof. We start with a calculation for a bounded and compactly supported potential V. In this case, the eigenvalue equation rewrites as

$$\Psi' = (\mathrm{i}\sigma_2 (V + \lambda) + m\sigma_1)\,\Psi.$$

We decompose Ψ on the (not-orthonormal) basis given by the eigenvectors $e_\pm$ of the matrix $\mathrm{i}\lambda\sigma_2 - m\sigma_1$ defined by

$$e_\pm := \begin{pmatrix} \sqrt{m^2-\lambda^2} \\ \pm(m-\lambda) \end{pmatrix} \quad \text{such that} \quad \begin{pmatrix} 0 & m+\lambda \\ m-\lambda & 0 \end{pmatrix} e_\pm = \pm\sqrt{m^2-\lambda^2}\, e_\pm.$$

Decomposing $\Psi(x) = \mathsf{a}(x)\, e_+ + \mathsf{b}(x)\, e_-$ and using the identities

$$\langle \mathrm{i}\sigma_2\, e_\pm, e_\pm\rangle = 0, \qquad \langle \mathrm{i}\sigma_2\, e_\pm, e_\mp\rangle = \pm 2\,(m-\lambda)\sqrt{m^2-\lambda^2},$$
$$\langle \mathrm{i}\sigma_2\, e_\pm, \mathrm{i}\sigma_2\, e_\pm\rangle = 2m\,(m-\lambda), \qquad \langle \mathrm{i}\sigma_2\, e_\pm, \mathrm{i}\sigma_2\, e_\mp\rangle = 2\lambda\,(m-\lambda),$$

gives, with $\mathsf{W} := V/\sqrt{m^2-\lambda^2}$,

$$\mathsf{a}' = \left(\sqrt{m^2-\lambda^2} - \lambda\,\mathsf{W}\right)\mathsf{a} - m\,\mathsf{W}\,\mathsf{b},$$
$$\mathsf{b}' = -\left(\sqrt{m^2-\lambda^2} - \lambda\,\mathsf{W}\right)\mathsf{b} + m\,\mathsf{W}\,\mathsf{a}.$$

Since V (and W) are compactly supported, a square-integrable solution must have $\mathsf{b} = 0$ in a neighborhood of $-\infty$ and $\mathsf{a} = 0$ in a neighborhood of $+\infty$. Without loss of generality, we take a solution with $\mathsf{a}(x) > 0$ for x near $-\infty$. Since

$$(\mathsf{a}\mathsf{b})' = m\,\mathsf{W}\,(\mathsf{a}^2 - \mathsf{b}^2)$$

is nonnegative if $|\mathsf{a}| > |\mathsf{b}|$, such a solution enters the first (a, b) quadrant and stays in the first quadrant until the first value of x such that $\mathsf{a}(x) = 0$. We denote this value by x_1, with $x_1 = +\infty$ if a does not change sign. In the interval $(-\infty, x_1)$, the ratio $\mathsf{t} := \mathsf{b}/\mathsf{a}$ is well-defined and satisfies

$$\begin{aligned}\mathsf{t}' &= \frac{1}{\mathsf{a}^2}\Big(-2\,\mathsf{ab}\big(\sqrt{m^2-\lambda^2} - \lambda\,\mathsf{W}\big) + m\,\mathsf{W}(\mathsf{a}^2+\mathsf{b}^2)\Big)\\ &= -2\sqrt{m^2-\lambda^2}\,\mathsf{t} + \mathsf{W}(m + m\,\mathsf{t}^2 + 2\,\lambda\,\mathsf{t}).\end{aligned}$$

We finally define the angle

$$\theta_\lambda(\mathsf{t}) := \arctan\Big(\frac{m\,\mathsf{t}+\lambda}{\sqrt{m^2-\lambda^2}}\Big),$$

such that $\lim_{x\to x_1}\theta_\lambda(\mathsf{t}(x)) = \pi/2$ and

$$\text{(D.1)}\qquad \frac{1}{\sqrt{m^2-\lambda^2}}\,(\theta_\lambda\circ\mathsf{t})' = \frac{\mathsf{t}'}{m+2\,\lambda\,\mathsf{t}+m\,\mathsf{t}^2} = \mathsf{W} - 2\,\frac{\sqrt{m^2-\lambda^2}\,\mathsf{t}}{m+2\,\lambda\,\mathsf{t}+m\,\mathsf{t}^2}.$$

Integrating for $x \in (-\infty, x_1)$, we obtain

$$\text{(D.2)}\qquad \frac{\pi/2-\theta_\lambda(0)}{\sqrt{m^2-\lambda^2}} = \int_{-\infty}^{x_1}\mathsf{W}(s)\,\mathrm{d}s - 2\int_{-\infty}^{x_1}\frac{\sqrt{m^2-\lambda^2}\,\mathsf{t}}{m+2\,\lambda\,\mathsf{t}+m\,\mathsf{t}^2}\,\mathrm{d}s < \frac{\alpha}{\sqrt{m^2-\lambda^2}}.$$

Since $\theta_\lambda(0) = \arcsin(\lambda/m)$, we obtain

$$\arccos(\lambda/m) < \alpha \quad \text{or} \quad \lambda > m\cos\alpha.$$

In order to approximate unbounded potentials, we need an estimate on the negative term in (D.2). Take any number $c > 1$. Since t is continuous, there is an interval $I_V(c) \subset (-\infty, x_1]$ such that $\mathsf{t}(x) \in (1/c, c)$ for all $x \in I_V(c)$. We have the bound (note that the integrand is symmetric under $\mathsf{t} \mapsto 1/\mathsf{t}$)

$$\int_{I_V(c)}\frac{2\,(m^2-\lambda^2)\,\mathsf{t}(s)}{m+m\,\mathsf{t}(s)^2+2\,\lambda\,\mathsf{t}(s)}\,\mathrm{d}s \geq \frac{2\,(m^2-\lambda^2)\,c}{m+mc^2+2\,\lambda\,c}\,|I_V(c)|,$$

and therefore

$$\text{(D.3)}\qquad \arccos(\lambda/m) \leq \alpha - \frac{2\,(m^2-\lambda^2)\,c}{m+mc^2+2\,\lambda\,c}\,|I_V(c)|.$$

To prove that $|I_V(c)|$ cannot be arbitrarily small, we integrate (D.1) on $I_V(c)$, which gives

$$\text{(D.4)}\qquad \theta_\lambda(c) - \theta_\lambda(1/c) \leq \int_{I_V(c)} V(s)\,\mathrm{d}s \leq Q_V(|I_V(c)|),$$

where we have defined

$$Q_V(r) := \sup_{x\in\mathbb{R}}\int_{x-r/2}^{x+r/2} V(s)\,\mathrm{d}s.$$

Now, assume that $(V_n)_{n\in\mathbb{N}}$ is a sequence of potentials with $\|V_n\|_1 = \alpha$ and eigenvalues λ_n converging to $\lambda := m\cos\alpha$. Without loss of generality, we may assume that each V_n is bounded and compactly supported. By (D.3), in order to approach the equality case, we need that $|I_{V_n}(c)|$ tends to zero for each $c > 1$. We now use (D.4) to show that this implies the convergence (after suitable translations) to a Dirac δ distribution.

Fix $\varepsilon > 0$ and $r > 0$. Fix $c > 1$ and n_0 such that for all $n \geq n_0$, we have

$$\theta_{\lambda_n}(c) - \theta_{\lambda_n}(1/c) \geq \theta_\lambda(+\infty) - \theta_\lambda(0) - \varepsilon = \alpha - \varepsilon.$$

Upon increasing n_0, we can assume $|I_{V_n}(c)| \leq r$ for all $n \geq n_0$. From (D.4), this gives

$$Q_{V_n}(r) \geq Q_{V_n}(|I_{V_n}(c)|) \geq \alpha - \varepsilon.$$

Since r and ε are arbitrary, we have shown that Q_{V_n} converges pointwise to α. In the language of concentration-compactness, this excludes *vanishing* and *dichotomy* and implies that, after a sequence of translations, V_n converges to a measure of total mass α supported at the origin, hence, to a Dirac δ distribution. ∎

Acknowledgments. The authors thank the referees for their suggestions, which have all been taken into account and led to a significant improvement of the paper.

Funding. This work was partially supported by the Project EFI (ANR-17-CE40-0030) of the French National Research Agency (ANR). HVDB received funding from the Center for Mathematical Modeling (Universidad de Chile and CNRS IRL 2807) through ANID/Basal projects #FB210005 and #ACE210010, as well as ANID/Fondecyt project #11220194, and MathAmSud project EEQUADDII 20-MATH-04. This work was partially developed when FP was employed at CNRS and CEREMADE-Université Paris-Dauphine, and supported by the project ANR-17-CE29-0004 molQED of the ANR and by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement MDFT no. 725528). He is also supported by the project PID2021-123034NB-I00 funded by MCIN/AEI/10.13039/501100011033/ FEDER, UE.

References

[1] Abramowitz, M. and Stegun, I. A.: *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. Tenth printing, with corrections. National Bureau of Standards. A Wiley-Interscience Publication, New York, 1972.

[2] Alama, S., Deift, P. A. and Hempel, R.: Eigenvalue branches of the Schrödinger operator $H - \lambda W$ in a gap of $\sigma(H)$. *Comm. Math. Phys.* **121** (1989), no. 2, 291–321.

[3] Arrizabalaga, N., Duoandikoetxea, J. and Vega, L.: Self-adjoint extensions of Dirac operators with Coulomb type singularity. *J. Math. Phys.* **54** (2013), no. 4, article no. 041504, 20 pp. [Erratum. *J. Math. Phys.* **59** (2018), no. 7, article no. 079902, 2 pp.]

[4] Balabane, M., Cazenave, T., Douady, A. and Merle, F.: Existence of excited states for a nonlinear Dirac field. *Comm. Math. Phys.* **119** (1988), no. 1, 153–176.

[5] Borrelli, W.: Stationary solutions for the 2D critical Dirac equation with Kerr nonlinearity. *J. Differential Equations* **263** (2017), no. 11, 7941–7964.

[6] Borrelli, W.: Multiple solutions for a self-consistent Dirac equation in two dimensions. *J. Math. Phys.* **59** (2018), no. 4, article no. 041503, 13 pp.

[7] Borrelli, W.: Weakly localized states for nonlinear Dirac equations. *Calc. Var. Partial Differential Equations* **57** (2018), no. 6, article no. 155, 21 pp.

[8] Borrelli, W. and Frank, R. L.: Sharp decay estimates for critical Dirac equations. *Trans. Amer. Math. Soc.* **373** (2020), no. 3, 2045–2070.

[9] Borrelli, W., Malchiodi, A. and Wu, R.: Ground state Dirac bubbles and Killing spinors. *Comm. Math. Phys.* **383** (2021), no. 2, 1151–1180.

[10] Bosi, R., Dolbeault, J. and Esteban, M. J.: Estimates for the optimal constants in multipolar Hardy inequalities for Schrödinger and Dirac operators. *Commun. Pure Appl. Anal.* **7** (2008), no. 3, 533–562.

[11] Brezis, H.: *Functional analysis, Sobolev spaces and partial differential equations*. Universitext, Springer, New York, 2011.

[12] Carlen, E. A., Frank, R. L. and Lieb, E. H.: Stability estimates for the lowest eigenvalue of a Schrödinger operator. *Geom. Funct. Anal.* **24** (2014), no. 1, 63–84.

[13] Cassano, B. and Pizzichillo, F.: Self-adjoint extensions for the Dirac operator with Coulomb-type spherically symmetric potentials. *Lett. Math. Phys.* **108** (2018), no. 12, 2635–2667.

[14] Cassano, B., Pizzichillo, F. and Vega, L.: A Hardy-type inequality and some spectral characterizations for the Dirac–Coulomb operator. *Rev. Mat. Complut.* **33** (2020), no. 1, 1–18.

[15] Cazenave, T. and Vázquez, L.: Existence of localized solutions for a classical nonlinear Dirac field. *Comm. Math. Phys.* **105** (1986), no. 1, 35–47.

[16] Coffman, C. V.: Uniqueness of the ground state solution for $\Delta u - u + u^3 = 0$ and a variational characterization of other solutions. *Arch. Rational Mech. Anal.* **46** (1972), 81–95.

[17] Cuenin, J.-C.: Eigenvalue bounds for Dirac and fractional Schrödinger operators with complex potentials. *J. Funct. Anal.* **272** (2017), no. 7, 2987–3018.

[18] Cuenin, J.-C., Laptev, A. and Tretter, C.: Eigenvalue estimates for non-selfadjoint Dirac operators on the real line. *Ann. Henri Poincaré* **15** (2014), no. 4, 707–736.

[19] Daubechies, I.: An uncertainty principle for fermions with generalized kinetic energy. *Comm. Math. Phys.* **90** (1983), no. 4, 511–520.

[20] Dolbeault, J., Esteban, M. and Loss, M.: Characterization of the critical magnetic field in the Dirac–Coulomb equation. *J. Phys. A* **41** (2008), no. 18, article no. 185303, 13 pp.

[21] Dolbeault, J., Esteban, M. J., Duoandikoetxea, J. and Vega, L.: Hardy-type estimates for Dirac operators. *Ann. Sci. École Norm. Sup. (4)* **40** (2007), no. 6, 885–900.

[22] Dolbeault, J., Esteban, M. J. and Laptev, A.: Spectral estimates on the sphere. *Anal. PDE* **7** (2014), no. 2, 435–460.

[23] Dolbeault, J., Esteban, M. J., Laptev, A. and Loss, M.: One-dimensional Gagliardo–Nirenberg–Sobolev inequalities: remarks on duality and flows. *J. Lond. Math. Soc. (2)* **90** (2014), no. 2, 525–550.

[24] Dolbeault, J., Esteban, M. J., Laptev, A. and Loss, M.: Interpolation inequalities and spectral estimates for magnetic operators. *Ann. Henri Poincaré* **19** (2018), no. 5, 1439–1463.

[25] Dolbeault, J., Esteban, M. J. and Loss, M.: Relativistic hydrogenic atoms in strong magnetic fields. *Ann. Henri Poincaré* **8** (2007), no. 4, 749–779.

[26] Dolbeault, J., Esteban, M. J. and Loss, M.: Critical magnetic field for 2D magnetic Dirac–Coulomb operators and Hardy inequalities. In *Partial differential equations, spectral theory, and mathematical physics – the Ari Laptev anniversary volume*, pp. 41–63. EMS Ser. Congr. Rep., EMS Press, Berlin, 2021.

[27] Dolbeault, J., Esteban, M. J., Loss, M. and Vega, L.: An analytical proof of Hardy-like inequalities related to the Dirac operator. *J. Funct. Anal.* **216** (2004), no. 1, 1–21.

[28] Dolbeault, J., Esteban, M. J. and Séré, E.: On the eigenvalues of operators with gaps. Application to Dirac operators. *J. Funct. Anal.* **174** (2000), no. 1, 208–226.

[29] Dolbeault, J., Esteban, M. J. and Séré, E.: General results on the eigenvalues of operators with gaps, arising from both ends of the gaps. Application to Dirac operators. *J. Eur. Math. Soc. (JEMS)* **8** (2006), no. 2, 243–251.

[30] Dolbeault, J., Esteban, M. J. and Séré, E.: Corrigendum to: "On the eigenvalues of operators with gaps. Application to Dirac operators" [J. Funct. Anal. 174 (1) (2000) 208–226]. *J. Funct. Anal.* **284** (2023), no. 1, article no. 109651, 6 pp.

[31] Dolbeault, J., Esteban, M. J. and Séré, E.: Distinguished self-adjoint extension and eigenvalues of operators with gaps. Application to Dirac–Coulomb operators. *J. Spectral Theory* **13** (2023), no. 2, 491–524.

[32] Esteban, M. J., Lewin, M. and Séré, E.: Variational methods in relativistic quantum mechanics. *Bull. Amer. Math. Soc. (N.S.)* **45** (2008), no. 4, 535–593.

[33] Esteban, M. J., Lewin, M. and Séré, É.: Domains for Dirac–Coulomb min-max levels. *Rev. Mat. Iberoam.* **35** (2019), no. 3, 877–924.

[34] Esteban, M. J., Lewin, M. and Séré, É.: Dirac–Coulomb operators with general charge distribution I. Distinguished extension and min-max formulas. *Ann. H. Lebesgue* **4** (2021), 1421–1456.

[35] Esteban, M. J., Lewin, M. and Séré, É.: Dirac–Coulomb operators with general charge distribution II. The lowest eigenvalue. *Proc. Lond. Math. Soc. (3)* **123** (2021), no. 4, 345–383.

[36] Esteban, M. J. and Loss, M.: Self-adjointness for Dirac operators via Hardy–Dirac inequalities. *J. Math. Phys.* **48** (2007), no. 11, article no. 112107, 8 pp.

[37] Esteban, M. J. and Loss, M.: Self-adjointness via partial Hardy-like inequalities. In *Mathematical results in quantum mechanics*, pp. 41–47. World Sci. Publ., Hackensack, NJ, 2008.

[38] Esteban, M. J. and Séré, É.: Stationary states of the nonlinear Dirac equation: a variational approach. *Comm. Math. Phys.* **171** (1995), no. 2, 323–350.

[39] Fanelli, L. and Krejčiřík, D.: Location of eigenvalues of three-dimensional non-self-adjoint Dirac operators. *Lett. Math. Phys.* **109** (2019), no. 7, 1473–1485.

[40] Frank, R. L. and Simon, B.: Critical Lieb–Thirring bounds in gaps and the generalized Nevai conjecture for finite gap Jacobi matrices. *Duke Math. J.* **157** (2011), no. 3, 461–493.

[41] Friedrich, T.: *Dirac operators in Riemannian geometry*. Graduate Studies in Mathematics 25, American Mathematical Society, Providence, RI, 2000.

[42] Gradshteyn, I. S. and Ryzhik, I. M.: *Table of integrals, series, and products*. Eighth edition. Elsevier/Academic Press, Amsterdam, 2015.

[43] Kato, T.: Fundamental properties of Hamiltonian operators of Schrödinger type. *Trans. Amer. Math. Soc.* **70** (1951), no. 2, 195–211.

[44] Kato, T.: Holomorphic families of Dirac operators. *Math. Z.* **183** (1983), no. 3, 399–406.

[45] Keller, J. B.: Lower bounds and isoperimetric inequalities for eigenvalues of the Schrödinger equation. *J. Mathematical Phys.* **2** (1961), 262–266.

[46] Klaus, M.: Dirac operators with several Coulomb singularities. *Helv. Phys. Acta* **53** (1980), no. 3, 463–482.

[47] Klaus, M. and Wüst, R.: Spectral properties of Dirac operators with singular potentials. *J. Math. Anal. Appl.* **72** (1979), no. 1, 206–214.

[48] Konno, R. and Kuroda, S. T.: On the finiteness of perturbed eigenvalues. *J. Fac. Sci. Univ. Tokyo Sect. I* **13** (1966), 55–63.

[49] Kwong, M. K.: Uniqueness of positive solutions of $\Delta u - u + u^p = 0$ in $\mathbb{R}^n$. *Arch. Rational Mech. Anal.* **105** (1989), no. 3, 243–266.

[50] Lieb, E. H., Loss, M. and Siedentop, H.: Stability of relativistic matter via Thomas–Fermi theory. *Helv. Phys. Acta* **69** (1996), no. 5-6, 974–984.

[51] Lieb, E. H. and Seiringer, R.: *The stability of matter in quantum mechanics*. Cambridge University Press, Cambridge, 2010.

[52] Lieb, E. and Thirring, W.: Bound on kinetic energy of fermions which proves stability of matter. *Phys. Rev. Lett.* **35** (1975), 687–689.

[53] Lieb, E. H. and Thirring, W. E.: Inequalities for the moments of the eigenvalues of the Schrödinger Hamiltonian and their relation to Sobolev inequalities. In *Studies in mathematical physics. Essays in honor of Valentine Bargmann*, pp. 269–303. Princeton University Press, Princeton, NJ, 1976.

[54] Lions, P.-L.: The concentration-compactness principle in the calculus of variations. The locally compact case, part 1. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **1** (1984), no. 2, 109–145.

[55] McLeod, K.: Uniqueness of positive radial solutions of $\Delta u + f(u) = 0$ in $\mathbb{R}^n$. II. *Trans. Amer. Math. Soc.* **339** (1993), no. 2, 495–505.

[56] Merle, F.: Existence of stationary states for nonlinear Dirac equations. *J. Differential Equations* **74** (1988), no. 1, 50–68.

[57] Nenciu, G.: Self-adjointness and invariance of the essential spectrum for Dirac operators defined as quadratic forms. *Comm. Math. Phys.* **48** (1976), no. 3, 235–247.

[58] Nenciu, G.: Distinguished self-adjoint extension for Dirac operator with potential dominated by multicenter Coulomb potentials. *Helv. Phys. Acta* **50** (1977), no. 1, 1–3.

[59] Pelinovsky, D.: Survey on global existence in the nonlinear Dirac equations in one spatial dimension. In *Harmonic analysis and nonlinear partial differential equations*, pp. 37–50. RIMS Kôkyûroku Bessatsu, B26, Res. Inst. Math. Sci. (RIMS), Kyoto, 2011.

[60] Reed, M. and Simon, B.: *Methods of modern mathematical physics. II. Fourier analysis, self-adjointness*. Academic Press [Harcourt Brace Jovanovich, Publs.], New York-London, 1975.

[61] Reed, M. and Simon, B.: *Methods of modern mathematical physics. IV. Analysis of operators*. Academic Press [Harcourt Brace Jovanovich, Publs.], New York-London, 1978.

[62] Rellich, F.: Störungstheorie der Spektralzerlegung. *Math. Ann.* **113** (1937), no. 1, 600–619.

[63] Schimmer, L., Solovej, J. P. and Tokus, S.: Friedrichs extension and min-max principle for operators with a gap. *Ann. Henri Poincaré* **21** (2020), no. 2, 327–357.

[64] Schmincke, U.-W.: Distinguished selfadjoint extensions of Dirac operators. *Math. Z.* **129** (1972), 335–349.

[65] Šeba, P.: Klein's paradox and the relativistic point interaction. *Lett. Math. Phys.* **18** (1989), no. 1, 77–86.

[66] Simon, B.: *Trace ideals and their applications*. Second edition. Mathematical Surveys and Monographs 120, American Mathematical Society, Providence, RI, 2005.

[67] Wakano, M.: Intensely localized solutions of the classical Dirac–Maxwell field equations. *Prog. Theor. Phys.* **35** (1966), no. 6, 1117–1141.

[68] Thaller, B.: *The Dirac equation*. Texts and Monographs in Physics, Springer, Berlin, 1992.

[69] Wüst, R.: A convergence theorem for selfadjoint operators applicable to Dirac operators with cutoff potentials. *Math. Z.* **131** (1973), 339–349.

[70] Wüst, R.: Distinguished self-adjoint extensions of Dirac operators constructed by means of cut-off potentials. *Math. Z.* **141** (1975), 93–98.

[71] Wüst, R.: Dirac operations with strongly singular potentials. Distinguished self-adjoint extensions constructed with a spectral gap theorem and cut-off potentials. *Math. Z.* **152** (1977), no. 3, 259–271.

Received October 6, 2022; revised July 27, 2023. Published online September 9, 2023.

Jean Dolbeault
CEREMADE (CNRS UMR no. 7534), PSL University, Université Paris-Dauphine
Place de Lattre de Tassigny, 75775 Paris 16, France;
dolbeaul@ceremade.dauphine.fr

David Gontier
CEREMADE (CNRS UMR no. 7534), PSL University, Université Paris-Dauphine,
Place de Lattre de Tassigny, 75775 Paris 16;
and ENS/PSL University, DMA, 75005 Paris, France;
gontier@ceremade.dauphine.fr

Fabio Pizzichillo
Departamento de Matemática e Informática aplicadas a las Ingenierías Civil y Naval,
Escuela Técnica Superior de Ingenieros de Caminos, Canales y Puertos,
Universidad Politécnica de Madrid, Calle Profesor Aranguren 3, 28040 Madrid, Spain;
fabio.pizzichillo@upm.es

Hanne Van Den Bosch
Departamento de Ingeniería Matemática and Center for Mathematical Modeling,
Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile, and CNRS IRL 2807,
Beauchef 851, Piso 5, Santiago, Chile;
hvdbosch@dim.uchile.cl