

Self-adjoint realisations of the Dirac-Coulomb Hamiltonian for heavy nuclei

Matteo Gallone¹ · Alessandro Michelangeli¹

Abstract We derive a classification of the self-adjoint extensions of the three-dimensional Dirac-Coulomb operator in the critical regime of the Coulomb coupling. Our approach is solely based upon the Kreĭn-Višik-Birman extension scheme, or also on Grubb's universal classification theory, as opposite to previous works within the standard von Neumann framework. This let the boundary condition of self-adjointness emerge, neatly and intrinsically, as a multiplicative constraint between regular and singular part of the functions in the domain of the extension, the multiplicative constant giving also immediate information on the invertibility property and on the resolvent and spectral gap of the extension.

Keywords Dirac-Coulomb operator · Self-adjoint extensions · Kreĭn-Višik-Birman extension theory · Grubb's universal classification

Mathematics Subject Classification 47B25 · 47N20 · 47N50 · 81Q10

This work was partially supported by the 2014-2017 MIUR-FIR grant “*Cond-Math: Condensed Matter and Mathematical Physics*” code RBFR13WAET.

✉ Alessandro Michelangeli
alemiche@sissa.it

Matteo Gallone
mgallone@sissa.it

¹ International School for Advanced Studies – SISSA, via Bonomea 265, 34136 Trieste, Italy

1 Introduction

In quantum mechanics a relativistic electron or positron (or more generally a relativistic spin- $\frac{1}{2}$ particle) which moves freely in the three-dimensional space is described by elements of the Hilbert space

$$\mathcal{H} := L^2(\mathbb{R}^3) \otimes \mathbb{C}^4 \cong L^2(\mathbb{R}^3, \mathbb{C}^4, dx) \quad (1.1)$$

and by the (formal) Hamiltonian

$$H_0 := -ic\hbar \boldsymbol{\alpha} \cdot \nabla + \beta mc^2 \quad (1.2)$$

acting on \mathcal{H} , where \hbar is Planck's constant, c is the speed of light, m is the mass of the particle, and $\boldsymbol{\alpha} \equiv (\alpha_1, \alpha_2, \alpha_3)$ and β are the 4×4 matrices

$$\beta = \begin{pmatrix} \mathbb{1} & \mathbb{O} \\ \mathbb{O} & -\mathbb{1} \end{pmatrix}, \quad \alpha_j = \begin{pmatrix} \mathbb{O} & \sigma_j \\ \sigma_j & \mathbb{O} \end{pmatrix}, \quad j \in \{1, 2, 3\}, \quad (1.3)$$

having denoted by $\mathbb{1}$ and \mathbb{O} , respectively, the identity and the zero 2×2 matrix, and by σ_j , as customary, the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (1.4)$$

Explicitly, the scalar product between any two elements $\psi \equiv (\psi_1, \psi_2, \psi_3, \psi_4)$ and $\phi \equiv (\phi_1, \phi_2, \phi_3, \phi_4)$ in \mathcal{H} is given by

$$\langle \psi, \phi \rangle_{\mathcal{H}} = \sum_{j=1}^4 \int_{\mathbb{R}^3} \overline{\psi_j(x)} \phi_j(x) dx, \quad (1.5)$$

and H_0 is the first order matrix-valued differential operator

$$H_0 = \begin{pmatrix} mc^2 \mathbb{1} & -i\hbar c \boldsymbol{\sigma} \cdot \nabla \\ -i\hbar c \boldsymbol{\sigma} \cdot \nabla & -mc^2 \mathbb{1} \end{pmatrix} \quad (1.6)$$

(where $\boldsymbol{\sigma} \equiv (\sigma_1, \sigma_2, \sigma_3)$), known as the free Dirac operator.

The properties of H_0 are well known [34]. H_0 is essentially self-adjoint on $C_0^\infty(\mathbb{R}^3 \setminus \{0\}) \otimes \mathbb{C}^4$ with domain of self-adjointness

$$H^1(\mathbb{R}^3) \otimes \mathbb{C}^4 \cong H^1(\mathbb{R}^3, \mathbb{C}^4), \quad (1.7)$$

and its spectrum (as a self-adjoint operator on \mathcal{H}) is purely absolutely continuous and given by

$$\sigma(H_0) = \sigma_{\text{ac}}(H_0) = (-\infty, -mc^2] \cup [mc^2, +\infty). \quad (1.8)$$

In fact, H_0 is unitarily equivalent to

$$\widetilde{H}_0 := \begin{pmatrix} \mathbb{1}\sqrt{-c^2\Delta + m^2c^4} & 0 \\ 0 & -\mathbb{1}\sqrt{-c^2\Delta + m^2c^4} \end{pmatrix}. \quad (1.9)$$

When the particle is subject to the external scalar field due to the Coulomb interaction with a nucleus of atomic number Z placed in the origin of \mathbb{R}^3 , this is accounted for by the so-called Dirac-Coulomb Hamiltonian

$$H := -ic\hbar \boldsymbol{\alpha} \cdot \nabla + \beta mc^2 - \frac{e^2 Z}{\hbar |x|} \mathbb{1} = H_0 - \frac{cZ\alpha_f}{|x|} \mathbb{1}, \quad (1.10)$$

where now $\mathbb{1}$ is the 4×4 identity matrix (no confusion should arise here and henceforth on the symbol $\mathbb{1}$, being its meaning of identity self-explanatory from the context), e is the elementary charge, and

$$\alpha_f = \frac{e^2}{\hbar c} \approx \frac{1}{137} \quad (1.11)$$

is the fine-structure constant. The operator H can at least be defined minimally on $C_0^\infty(\mathbb{R}^3 \setminus \{0\}) \otimes \mathbb{C}^4$, in which case it is densely defined and symmetric on \mathcal{H} . However, the possibility that this yields an unambiguous physical realisation of H depends on the magnitude of the coupling $Z\alpha_f$, hence of the nuclear charge Z . It is indeed well known [34] that the formal operator (1.10) is essentially self-adjoint on $C_0^\infty(\mathbb{R}^3 \setminus \{0\}, \mathbb{C}^4)$ *only* when $Z\alpha_f \leq \frac{\sqrt{3}}{2}$ (i.e., $Z \leq 118$), in which case the domain of self-adjointness is $\mathcal{D}(H) = \mathcal{D}(H_0) = H^1(\mathbb{R}^3, \mathbb{C}^4)$ and the spectrum consists of the same essential part $\sigma_{\text{ess}}(H) = (-\infty, -mc^2] \cup [mc^2, +\infty)$ as for H_0 , plus a discrete spectrum in the ‘gap’ $(-mc^2, mc^2)$ consisting of eigenvalues $E_{n,\kappa}$ given by Sommerfeld’s celebrated fine-structure formula

$$E_{n,\kappa} = mc^2 \left(1 + \frac{(Z\alpha_f)^2}{(n + \sqrt{\kappa^2 - (Z\alpha_f)^2})^2} \right)^{-\frac{1}{2}}, \quad n \in \mathbb{N}_0, \kappa \in \mathbb{Z} \setminus \{0\}. \quad (1.12)$$

Although the above regime of Z covers all currently known elements (the last one to be discovered, the Oganesson ${}_{118}^{294}\text{Og}$, thus $Z = 118$, was first synthesized in 2002 and formally named in 2016), the problem of the self-adjoint realisation of the Dirac-Coulomb Hamiltonian above the threshold $Z\alpha_f = \frac{\sqrt{3}}{2}$ has been topical since long and so is still today. Even the consideration that the problem only arises due to the idealisation of point-like nuclei (and also because one neglects the anomalous magnetic moment of the electron) does not diminish its relevance, given the extreme experimental precision, for example, of Sommerfeld’s fine-structure formula for the eigenvalues of H when $Z \leq 118$.

From the mathematical side, the study of the self-adjoint extensions of the Dirac-Coulomb Hamiltonian has a long and active history [3–7, 9–13, 17, 20–27, 29, 31–35, 37, 39–41]. A concise survey of this vast literature is discussed in [15]. Let us cast in Theorem 1.1 here below the main relevant facts known today. For the clarity of presentation, let us adopt natural units $c = \hbar = m = e = 1$ henceforth, so as to get

rid of mathematically inessential parameters, the coupling constant of relevance thus becoming $\nu \equiv -Z\alpha_f$.

Theorem 1.1 (Self-adjoint extensions of the minimal Dirac-Coulomb) *On the Hilbert space $\mathcal{H} = L^2(\mathbb{R}^3, \mathbb{C}^4, dx)$ consider, for fixed $\nu \in \mathbb{R}$, the operator*

$$\begin{aligned} H &= H_0 + \frac{\nu}{|x|} \mathbb{1}, & H_0 &= -i\boldsymbol{\alpha} \cdot \nabla + \beta, \\ \mathcal{D}(H) &= \mathcal{D}(H_0) = C_0^\infty(\mathbb{R}^3 \setminus \{0\}, \mathbb{C}^4). \end{aligned} \quad (1.13)$$

H_0 is essentially self-adjoint and the domain (of self-adjointness) of its operator closure $\overline{H_0}$ is $H^1(\mathbb{R}^3, \mathbb{C}^4)$. Moreover, the following holds.

- (i) (Sub-critical regime.) *If $|\nu| \leq \frac{\sqrt{3}}{2}$, then H is essentially self-adjoint and $\mathcal{D}(\overline{H}) = H^1(\mathbb{R}^3, \mathbb{C}^4)$.*
- (ii) (Critical regime.) *If $\frac{\sqrt{3}}{2} < |\nu| < 1$, then H admits an infinity of self-adjoint extensions, among which there is a ‘distinguished’ one, H_D , uniquely characterised by the properties*

$$\mathcal{D}(H_D) \subset \mathcal{D}(|H_0|^{1/2}) \quad \text{or} \quad \mathcal{D}(H_D) \subset \mathcal{D}(|x|^{-1/2}), \quad (1.14)$$

that is, the unique extension whose operator domain is both in the kinetic energy form domain $\mathcal{D}[H_0] = \mathcal{D}(|H_0|^{1/2})$ and in the potential energy form domain $\mathcal{D}[|x|^{-1}] = \mathcal{D}(|x|^{-1/2})$. Moreover, $0 \notin \sigma(H_D)$.

- (iii) (Super-critical regime.) *If $|\nu| \geq 1$, then H admits an infinity of self-adjoint extensions, without a distinguished one in the sense of the operator H_D in the critical regime. In fact, when $|\nu| > 1$ every self-adjoint extension of H has infinitely many eigenfunctions not belonging to $\mathcal{D}(|x|^{-1/2})$.*

In either regime, the spectrum of any self-adjoint extension \tilde{H} of H is such that

$$\begin{aligned} \sigma_{\text{ess}}(\tilde{H}) &= \sigma(\overline{H_0}) = (-\infty, -1] \cup [1, +\infty) \\ \sigma_{\text{disc}}(\tilde{H}) &\subset (-1, 1). \end{aligned} \quad (1.15)$$

It is worth remarking that for Coulomb-like matrix-valued interactions $V(x)$ that are *not* of the form $\nu|x|^{-1}\mathbb{1}$ but still satisfy $|V(x)| \leq \nu|x|^{-1}$, the sub-critical regime described in Theorem 1.1(i) only ranges up to $|\nu| < \frac{1}{2}$, and counterexamples are well known of operators $H_0 + V$ with $|V(x)| \leq (\frac{1}{2} + \varepsilon)|x|^{-1}$ for arbitrary $\varepsilon > 0$ and failing to be essentially self-adjoint on $C_0^\infty(\mathbb{R}^3 \setminus \{0\}, \mathbb{C}^4)$ [3].

In this work we are primarily focused on the critical regime, $|\nu| \in (\frac{\sqrt{3}}{2}, 1)$. This is a regime of ultra-heavy nuclei, in fact nuclei of elements that one expects to discover in the next future. It is the first regime where the Kato-Rellich-like perturbative arguments, applicable for small ν 's, cease to work. It is also regarded as a physically meaningful regime, because as long as $|\nu| < 1$ Sommerfeld's fine-structure formula still provides, formally, bound states for real energy levels, which only become

complex when $|\nu| > 1$, thus predicting an instability of the atom (the ‘ $Z = 137$ catastrophe’).

In fact, the critical regime for the Dirac-Coulomb operator is already intensively studied, with a special focus on the ‘distinguished’ self-adjoint extension H_D [3, 5, 6, 11–13, 21, 23–25, 29, 32, 35, 37, 39–41]. Conceptually, and qualitatively, this has very much in common with the analogous, scaling-critical problem of the self-adjoint realisation of the (formal) non-relativistic and pseudo-relativistic Schrödinger operators

$$-\Delta + \frac{\nu}{|x|^2} \quad \text{or} \quad \sqrt{1 - \Delta} + \frac{\nu}{|x|} \quad \text{on } L^2(\mathbb{R}^3)$$

when ν is out of the perturbative regime, an issue that is both standard textbook material [30, Appendix to X.1] and object of recent research [8, 14, 28].

Our perspective in the present work is that of a convenient *classification* of all self-adjoint extensions of the minimally defined H , both in terms of explicit *boundary conditions* for the functions in the domain of each extension, and in terms of an intrinsic, canonical structure of the domain of each extension. Moreover, unlike recent classifications [21, 35] based on von Neumann’s extension theory, we put the emphasis on the straightforward applicability of the so-called Kreĭn-Višik-Birman extension theory [16], and in fact of its non-semi-bounded version, namely Grubb’s universal classification theory [19, Chapter 13], which as a matter of fact turns out to be particularly versatile and informative in this context.

In order to give a first formulation of our main result, let us exploit, as customary, the canonical decomposition of H into partial wave operators [34, Section 4.6], which is induced by its spherical symmetry. By expressing $x \equiv (x_1, x_2, x_3) \in \mathbb{R}^3$ in polar coordinates $x = (r, \Omega) \in \mathbb{R}^+ \times \mathbb{S}^2$, $r := |x|$, the map $\psi(x) \mapsto r\psi(x_1(r, \Omega), x_2(r, \Omega), x_3(r, \Omega))$ induces a unitary isomorphism

$$L^2(\mathbb{R}^3, \mathbb{C}^4, dx) \xrightarrow{\cong} L^2(\mathbb{R}^+, dr) \otimes L^2(\mathbb{S}^2, \mathbb{C}^4, d\Omega).$$

In terms of the observables

$$\begin{aligned} \mathbf{L} &= \mathbf{x} \times (-i\nabla), & \mathbf{S} &= -\frac{1}{4}\boldsymbol{\alpha} \times \boldsymbol{\alpha}, \\ \mathbf{J} &= \mathbf{L} + \mathbf{S} \equiv (J_1, J_2, J_3), & K &= \beta(2\mathbf{L} \cdot \mathbf{S} + \mathbb{1}), \end{aligned}$$

one further decomposes

$$L^2(\mathbb{S}^2, \mathbb{C}^4, d\Omega) \cong \bigoplus_{j \in \frac{1}{2}\mathbb{N}} \bigoplus_{m_j = -j}^j \bigoplus_{\kappa_j = \pm(j + \frac{1}{2})} \mathcal{K}_{m_j, \kappa_j}, \quad (1.16)$$

where

$$\mathcal{K}_{m_j, \kappa_j} := \text{span}\{\Psi_{m_j, \kappa_j}^+, \Psi_{m_j, \kappa_j}^-\} \cong \mathbb{C}^2 \quad (1.17)$$

and Ψ_{m_j, κ_j}^+ and Ψ_{m_j, κ_j}^- are two orthonormal vectors in \mathbb{C}^4 , and simultaneous eigenvectors of the observables $J^2 \upharpoonright L^2(\mathbb{S}^2, \mathbb{C}^4, d\Omega)$, $J_3 \upharpoonright L^2(\mathbb{S}^2, \mathbb{C}^4, d\Omega)$, and $K \upharpoonright L^2(\mathbb{S}^2, \mathbb{C}^4, d\Omega)$ with eigenvalue, respectively, $j(j+1)$, m_j , and κ_j . It then turns out that each subspace

$$\mathcal{H}_{m_j, \kappa_j} := L^2(\mathbb{R}^+, dr) \otimes \mathcal{K}_{m_j, \kappa_j} \cong L^2(\mathbb{R}^+, \mathbb{C}^2, dr) \quad (1.18)$$

is a reducing subspace for the Dirac-Coulomb Hamiltonian H , which, through the overall isomorphism

$$U : L^2(\mathbb{R}^3, \mathbb{C}^4, dx) \xrightarrow{\cong} \bigoplus_{j \in \frac{1}{2}\mathbb{N}} \bigoplus_{m_j = -j}^j \bigoplus_{\kappa_j = \pm(j + \frac{1}{2})} \mathcal{H}_{m_j, \kappa_j}, \quad (1.19)$$

is therefore unitarily equivalent to

$$UHU^* = \bigoplus_{j \in \frac{1}{2}\mathbb{N}} \bigoplus_{m_j = -j}^j \bigoplus_{\kappa_j = \pm(j + \frac{1}{2})} h_{m_j, \kappa_j}, \quad (1.20)$$

where

$$h_{m_j, \kappa_j} := \begin{pmatrix} 1 + \frac{v}{r} & -\frac{d}{dr} + \frac{\kappa_j}{r} \\ \frac{d}{dr} + \frac{\kappa_j}{r} & -1 + \frac{v}{r} \end{pmatrix}, \quad (1.21)$$

$$\mathcal{D}(h_{m_j, \kappa_j}) := C_0^\infty(\mathbb{R}^+) \otimes \mathcal{K}_{m_j, \kappa_j} \cong C_0^\infty(\mathbb{R}^+, \mathbb{C}^2).$$

Thus, (1.21) defines a densely defined and symmetric operator on the Hilbert space (1.18) and the overall problem of the self-adjoint realisation of H is reduced to the same problem in each reducing subspace.

In particular, it is of physical relevance to consider each operator

$$h_{m_j} := h_{m_j, \kappa_j = j + \frac{1}{2}} \oplus h_{m_j, \kappa_j = -(j + \frac{1}{2})} \quad (1.22)$$

acting block-diagonal-wise, with the two different spin-orbit components, on the Hilbert eigenspace $L^2(\mathbb{R}^+, \mathbb{C}^4, dr)$ of (j, m_j) -eigenvalue for J^2 and J_3 .

Now, the following property is well known, as one can see by means of standard limit-point limit-circle arguments [38, Chapter 6.B]. Its proof is discussed, e.g., in [15, Section 2].

Proposition 1.2 *The operator h_{m_j, κ_j} is essentially self-adjoint on its domain with respect to the Hilbert space $\mathcal{H}_{m_j, \kappa_j}$ if and only if*

$$v^2 \leq \kappa_j^2 - \frac{1}{4}, \quad (1.23)$$

and it has deficiency indices $(1, 1)$ otherwise. In particular, in the regime $|v| \in (\frac{\sqrt{3}}{2}, 1)$ only the operators of the decomposition (1.20) with $\kappa_j^2 = 1$, thus

$$h_{\frac{1}{2},1}, \quad h_{-\frac{1}{2},1}, \quad h_{\frac{1}{2},-1}, \quad h_{-\frac{1}{2},-1}, \quad (1.24)$$

have deficiency indices $(1, 1)$, all others being essentially self-adjoint.

Therefore the operator $h_{\frac{1}{2},1} \oplus h_{\frac{1}{2},-1} \oplus h_{-\frac{1}{2},1} \oplus h_{-\frac{1}{2},-1}$, and hence H itself, has deficiency indices $(4, 4)$. This means that there is a 16-real-parameter family of self-adjoint extensions of H , hence of physically inequivalent realisations of the Dirac-Coulomb Hamiltonian. From the operator-theoretic point of view, the analysis of the self-adjoint extensions of $h_{\frac{1}{2},1}$ is the very same as for the other three operators (and in fact $h_{\frac{1}{2},1}$ and $h_{-\frac{1}{2},1}$ have the same formal action on $L^2(\mathbb{R}^+, \mathbb{C}^2)$, and so have $h_{\frac{1}{2},-1}$ and $h_{-\frac{1}{2},-1}$), and hence we will discuss only the first case.

There is room for extensions only on the sector $j = \frac{1}{2}$ of lowest total angular momentum J^2 and, as we shall discuss in Sect. 2, each extension corresponds to a particular prescription on the wave functions of the domain in the vicinity of the centre $x = 0$ of the Coulomb interaction. For higher j 's the large angular momentum makes the Coulomb singularity lesser and lesser relevant, and on such sectors H is already essentially self-adjoint.

Physically, the relevant class of extensions is rather the *one*-parameter sub-family consisting of the same extension for each elementary operators (1.24), in a sense that will be evident in the next Section, that is, extensions where the same boundary conditions of self-adjointness occurs on each block of H – it would be non-physical to have a different behaviour of the physical Hamiltonian on different sectors $\mathcal{H}_{m_j, \kappa_j}$ of its symmetry.

We can now conclude this Introduction by anticipating an informal version of the main results that we will present rigorously in Sect. 2. First and foremost, we do *not* apply the self-adjoint extension theory of von Neumann, unlike what is done ubiquitously in the previous literature, and we exploit instead (and to our knowledge for the first time) the Kreĭn-Višik-Birman/Grubb extension scheme. To our taste such a scheme produces in this context the most informative version of the classification of the self-adjoint extensions of H in a considerably less laborious way.

Informally, our results can be summarised as follows.

Theorem 1.3 (Classification of Dirac-Coulomb extensions—informal version)

Let $|v| \in (\frac{\sqrt{3}}{2}, 1)$.

- (i) On each of the four sectors $(j, m_j, \kappa_j) = (\frac{1}{2}, \pm\frac{1}{2}, \pm 1)$ of non-self-adjointness, the operator H admits a one-parameter family $(S_\beta)_{\beta \in \mathbb{R} \cup \{\infty\}}$ of self-adjoint extensions—which are then restrictions of H^* .
- (ii) Whereas the domain of H^* in each sector consists of spinors g with H^1 -regularity on $[\varepsilon, +\infty)$ for all $\varepsilon > 0$ and the short-distance asymptotics

$$g(r) = g_0 r^{-\sqrt{1-v^2}} + g_1 r^{\sqrt{1-v^2}} + o(r^{1/2}) \quad \text{as } r \downarrow 0 \quad (1.25)$$

for some $g_0, g_1 \in \mathbb{C}^2$ dependent on v only, the domain $\mathcal{D}(S_\beta)$ of the extension S_β consists of those such spinors for which a prescribed ratio holds between the corresponding components of g_1 and g_0 , for concreteness

$$\frac{g_1^+}{g_0^+} = c_\nu \beta + d_\nu \quad (1.26)$$

for some explicitly known constants $c_\nu, d_\nu \in \mathbb{C}$.

- (iii) The extension $\beta = \infty$ is the restriction S_D , on the considered sector (j, m_j, κ_j) , of the distinguished extension H_D of H discussed in Theorem 1.1(ii): the functions in its domain have the asymptotics (1.25) with $g_0 \equiv 0$, i.e., without singular term.
- (iv) All those extensions S_β with $\beta \neq 0$ are invertible with everywhere defined and bounded inverse, in which case the inverse S_β^{-1} is an explicit rank-one perturbation of S_D^{-1} .
- (v) The gap in the spectrum of S_β around $\lambda = 0$ has a direct estimate in terms of β and $\|S_D^{-1}\|$ and must be at least the interval

$$(-E_0(\beta), E_0(\beta)), \quad E_0(\beta) := \frac{|\beta|}{|\beta| \|S_D^{-1}\| + 1}. \quad (1.27)$$

Last, here is how the material is organised. As mentioned already, in Sect. 2 we state rigorously our main results, whose proof, outlined in Sect. 2 itself, is based on intermediate results that we prove in Sects. 3, 4, and 5. In Sect. 6 we discuss further properties of the Dirac-Coulomb extensions involving the resolvent and the spectral gap at zero.

Notation Essentially all the notation adopted here is standard, let us only emphasize the following. Concerning the various sums of spaces that will occur, we denote by $\dot{+}$ the direct sum of vector spaces, by \oplus the direct orthogonal sum of *closed* Hilbert subspaces of the same underlying Hilbert space, and by \boxplus the direct sum of subspaces of \mathcal{H} that are orthogonal to each other but are not a priori all closed. *Operator* domain and *form* domain of any given densely defined and symmetric operator S are denoted, respectively, by $\mathcal{D}(S)$ and $\mathcal{D}[S]$. As customary, $\mathbb{R}^+ = (0, +\infty)$, and $\sigma(T)$ and $\rho(T)$ denote, respectively, the spectrum and the resolvent set of an operator T on Hilbert space. The notation $A \lesssim B$ stands for the inequality $A \leq cB$ for some constant that is universal or is clear from the context not to depend on the other variables of the inequality itself. We refer to elements $g \in \mathbb{C}^2$ or in $L^2(\mathbb{R}^+, \mathbb{C}^2)$ as spinors $g = \begin{pmatrix} g^+ \\ g^- \end{pmatrix}$.

G^T denotes the transpose of a matrix G .

2 Classification scheme and main results

The original Kreĭn-Višik-Birman scheme [2, 16] for the determination and classification of the self-adjoint extensions of a given densely defined and symmetric operator on Hilbert space was developed for *semi-bounded* operators: for this case one can non-restrictively assume that the bottom of the operator S to extend is strictly positive and hence a canonical extension exists, the Friedrichs extension S_F , with the same bottom and hence with everywhere defined bounded inverse S_F^{-1} .

In fact, to a large extent, the role of S_F in the theory can be played as well by any other ‘distinguished’ self-adjoint extension S_D of S which is itself invertible with everywhere defined and bounded inverse S_D^{-1} , and this makes many results of the theory applicable also to a (densely defined and symmetric) non-semi-bounded S , provided that S admits such an extension S_D . In this spirit, Grubb’s ‘universal classification’ scheme was later developed [18] (a modern survey of which may be found in [19, Chapter 13]), which only makes reference to the existence of an invertible extension and, in the case of symmetric operators, it reproduces many features of the Kreĭn-Višik-Birman scheme.

For our next purposes, let us single out from such extension theories the following result, for a discussion of which we refer to [16, Theorem 3.4].

Theorem 2.1 (Classification of self-adjoint extensions—operator version) *Let S be a densely defined symmetric operator on a Hilbert space \mathcal{H} and assume that S admits a self-adjoint extension S_D which is invertible with everywhere defined inverse. Then there is a one-to-one correspondence between the family of all self-adjoint extensions of S on \mathcal{H} and the family of the self-adjoint operators on Hilbert subspaces of $\ker S^*$. If T is any such operator, in the correspondence $T \leftrightarrow S_T$ each self-adjoint extension S_T of S is given by*

$$S_T = S^* \upharpoonright \mathcal{D}(S_T)$$

$$\mathcal{D}(S_T) = \left\{ f + S_D^{-1}(Tv + w) + v \left| \begin{array}{l} f \in \mathcal{D}(\bar{S}), \quad v \in \mathcal{D}(T) \\ w \in \ker S^* \cap \mathcal{D}(T)^\perp \end{array} \right. \right\}. \quad (2.1)$$

When applying Theorem 2.1 to the extension problem for the operator h_{m_j, κ_j} defined in (1.21) on the Hilbert space $\mathcal{H}_{m_j, \kappa_j}$ defined in (1.18), it is natural that the reference extension is the distinguished extension of Theorem 1.1(ii).

Acting on the Hilbert space $L^2(\mathbb{R}^+, \mathbb{C}^2, dr)$ with scalar product

$$\langle \psi, \phi \rangle_{L^2(\mathbb{R}^+, \mathbb{C}^2)} = \int_0^{+\infty} \langle \psi(r), \phi(r) \rangle_{\mathbb{C}^2} dr = \sum_{\alpha=\pm} \int_0^{+\infty} \overline{\psi^\alpha(r)} \phi^\alpha(r) dr \quad (2.2)$$

$$\psi \equiv \begin{pmatrix} \psi^+ \\ \psi^- \end{pmatrix}, \quad \phi \equiv \begin{pmatrix} \phi^+ \\ \phi^- \end{pmatrix} \in L^2(\mathbb{R}^+, \mathbb{C}^2),$$

we consider the operator

$$S := \begin{pmatrix} 1 + \frac{v}{r} & -\frac{d}{dr} + \frac{1}{r} \\ \frac{d}{dr} + \frac{1}{r} & -1 + \frac{v}{r} \end{pmatrix}, \quad (2.3)$$

$$\mathcal{D}(S) := C_0^\infty(\mathbb{R}^+, \mathbb{C}^2).$$

S is non-semi-bounded, densely defined, and symmetric, and following from Proposition 1.2 it has deficiency indices (1, 1). For convenience, let us also denote by \tilde{S} the differential operator defined by

$$\tilde{S} \begin{pmatrix} f^+ \\ f^- \end{pmatrix} := \begin{pmatrix} 1 + \frac{\nu}{r} & -\frac{d}{dr} + \frac{1}{r} \\ \frac{d}{dr} + \frac{1}{r} & -1 + \frac{\nu}{r} \end{pmatrix} \begin{pmatrix} f^+ \\ f^- \end{pmatrix}. \quad (2.4)$$

Since \tilde{S} has real smooth coefficients, and is formally self-adjoint, it is a standard fact [19, Section 4.1] that the operator closure \bar{S} and the adjoint S^* of S are nothing but, respectively, the *minimal* and the *maximal realisation* of \tilde{S} , that is, they both act as \tilde{S} respectively on

$$\begin{aligned} \mathcal{D}(\tilde{S}) &= \overline{C_0^\infty(\mathbb{R}^+, \mathbb{C}^2)}^{\|\cdot\|_S} \\ \mathcal{D}(S^*) &= \{\psi \in L^2(\mathbb{R}^+, \mathbb{C}^2) \mid \tilde{S}\psi \in L^2(\mathbb{R}^+, \mathbb{C}^2)\}, \end{aligned} \quad (2.5)$$

where $\|\cdot\|_S$ is the graph norm associated with S . One has $\bar{S} \subset S^*$, and by *self-adjoint realisation* of S we shall mean any operator $R = R^*$ on $L^2(\mathbb{R}^+, \mathbb{C}^2)$ such that $\bar{S} \subset R \subset S^*$.

In order to identify the self-adjoint realisations of S using the Kreĭn-Višik-Birman scheme of Theorem 2.1, we shall collect the intermediate results of Propositions 2.2, 2.3, and 2.4 below, whose proof is deferred to the following Sections.

For convenience, let us introduce the parameter

$$B := \sqrt{1 - \nu^2}. \quad (2.6)$$

It will be important to remember throughout our analysis that $B \in (0, \frac{1}{2})$.

First one needs a characterisation of $\ker S^*$.

Proposition 2.2 *For every $|\nu| \in (\frac{\sqrt{3}}{2}, 1)$ the operator S^* has a one dimensional kernel, spanned by the function*

$$\Phi = \begin{pmatrix} \Phi^+ \\ \Phi^- \end{pmatrix} \quad (2.7)$$

with

$$\Phi^\pm(r) := e^{-r} r^{-B} \left(\frac{\pm(1+\nu)+B}{1+\nu} U_{-B,1-2B}(2r) - \frac{2rB}{1+\nu} U_{1-B,2-2B}(2r) \right), \quad (2.8)$$

where $U_{a,b}(r)$ is the Tricomi function [1, Sec. 13.1.3]. Φ is analytic on $(0, +\infty)$ with asymptotics

$$\begin{aligned} \Phi(r) &= r^{-B} \frac{\Gamma(2B)}{\Gamma(B)} \left(-\frac{1+\nu+B}{1+\nu} \right) + \begin{pmatrix} q^+ \\ q^- \end{pmatrix} r^B + O(r^{1-B}) \quad \text{as } r \downarrow 0 \\ \Phi(r) &= 2^B \begin{pmatrix} 1 \\ -1 \end{pmatrix} r^{-B} e^{-r} (1 + O(r^{-1})) \quad \text{as } r \rightarrow +\infty, \end{aligned} \quad (2.9)$$

where q^\pm are both non-zero and explicitly given by (3.11) below.

Next, one needs to identify a reference extension S_D of S which be self-adjoint and with everywhere defined inverse, and to characterise the action of S_D on $\ker S^*$.

Proposition 2.3 (i) *There exists a self-adjoint realisation S_D of S with the property that*

$$\mathcal{D}(S_D) \subset H^{1/2}(\mathbb{R}^+, \mathbb{C}^2) \quad \text{or} \quad \mathcal{D}(S_D) \subset \mathcal{D}[r^{-1}], \quad (2.10)$$

where the latter is the form domain of the multiplication operator by r^{-1} on each component of $L^2(\mathbb{R}^+, \mathbb{C}^2)$ (the space of ‘finite potential energy’). S_D is the only self-adjoint realisation of S satisfying (2.10).

- (ii) S_D is invertible on $L^2(\mathbb{R}^+, \mathbb{C}^2)$ with everywhere defined and bounded inverse. The explicit integral kernel of S_D is given by (4.10).
 (iii) In terms of the spaces $\mathcal{D}(\bar{S})$ and $\ker S^*$ one has

$$\mathcal{D}(S_D) = \mathcal{D}(\bar{S}) \dot{+} S_D^{-1} \ker S^*. \quad (2.11)$$

Moreover,

$$\begin{aligned} \mathcal{D}(S^*) &= \mathcal{D}(S_D) \dot{+} \ker S^*, \\ &= \mathcal{D}(\bar{S}) \dot{+} S_D^{-1} \ker S^* \dot{+} \ker S^*. \end{aligned} \quad (2.12)$$

- (iv) For the vector $S_D^{-1}\Phi$, where $\Phi \in \ker S^*$ is given by (2.7)–(2.8), one has the following point-wise asymptotics:

$$S_D^{-1}\Phi(r) \sim \begin{pmatrix} p^+ \\ p^- \end{pmatrix} r^B + o(r^{1/2}) \quad \text{as } r \downarrow 0, \quad (2.13)$$

where p^\pm are both non-zero and explicitly given in (4.17) below.¹

Last, an amount of information is needed on the domain of the operator closure \bar{S} of S . Although $\mathcal{D}(\bar{S})$ is canonically constructed as the closure of $\mathcal{D}(S)$ in the operator norm, it does not correspond to a standard functional space. In fact in Sect. 5 we will present a complete characterisation of $\mathcal{D}(\bar{S})$, from which we will be able to deduce the following properties, relevant for our main results.

Proposition 2.4 *Let $f \in \mathcal{D}(\bar{S})$. Then $f \in H_{\text{loc}}^1(\mathbb{R}^+)$ and*

$$f(r) = o(r^{1/2}) \quad \text{as } r \downarrow 0. \quad (2.14)$$

With Propositions 2.2, 2.3, and 2.4 at hand, we can now formulate a general classification as follows.

Theorem 2.5 (Classification of the self-adjoint realisations for the Dirac-Coulomb Hamiltonian—structural version)

The self-adjoint extensions of the operator S on $L^2(\mathbb{R}^+, \mathbb{C}^2, dr)$ defined in (2.3) constitute a one-parameter family $(S_\beta)_{\beta \in \mathbb{R} \cup \{\infty\}}$ of restrictions of the adjoint operator S^ determined in (2.5), each of which is given by*

¹ In fact, with a slightly more elaborate argument we can better estimate the remainder in (2.13) as a $O(r^{1-B})$ term; however, this is not needed in the analysis that follows.

$$\begin{aligned}
S_\beta &:= S^* \upharpoonright \mathcal{D}(S_\beta) \\
\mathcal{D}(S_\beta) &:= \left\{ g = f + c(\beta S_D^{-1} \Phi + \Phi) \left| \begin{array}{l} f \in \mathcal{D}(\bar{S}) \\ c \in \mathbb{C} \end{array} \right. \right\}.
\end{aligned} \tag{2.15}$$

Here S_D is the distinguished self-adjoint extension of S identified in Proposition 2.3 and Φ is the spanning element of $\ker S^*$ identified in Proposition 2.2. In this parametrisation the distinguished extension S_D corresponds to $\beta = \infty$. For each $g \in \mathcal{D}(S_\beta)$ the function $f \in \mathcal{D}(\bar{S})$ and the constant $c \in \mathbb{C}$ are uniquely determined.

In analogy with the parametrisation formula (2.15), we can re-express the above result in terms of boundary conditions at the centre of the Coulomb singularity.

Theorem 2.6 (Classification of the self-adjoint realisations for the Dirac-Coulomb Hamiltonian – boundary condition version)

- (i) Any function $g = \begin{pmatrix} g^+ \\ g^- \end{pmatrix} \in \mathcal{D}(S^*)$ satisfies the short-distance asymptotics

$$\begin{aligned}
\lim_{r \downarrow 0} r^B g(r) &= g_0 \\
\lim_{r \downarrow 0} r^{-B} (g(r) - g_0 r^{-B}) &= g_1
\end{aligned} \tag{2.16}$$

for some $g_0, g_1 \in \mathbb{C}^2$. In particular,

$$g(r) = g_0 r^{-B} + g_1 r^B + o(r^{1/2}) \quad \text{as } r \downarrow 0. \tag{2.17}$$

- (ii) The self-adjoint extensions of the operator S on $L^2(\mathbb{R}^+, \mathbb{C}^2)$ defined in (2.3) constitute a one-parameter family $(S_\beta)_{\beta \in \mathbb{R} \cup \{\infty\}}$ of restrictions of the adjoint operator S^* , each of which is given by

$$\begin{aligned}
S_\beta &:= S^* \upharpoonright \mathcal{D}(S_\beta) \\
\mathcal{D}(S_\beta) &:= \left\{ g \in \mathcal{D}(S^*) \left| \begin{array}{l} \frac{g_1^+}{g_0^+} = c_\nu \beta + d_\nu \end{array} \right. \right\},
\end{aligned} \tag{2.18}$$

where

$$\begin{aligned}
c_\nu &= p^+ \left(\frac{\Gamma(2B)}{\Gamma(B)} \frac{1+\nu+B}{1+\nu} \right)^{-1} \\
d_\nu &= q^+ \left(\frac{\Gamma(2B)}{\Gamma(B)} \frac{1+\nu+B}{1+\nu} \right)^{-1},
\end{aligned} \tag{2.19}$$

and p^+ and q^+ are given, respectively, by (4.17) and (3.11). This is precisely the same parametrisation of the extension as in Theorem 2.5.

The proofs of Theorems 2.5 and 2.6 are an application of the general Kreĭn-Višik-Birman Theorem 2.1, through the intermediate results of Propositions 2.2, 2.3, and 2.4, as we shall show in a moment. Owing to further corollaries of Theorem 2.1, which we work out in detail in Sect. 6 (Theorem 6.1 therein), we can supplement

the above extension picture with an additional amount of information concerning the invertibility, the resolvent, and the spectral gap of each realisation S_β . This too is an example of relevant and non-trivial features of the self-adjoint extensions that can be established in a relatively cheap and elementary manner, unlike the counterpart way based on von Neumann's extension theory.

Theorem 2.7 (Invertibility, resolvent, and estimate on the spectral gap)

The elements of the family $(S_\beta)_{\beta \in \mathbb{R} \cup \{\infty\}}$ of the self-adjoint extensions of the operator S on $L^2(\mathbb{R}^+, \mathbb{C}^2, dr)$ defined in (2.3), labelled according to the parametrisation of Theorem 2.5, have the following properties.

- (i) S_β is invertible on the whole $L^2(\mathbb{R}^+, \mathbb{C}^2)$ if and only if $\beta \neq 0$.
- (ii) For each invertible extension S_β ,

$$S_\beta^{-1} = S_D^{-1} + \frac{1}{\beta \|\Phi\|^2} |\Phi\rangle\langle\Phi|. \quad (2.20)$$

- (iii) For each invertible extension S_β ,

$$\sigma_{\text{ess}}(S_\beta) = \sigma_{\text{ess}}(S_D) = (-\infty, -1] \cup [1, +\infty), \quad (2.21)$$

and the gap in the spectrum $\sigma(S_\beta)$ around $E = 0$ is at least the interval $(-E(\beta), E(\beta))$, where

$$E(\beta) := \frac{|\beta|}{|\beta| \|S_D^{-1}\| + 1}. \quad (2.22)$$

We conclude this Section with the proof of Theorems 2.5 and 2.6, and we defer the proof of the technical intermediate results and of Theorem 2.7 to the following Sections.

Proof of Theorem 2.5 One extension is surely the distinguished extension S_D , with domain $\mathcal{D}(S_D) = \mathcal{D}(\bar{S}) \dot{+} S_D^{-1} \ker S^*$ (Proposition 2.3(iii)), which is of the form (2.15) for $\beta = \infty$: indeed, with respect to the general formula (2.1), this is the extension that corresponds to an operator T defined on $\{0\} \subset \ker S^*$. Since $\dim \ker S^* = 1$ (Proposition 2.2), for all other extensions of S the parametrising operator T , in the sense of the general formula (2.1), must be self-adjoint on the whole one-dimensional span $\{\Phi\}$, and therefore is the multiplication operator by a scalar β . Then (2.1) takes the form (2.15). The uniqueness of the decomposition of $g \in \mathcal{D}(S_\beta)$ into $g \in \mathcal{D}(S_\beta)$ is a direct consequence of the direct sum decomposition (2.12) of Proposition 2.3(iii). \square

Proof of Theorem 2.6 (i) It was determined in Propositions 2.2, 2.3, and 2.4 that a generic $g \in \mathcal{D}(S^*)$ decomposes with respect to (2.12) as

$$\begin{pmatrix} g^+ \\ g^- \end{pmatrix} = \begin{pmatrix} f^+ \\ f^- \end{pmatrix} + a S_D^{-1} \begin{pmatrix} \Phi^+ \\ \Phi^- \end{pmatrix} + \frac{b}{\gamma} \begin{pmatrix} \Phi^+ \\ \Phi^- \end{pmatrix} \quad \gamma := \frac{\Gamma(2B)}{\Gamma(B)} \frac{1+\nu+B}{1+\nu} \quad (2.23)$$

for some $a, b \in \mathbb{C}$, and moreover, as $r \downarrow 0$,

$$\begin{aligned} f(r) &= o(r^{1/2}), \\ (S_D^{-1}\Phi)(r) &= \begin{pmatrix} p^+ \\ p^- \end{pmatrix} r^B + o(r^{1/2}), \\ r^B \Phi(r) &= \begin{pmatrix} 1 \\ -\frac{1+\nu-B}{1+\nu+B} \end{pmatrix} \gamma + \begin{pmatrix} q^+ \\ q^- \end{pmatrix} r^{2B} + o(r^{1/2+B}) \end{aligned}$$

(see, respectively, (2.14), (2.13), and (2.9) above). Therefore, the limit in the first component yields $r^B g^+(r) \xrightarrow{r \downarrow 0} b$, and also

$$\begin{aligned} r^{-B}(g^+(r) - br^{-B}) &= r^{-B}(f^+(r) + a(S_D^{-1}\Phi)^+(r) + b\gamma^{-1}\Phi^+(r) - br^{-B}) \\ &= a p^+ + b q^+ \gamma^{-1} + o(r^{1/2-B}), \end{aligned}$$

that is, $r^{-B}(g^+(r) - br^{-B}) \xrightarrow{r \downarrow 0} a p^+ + b q^+ \gamma^{-1}$. Thus, (2.16) and (2.17) follow by setting

$$g_0^+ := b, \quad g_1^+ := a p^+ + b q^+ \gamma^{-1} \quad (*)$$

(an analogous argument holds for the lower components).

(ii) Necessary and sufficient condition for $g \in \mathcal{D}(S^*)$ to belong to the domain $\mathcal{D}(S_\beta)$ of the extension S_β determined by (2.15) of Theorem 2.5 is that in the decomposition (2.23) above the coefficients a and b satisfy $a = \beta b \gamma^{-1}$. Owing to (*) and (2.19), the latter condition reads $g_1^+ / g_0^+ = c_\nu \beta + d_\nu$. \square

Last, it is worth highlighting a couple of important remarks.

Remark 2.8 The proof of Theorem 2.6 shows that the decomposition of $g \in \mathcal{D}(S_\beta)$ determined by (2.15), and hence c and f , are explicitly given by

$$\begin{aligned} c &= \left(\frac{\Gamma(B)}{\Gamma(2B)} \frac{1+\nu}{1+\nu+B} \right) \cdot \lim_{r \downarrow 0} r^B g^+(r) \\ f &= g - c(\beta S_D^{-1}\Phi + \Phi). \end{aligned} \quad (2.24)$$

Indeed, in the notation of (2.23) therein, $b = \gamma c$. In fact, the same argument shows that the first equation in (2.24) determines the component $c\Phi \in \ker S^*$ of a generic $g \in \mathcal{D}(S^*)$, and hence defines the (non-orthogonal) projection $\mathcal{D}(S^*) \rightarrow \ker S^*$, $g \mapsto c\Phi$ induced by the decomposition formula (2.12). When $\beta \neq 0$, one has equivalently

$$c = \beta^{-1} \int_0^{+\infty} \langle \Phi(r), (\tilde{S}g)(r) \rangle_{\mathbb{C}^2} dr. \quad (2.25)$$

Indeed $\tilde{S}g = S_\beta g = S^*g = \bar{S}f + c\beta\Phi$ and $\text{ran } \bar{S} \perp \ker S^*$, whence it follows that $\langle \Phi, \tilde{S}g \rangle_{L^2(\mathbb{R}^3, \mathbb{C}^2)} = c\beta$.

Remark 2.9 As typical when the operator which one studies the self-adjoint extensions of is a differential operator, one interprets (2.12) as the canonical decomposition of an element $g \in \mathcal{D}(S^*)$ into a ‘regular’ and a ‘singular’ part

$$\begin{aligned} g_{\text{reg}} &:= f + a S_D^{-1} \Phi \in \mathcal{D}(S_D) \\ g_{\text{sing}} &:= \frac{b}{\gamma} \Phi \in \ker S^*, \end{aligned} \quad (2.26)$$

where $a, b \in \mathbb{C}$ and $f \in \mathcal{D}(\bar{S})$ are determined by g and $\gamma = \frac{\Gamma(2B)}{\Gamma(B)} \frac{1+\nu-B}{1+\nu}$. Indeed $\mathcal{D}(S_D)$ has a higher regularity than $\ker S^*$: functions in the former space vanish at zero, as follows from (2.13)–(2.14), whereas Φ diverges at zero, as seen in (2.9). In this language, $r^{-B} g_{\text{reg}}^+(r) \xrightarrow{r \downarrow 0} ap^+$ and $r^B g_{\text{sing}}^+(r) \xrightarrow{r \downarrow 0} b$, and the self-adjointness condition (2.18) that selects, among the elements in $\mathcal{D}(S^*)$, only those in $\mathcal{D}(S_\beta)$ reads

$$\left(\frac{\gamma}{p^+} \lim_{r \downarrow 0} r^{-B} g_{\text{reg}}^+(r) \right) = (c_\nu \beta + d_\nu) \left(\lim_{r \downarrow 0} r^B g_{\text{sing}}^+(r) \right), \quad (2.27)$$

that is, the ratio between $\gamma(p^+)^{-1}$ times the coefficient of the leading vanishing term of g_{reg}^+ and the coefficient of the leading divergent term of g_{sing}^+ is indexed by the real extension parameter β .

3 Homogeneous problem $\tilde{S}u = 0$

In this Section we identify the dimension and the basis of the subspace $\ker S^*$, and prove Proposition 2.2. One has to solve the homogeneous differential equation $\tilde{S}u = 0$, where \tilde{S} is the differential operator (2.4) and the function $r \mapsto u(r) = \begin{pmatrix} u^+(r) \\ u^-(r) \end{pmatrix}$ on \mathbb{R}^+ is the spinorial unknown. The needed ODE technique is classical and we include it concisely for completeness. Observe, however, that for the application of von Neumann’s theory of self-adjoint extension one has to solve the ODE problem $\tilde{S}u = zu$ for non-real $z \in \mathbb{C}$, say, $z = \pm i$, which requires a somewhat more extended discussion – see, e.g., [35, Section 3] or [21, Sections 4-6].

Upon transforming the unknown u into φ , where

$$\varphi(r) := \frac{1}{2} (\mathbf{A}u) \begin{pmatrix} 1 \\ 2 \end{pmatrix} e^{r/2}, \quad \mathbf{A} := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad (3.1)$$

the differential system $\tilde{S}u = 0$ takes the form

$$\begin{cases} (\varphi^+)' = \varphi^+ - \frac{1-\nu}{r} \varphi^- \\ (\varphi^-)' = -\frac{1+\nu}{r} \varphi^+. \end{cases} \quad (3.2)$$

Therefore, φ^- is a solution to

$$r(\varphi^-)'' + (1-r)(\varphi^-)' - \frac{v^2-1}{r}\varphi^- = 0, \quad (3.3)$$

equivalently,

$$\xi(r) := r^B \varphi^-(r) \quad (3.4)$$

is a solution to

$$r\xi'' + (1-2B-r)\xi' + B\xi = 0. \quad (3.5)$$

The second order ODE (3.5) is the confluent hypergeometric equation [1, (13.1.1) and (13.1.11)], two linearly independent solutions of which are the confluent hypergeometric functions of first and second kind, that is, respectively, the Kummer function $M_{a,b}(r)$ [1, (13.1.2)] and the Tricomi function $U_{a,b}(r)$ [1, (13.1.3)], with $a = -B$ and $b = 1 - 2B$.

The solutions $\xi_0(r) := M_{-B,1-2B}(r)$ and $\xi_\infty(r) := U_{-B,1-2B}(r)$ to (3.5) determine, via (3.4) and the second of (3.2), two linearly independent solutions $\varphi_0 = \begin{pmatrix} \varphi_0^+ \\ \varphi_0^- \end{pmatrix}$ and $\varphi_\infty = \begin{pmatrix} \varphi_\infty^+ \\ \varphi_\infty^- \end{pmatrix}$ to (3.2). Using the properties

$$M'_{a,b}(r) = \frac{a}{b} M_{a+1,b+1}(r), \quad U'_{a,b}(r) = -a U_{a+1,b+1}(r)$$

([1, (13.4.8) and (13.4.21)]), and the inverse transformation of (3.1), that is, $u(r) = 2e^{-r/2}(\mathbf{A}^{-1}\varphi)(2r)$, where $\mathbf{A}^{-1} = \mathbf{A}$, yields the following two linearly independent solutions to the original problem $\tilde{S}u = 0$:

$$\begin{aligned} u_0(r) &:= \frac{1}{e^r r^B} \begin{pmatrix} \frac{1+v+B}{1+v} M_{-B,1-2B}(2r) + \frac{2rB}{(1+v)(1-2B)} M_{1-B,2-2B}(2r) \\ -\frac{1+v-B}{1+v} M_{-B,1-2B}(2r) + \frac{2rB}{(1+v)(1-2B)} M_{1-B,2-2B}(2r) \end{pmatrix} \\ u_\infty(r) &:= \frac{1}{e^r r^B} \begin{pmatrix} \frac{1+v+B}{1+v} U_{-B,1-2B}(2r) - \frac{2rB}{1+v} U_{1-B,2-2B}(2r) \\ -\frac{1+v-B}{1+v} U_{-B,1-2B}(2r) - \frac{2rB}{1+v} U_{1-B,2-2B}(2r) \end{pmatrix} \end{aligned} \quad (3.6)$$

(in fact, an irrelevant common pre-factor 2^{-B} has been neglected). Both u_0 and u_∞ are real-valued and smooth on \mathbb{R}^+ .

Because of the asymptotics [1, (13.1.2), (13.5.1), and (13.5.5)]

$$\begin{aligned} M_{a,b}(r) &= \frac{e^r r^{a-b}}{\Gamma(a)} (1 + O(r^{-1})) \quad \text{as } r \rightarrow +\infty \\ M_{a,b}(r) &= 1 + O(r) \quad \text{as } r \downarrow 0 \quad \text{and } -b \notin \mathbb{N} \end{aligned} \quad (3.7)$$

and [1, (13.1.2), (13.1.3), (13.5.2), (13.5.8), and (13.5.10)]

$$\begin{aligned}
 U_{a,b}(r) &= r^{-a}(1 + O(r^{-1})) && \text{as } r \rightarrow +\infty \\
 U_{a,b}(r) &= \frac{\Gamma(1-b)}{\Gamma(1+a-b)} + \frac{\Gamma(b-1)}{\Gamma(a)} r^{1-b} + O(r) && \begin{array}{l} \text{as } r \downarrow 0 \\ \text{and } b \in (0, 1) \end{array} \\
 U_{a,b}(r) &= \frac{\Gamma(b-1)}{\Gamma(a)} r^{-(b-1)} + O(1) && \begin{array}{l} \text{as } r \downarrow 0 \\ \text{and } b \in (1, 2), \end{array}
 \end{aligned} \tag{3.8}$$

one deduces that both u_0 and u_∞ are square-integrable around $r = 0$, whereas only u_∞ is square-integrable at infinity, and moreover

$$\begin{aligned}
 u_0(r) &= \left(\begin{array}{c} \frac{1+\nu+B}{1+\nu} \\ -\frac{1+\nu-B}{1+\nu} \end{array} \right) r^{-B} + O(r^{1-B}) && \text{as } r \downarrow 0 \\
 u_0(r) &= -\frac{2^B(1-2B)}{\Gamma(-B)(1+\nu)} \begin{pmatrix} 1 \\ 1 \end{pmatrix} r^B e^r (1 + O(r^{-1})) && \text{as } r \rightarrow +\infty,
 \end{aligned} \tag{3.9}$$

and

$$\begin{aligned}
 u_\infty(r) &= \frac{\Gamma(2B)}{\Gamma(B)} \left(\begin{array}{c} \frac{1+\nu-B}{1+\nu} \\ -\frac{1+\nu+B}{1+\nu} \end{array} \right) r^{-B} + \begin{pmatrix} q^+ \\ q^- \end{pmatrix} r^B + O(r^{1-B}) && \text{as } r \downarrow 0 \\
 u_\infty(r) &= 2^B \begin{pmatrix} 1 \\ -1 \end{pmatrix} r^{-B} e^{-r} (1 + O(r^{-1})) && \text{as } r \rightarrow +\infty,
 \end{aligned} \tag{3.10}$$

where

$$q^\pm := \frac{4^B(-B \pm (1+\nu))\Gamma(-2B)}{(1+\nu)\Gamma(-B)}. \tag{3.11}$$

Observe that $q^\pm \neq 0$.

Therefore, there is only a *one*-dimensional space of solutions to $\tilde{S}u = 0$ which are square integrable, and hence, owing to (2.5), $\ker S^*$ is one-dimensional. For convenience, let us choose as the spanning vector the function $\Phi := u_\infty$. Then (3.10) implies (2.9) and Proposition 2.2 is proved.

4 Distinguished extension S_D

In this Section we qualify the distinguished extension S_D of the operator S , and prove Proposition 2.3.

When comparing the approach based on von Neumann's theory, as developed, e.g., in [21], with the present one based on the Kreĭn-Višik-Birman theory, to solve the homogeneous problem $S^*u = 0$ (in the KVB strategy) or to solve the deficiency space problem $S^*u = \pm iu$ (in the von Neumann strategy) are two essentially analogous versions of the same step, from the ODE viewpoint. In contrast, the qualification of S_D (in view of Theorem 2.1, strictly speaking one only needs to qualify the action of S_D^{-1} on $\ker S^*$) is a specific step of the KVB strategy, and it boils down to solving the

ODE problem $\tilde{S}f = g$ for given g . Along this line, we adapt to our case the analysis done in [8] for homogeneous Schrödinger operators on half-line.

In order to set up the problem conveniently, let us first replace the pair (u_0, u_∞) of linearly independent solutions (3.6) to $\tilde{S}u = 0$ to the new pair (v_0, v_∞) given by

$$\begin{aligned} v_0 &:= u_\infty - \frac{\Gamma(2B)}{\Gamma(B)} u_0 \\ v_\infty &:= u_\infty. \end{aligned} \quad (4.1)$$

This preserves the linear independence of v_0 and v_∞ with the virtue of having two solutions with different power-law in the asymptotics as $r \downarrow 0$: from (4.1) and (3.9)-(3.10) we find

$$\begin{aligned} v_0(r) &= \begin{pmatrix} q^+ \\ q^- \end{pmatrix} r^B + O(r^{1-B}) \\ v_\infty(r) &= \frac{\Gamma(2B)}{\Gamma(B)} \begin{pmatrix} \frac{1+v+B}{1+v} \\ -\frac{1+v-B}{1+v} \end{pmatrix} r^{-B} + O(r^B) \end{aligned} \quad \text{as } r \downarrow 0, \quad (4.2)$$

where q^\pm is given by (3.11). At large distances, v_0 and v_∞ have exponential asymptotics as u_0 and u_∞ , namely

$$\begin{aligned} v_0(r) &= -\frac{1}{2} \frac{2^B B}{(1+v) \cos(B\pi)} \begin{pmatrix} 1 \\ 1 \end{pmatrix} r^B e^r (1 + O(r^{-1})) \\ v_\infty(r) &= 2^B \begin{pmatrix} 1 \\ -1 \end{pmatrix} r^{-B} e^{-r} (1 + O(r^{-1})) \end{aligned} \quad \text{as } r \rightarrow +\infty. \quad (4.3)$$

We then proceed with standard ODE arguments. With respect to the fundamental system (v_0, v_∞) , the general solution to the inhomogeneous problem $\tilde{S}f = g$ has the form

$$f = A_0 v_0 + A_\infty v_\infty + f_{\text{part}}, \quad (4.4)$$

where A_0 and A_∞ run over \mathbb{C} and f_{part} is a particular solution, namely, $\tilde{S}f_{\text{part}} = g$. Let us determine it through the variation of constants [36, Section 2.4].

First we re-write $\tilde{S}f = g$ in normal form as

$$y' + \mathbf{V}(r)y = g, \quad y := \mathbf{E}f, \quad (4.5)$$

where

$$\mathbf{V}(r) := \frac{1}{r} \begin{pmatrix} -1 & v \\ -v & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \mathbf{E} := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \quad (4.6)$$

We also introduce the Wronskian

$$\mathbb{R}^+ \ni r \mapsto W_r(v_0, v_\infty) := \det \begin{pmatrix} v_0^+(r) & v_\infty^+(r) \\ v_0^-(r) & v_\infty^-(r) \end{pmatrix}. \quad (4.7)$$

This is precisely the Wronskian $W_r(\mathbf{E}v_0, \mathbf{E}v_\infty)$ of two fundamental solutions of the problem written in normal form, because $\det \mathbf{E} = 1$ and hence $W_r(\mathbf{E}v_0, \mathbf{E}v_\infty) =$

$W_r(v_0, v_\infty)$. Moreover, since $\mathbf{V}(r)$ is traceless for any $r \in \mathbb{R}^+$, Liouville's theorem implies that $W_r(-\mathbf{E}v_0, -\mathbf{E}v_\infty)$ is constant, and so is also $W_r(v_0, v_\infty)$. Therefore,

$$W_r(v_0, v_\infty) = \lim_{r \downarrow 0} W_r(v_0, v_\infty) = \frac{4^B B}{(1+v)\cos(B\pi)} =: W_0^\infty. \quad (4.8)$$

The limit in (4.8) above follows straightforwardly from the asymptotics (4.2) and from the expression (3.11) for q^\pm . Clearly, $W_0^\infty \neq 0$. Then a standard application of the method of variation of constants for the differential problem (4.5) and the further transformation $f = \mathbf{E}y$ yields eventually

$$f_{\text{part}}(r) = \int_0^{+\infty} G(r, \rho) g(\rho) d\rho, \quad (4.9)$$

where

$$G(r, \rho) := \begin{cases} \frac{1}{W_0^\infty} \begin{pmatrix} v_\infty^+(r)v_0^+(\rho) & v_\infty^+(r)v_0^-(\rho) \\ v_\infty^-(r)v_0^+(\rho) & v_\infty^-(r)v_0^-(\rho) \end{pmatrix} & \text{if } 0 < \rho < r \\ \frac{1}{W_0^\infty} \begin{pmatrix} v_0^+(r)v_\infty^+(\rho) & v_0^+(r)v_\infty^-(\rho) \\ v_0^-(r)v_\infty^+(\rho) & v_0^-(r)v_\infty^-(\rho) \end{pmatrix} & \text{if } 0 < r < \rho. \end{cases} \quad (4.10)$$

Next, we observe the following.

Lemma 4.1 *The integral operator R_G on $L^2(\mathbb{R}^+, \mathbb{C}^2, dr)$ with kernel $G(r, \rho)$ given by (4.10) is bounded and self-adjoint.*

Proof For each $r, \rho \in \mathbb{R}^+$, $G(r, \rho)$ is the sum of the four terms

$$\begin{aligned} G^{++}(r, \rho) &:= G(r, \rho) \mathbf{1}_{(1,+\infty)}(r) \mathbf{1}_{(1,+\infty)}(\rho) \\ G^{+-}(r, \rho) &:= G(r, \rho) \mathbf{1}_{(1,+\infty)}(r) \mathbf{1}_{(0,1)}(\rho) \\ G^{-+}(r, \rho) &:= G(r, \rho) \mathbf{1}_{(0,1)}(r) \mathbf{1}_{(1,+\infty)}(\rho) \\ G^{--}(r, \rho) &:= G(r, \rho) \mathbf{1}_{(0,1)}(r) \mathbf{1}_{(0,1)}(\rho), \end{aligned} \quad (4.11)$$

where $\mathbf{1}_J$ denotes the characteristic function of the interval $J \subset \mathbb{R}^+$, and correspondingly R_G splits into the sum of four integral operators with kernel given by (4.11).

Now, for each entry of $G^{LM}(r, \rho)$, with $L, M \in \{+, -\}$, a point-wise estimate in (r, ρ) can be derived from the short and large distance asymptotics for v_0 and v_∞ . For example, the entry $G_{11}^{++}(r, \rho)$ in the first row and first column of $G^{++}(r, \rho)$ is controlled as

$$\begin{aligned} |v_\infty^+(r) v_0^+(\rho) \mathbf{1}_{(1,+\infty)}(r) \mathbf{1}_{(1,+\infty)}(\rho)| &\lesssim e^{-r} e^\rho (\rho/r)^B & \text{if } 0 < \rho < r \\ |v_0^+(r) v_\infty^+(\rho) \mathbf{1}_{(1,+\infty)}(r) \mathbf{1}_{(1,+\infty)}(\rho)| &\lesssim e^r e^{-\rho} (r/\rho)^B & \text{if } 0 < r < \rho, \end{aligned}$$

because v_0 diverges exponentially and v_∞ vanishes exponentially as $r \rightarrow +\infty$, (4.3); thus,

$$|G_{11}^{++}(r, \rho)| \lesssim e^{-|r-\rho|} (\rho r)^B.$$

In fact, the asymptotics for v_0 and v_∞ are the same for both components, so we can also conclude that

$$\|G^{++}(r, \rho)\|_{M_2(\mathbb{C})} \lesssim e^{-|r-\rho|} (\rho r)^B,$$

where $\|\cdot\|_{M_2(\mathbb{C})}$ denotes the matrix norm. The estimate of the other kernels is perfectly analogous, and we find

$$\begin{aligned} \|G^{++}(r, \rho)\|_{M_2(\mathbb{C})} &\lesssim (r\rho)^B e^{-|r-\rho|} \mathbf{1}_{(1,+\infty)}(r) \mathbf{1}_{(1,+\infty)}(\rho) \\ \|G^{+-}(r, \rho)\|_{M_2(\mathbb{C})} &\lesssim r^B e^{-\rho} \mathbf{1}_{(1,+\infty)}(r) \mathbf{1}_{(0,1)}(\rho) \\ \|G^{-+}(r, \rho)\|_{M_2(\mathbb{C})} &\lesssim e^{-r} \rho^B \mathbf{1}_{(0,1)}(r) \mathbf{1}_{(1,+\infty)}(\rho) \\ \|G^{--}(r, \rho)\|_{M_2(\mathbb{C})} &\lesssim (r^B \rho^{-B} + r^{-B} \rho^B) \mathbf{1}_{(0,1)}(r) \mathbf{1}_{(0,1)}(\rho). \end{aligned} \quad (4.12)$$

The last three estimates in (4.12) show at once that the kernels $G^{+-}(r, \rho)$, $G^{-+}(r, \rho)$, and $G^{--}(r, \rho)$ are in $L^2(\mathbb{R}^+ \times \mathbb{R}^+, M_2(\mathbb{C}), dr d\rho)$ and therefore the corresponding integral operators are Hilbert-Schmidt operators, hence bounded, on $L^2(\mathbb{R}^+, \mathbb{C}^2, dr)$. The first estimate in (4.12) allows to conclude, by an obvious Schur test, that also the integral operator with kernel $G^{++}(r, \rho)$ is bounded on $L^2(\mathbb{R}^+, \mathbb{C}^2, dr)$. This proves the overall boundedness of R_G .

The self-adjointness of R_G is clear from (4.10): the adjoint R_G^* of R_G has kernel $\overline{G(\rho, r)^T}$, but G is real-valued and $G(\rho, r) = G(r, \rho)$, thus proving that $R_G^* = R_G$. \square

The integral operator R_G has a relevant mapping property that is more directly read out from the following alternative representation. If $f_{\text{part}} = R_G g$, then

$$f_{\text{part}}(r) = \Theta_\infty^{(g)}(r) v_0(r) + \Theta_0^{(g)}(r) v_\infty(r), \quad (4.13)$$

where

$$\begin{aligned} \Theta_0^{(g)}(r) &:= \frac{1}{W_0^\infty} \int_0^r \langle \overline{v_0(\rho)}, g(\rho) \rangle_{\mathbb{C}^2} d\rho \\ \Theta_\infty^{(g)}(r) &:= \frac{1}{W_0^\infty} \int_r^{+\infty} \langle \overline{v_\infty(\rho)}, g(\rho) \rangle_{\mathbb{C}^2} d\rho \end{aligned} \quad (4.14)$$

and W_0^∞ is the constant computed in (4.8). Indeed, from (4.10),

$$\begin{aligned}
 f_{\text{part}}(r) &= \int_0^{+\infty} G(r, \rho) g(\rho) \, d\rho \\
 &= \frac{1}{W_0^\infty} \begin{pmatrix} v_\infty^+(r) \\ v_\infty^-(r) \end{pmatrix} \int_0^r (v_0^+(\rho)g^+(\rho) + v_0^-(\rho)g^-(\rho)) \, d\rho \\
 &\quad + \frac{1}{W_0^\infty} \begin{pmatrix} v_0^+(r) \\ v_0^-(r) \end{pmatrix} \int_r^{+\infty} (v_\infty^+(\rho)g^+(\rho) + v_\infty^-(\rho)g^-(\rho)) \, d\rho,
 \end{aligned}$$

that is, (4.13).

Lemma 4.2 *For every $g \in L^2(\mathbb{R}^+, \mathbb{C}^2)$ one has*

$$\int_0^{+\infty} \frac{\|(R_G g)(r)\|_{\mathbb{C}^2}^2}{r} \, dr < +\infty, \quad (4.15)$$

i.e.,

$$\text{ran } R_G \subset \mathcal{D}[r^{-1}]. \quad (4.16)$$

Proof It suffices to prove the finiteness of the integral in (4.15) only for $r \in (0, 1)$, since $\int_1^{+\infty} r^{-1} \|(R_G g)(r)\|_{\mathbb{C}^2}^2 \, dr \leq \|R_G\|^2 \|g\|_{L^2(\mathbb{R}^+, \mathbb{C}^2)}^2$. Let us represent $f = R_G g \in \text{ran } R_G$ as in (4.13)-(4.14). For $r \in (0, 1)$ one has

$$\begin{aligned}
 |\Theta_0^{(g)}(r)| &\leq |W_0^\infty|^{-1} \|v_0 \mathbf{1}_{(0,r)}\|_{L^2(\mathbb{R}^+, \mathbb{C}^2)} \|g\|_{L^2(\mathbb{R}^+, \mathbb{C}^2)} \leq C_{g,v} r^{B+\frac{1}{2}} \\
 |\Theta_\infty^{(g)}(r)| &\leq |W_0^\infty|^{-1} \|v_\infty \mathbf{1}_{(r,\infty)}\|_{L^2(\mathbb{R}^+, \mathbb{C}^2)} \|g\|_{L^2(\mathbb{R}^+, \mathbb{C}^2)} \leq C_{g,v}
 \end{aligned}$$

for some constant $C_{g,v} > 0$ depending on g and v only, having used the short distance asymptotics (4.2) for v_0 and v_∞ . Combining now the above bounds again with (4.2) we see that on the interval $(0, 1)$ the functions $r \mapsto \Theta_0^{(g)}(r) v_\infty(r)$ and $r \mapsto \Theta_\infty^{(g)}(r) v_0(r)$ are continuous and vanish when $r \rightarrow 0$, respectively, as $r^{1/2}$ and r^B , which makes the integral $\int_0^1 r^{-1} \|(R_G g)(r)\|_{\mathbb{C}^2}^2 \, dr$ finite. \square

Combining Lemmas 4.1 and 4.2 together, we are now in the condition to prove Proposition 2.3.

Proof of Proposition 2.3 (i) and (ii). The integral operator R_G on $L^2(\mathbb{R}^+, \mathbb{C})$ with kernel given by (4.10) is bounded and self-adjoint owing to Lemma 4.1, and by construction satisfies $\tilde{S} R_G g = g \forall g \in L^2(\mathbb{R}^+, \mathbb{C})$. Therefore, there is one self-adjoint extension \mathcal{S} of $S_{\min} = \tilde{S}$ such that $\mathcal{S} R_G g = g \forall g \in L^2(\mathbb{R}^+, \mathbb{C})$, whence, by self-adjointness, also $R_G \mathcal{S} h = h \forall h \in \mathcal{D}(\mathcal{S})$. Thus, $R_G = \mathcal{S}^{-1}$ for some invertible self-adjoint realisation \mathcal{S} of S . Because of Lemma 4.2, the space $\mathcal{D}(\mathcal{S}) = \text{ran } R_G$ is contained in the potential energy form domain $\mathcal{D}[r^{-1}]$: owing to Theorem 1.1(ii) then \mathcal{S} must be the reduction to the subspace $\mathcal{H}_{\frac{1}{2},1}$ of the distinguished self-adjoint extension of the Dirac-Coulomb operator H : we shall denote it with S_D . As such, S_D is the unique self-adjoint realisation of S satisfying the property (2.10), it is invertible, and its kernel is precisely given by (4.10).

(iii) The decompositions (2.11) and (2.12) are canonical, once a self-adjoint extension of S is given with everywhere defined and bounded inverse: see, for instance, [16, (2.4) and (2.5)].

(iv) From the previous discussion, $\Phi = u_\infty = v_\infty$ and $S_D^{-1}\Phi = R_G v_\infty$. A closed expression for the latter function is given by (4.13) above, which now reads

$$S_D^{-1}\Phi = \Theta_\infty^{(v_\infty)}(r) v_0(r) + \Theta_0^{(v_\infty)}(r) v_\infty(r).$$

From (4.14) and (4.2) we deduce

$$\begin{aligned} |\Theta_0^{(v_\infty)}(r)| &\leq |W_0^\infty|^{-1} \int_0^r |\langle \overline{v_0(\rho)}, v_\infty(\rho) \rangle_{\mathbb{C}^2}| d\rho \\ &\lesssim \int_0^r (\rho^B + O(\rho^{1-B}))(\rho^{-B} + O(\rho^B)) d\rho \\ &= r + o(r) \quad \text{as } r \downarrow 0 \end{aligned}$$

and

$$\begin{aligned} \Theta_\infty^{(v_\infty)}(r) &= \frac{1}{W_0^\infty} \int_r^{+\infty} \langle \overline{v_\infty(\rho)}, v_\infty(\rho) \rangle_{\mathbb{C}^2} d\rho \\ &= \frac{1}{W_0^\infty} \|v_\infty\|_{L^2(\mathbb{R}^+, \mathbb{C}^2)}^2 (1 + o(1)) \quad \text{as } r \downarrow 0. \end{aligned}$$

Therefore, using again the short distance asymptotics (4.2),

$$(S_D^{-1}\Phi)(r) = \frac{\|v_\infty\|_{L^2(\mathbb{R}^+, \mathbb{C}^2)}^2}{W_0^\infty} \begin{pmatrix} q^+ \\ q^- \end{pmatrix} r^B + o(r^B) \quad (*)$$

where q^\pm is given by (3.11). Upon setting

$$p^\pm := q^\pm (W_0^\infty)^{-1} \|v_\infty\|_{L^2(\mathbb{R}^+, \mathbb{C}^2)}^2 \quad (4.17)$$

we then obtain the leading term of (2.13). The remainder is in fact smaller than $o(r^B)$. This can be seen by comparing the above asymptotics for $S_D^{-1}\Phi$ with the expansion (5.2) established in the next Section (which is valid because $S_D^{-1}\Phi \in \mathcal{D}(S_D) \subset \mathcal{D}(S^*)$), namely

$$S_D^{-1}\Phi = a_0^{(S_D^{-1}\Phi)} v_0 + a_\infty^{(S_D^{-1}\Phi)} v_\infty + b_\infty^{(S_D^{-1}\Phi)} v_0 + b_0^{(S_D^{-1}\Phi)} v_\infty.$$

For the latter, we have the asymptotics

$$\begin{aligned} (S_D^{-1}\Phi)(r) &= \mathbf{c}_0 a_0^{(S_D^{-1}\Phi)} (r^B + O(r^{1-B})) + \mathbf{c}_\infty a_\infty^{(S_D^{-1}\Phi)} (r^{-B} + O(r^B)) \\ &\quad + o(r^{1/2}) \quad \text{as } r \downarrow 0. \end{aligned} \quad (**)$$

as follows from (4.2) and (5.3) for some non-zero constants $\mathbf{c}_0, \mathbf{c}_\infty \in \mathbb{C}^2$. In order for (*) and (**) to be compatible, necessarily $a_\infty^{(S_D^{-1}\Phi)} = 0$. This implies that after the leading order r^B there comes a remainder $o(r^{1/2})$, thus completing the proof of (2.13). \square

5 Operator closure \overline{S}

This Section is devoted to the proof of Proposition 2.4. In fact we will prove a stronger result of characterisation of $\mathcal{D}(\overline{S})$, namely Proposition 5.3 below, from which Proposition 2.4 follows as a corollary.

Let us start with another useful representation of $\mathcal{D}(S^*)$. It is analogous to the operator-theoretic decomposition (2.11), but its formulation (and proof) is more in the ODE spirit.

Lemma 5.1 *For each $g \in \mathcal{D}(S^*)$ there exist, uniquely determined, constants $a_0^{(g)}, a_\infty^{(g)} \in \mathbb{C}$ and functions*

$$\begin{aligned} b_0^{(g)}(r) &:= \frac{1}{W_0^\infty} \int_0^r \langle \overline{v_0(\rho)}, (S^*g)(\rho) \rangle_{\mathbb{C}^2} d\rho \\ b_\infty^{(g)}(r) &:= -\frac{1}{W_0^\infty} \int_0^r \langle \overline{v_\infty(\rho)}, (S^*g)(\rho) \rangle_{\mathbb{C}^2} d\rho \end{aligned} \quad (5.1)$$

on \mathbb{R}^+ such that

$$g = a_0^{(g)} v_0 + a_\infty^{(g)} v_\infty + b_\infty^{(g)} v_0 + b_0^{(g)} v_\infty, \quad (5.2)$$

where v_0 and v_∞ are the two linearly independent solutions (4.1) to the homogeneous problem $\tilde{S}v = 0$ (recall that they are real and smooth on \mathbb{R}^+) and W_0^∞ is the constant computed in (4.8). Moreover, both $b_0^{(g)}(r)$ and $b_\infty^{(g)}(r)$ vanish as $r \downarrow 0$, and

$$b_\infty^{(g)}(r) v_0(r) + b_0^{(g)}(r) v_\infty(r) = o(r^{1/2}) \quad \text{as } r \downarrow 0. \quad (5.3)$$

Proof Let $h := S^*g = \tilde{S}g$. Then, as already argued in (4.4) and (4.13)–(4.14), g is expressed in terms of h as

$$g = A_0 v_0 + A_\infty v_\infty + \Theta_\infty^{(h)} v_0 + \Theta_0^{(h)} v_\infty \quad (*)$$

for some $A_0, A_\infty \in \mathbb{C}$ that are now uniquely identified by g . From (4.14) and (5.1) we see that

$$\begin{aligned} \Theta_\infty^{(h)}(r) &= b_\infty^{(g)}(r) \\ \Theta_0^{(h)}(r) &= -\frac{1}{W_0^\infty} \int_0^r \langle \overline{v_0(\rho)}, (S^*g)(\rho) \rangle_{\mathbb{C}^2} d\rho \\ &= b_0^{(g)}(r) - \frac{1}{W_0^\infty} \int_0^{+\infty} \langle \overline{v_0(\rho)}, (S^*g)(\rho) \rangle_{\mathbb{C}^2} d\rho. \end{aligned}$$

Then (*) implies (5.2) at once, upon setting

$$\begin{aligned} a_0^{(g)} &:= A_0 - \frac{1}{W_0^\infty} \int_0^{+\infty} \overline{\langle v_0(\rho), (S^*g)(\rho) \rangle_{\mathbb{C}^2}} d\rho \\ a_\infty^{(g)} &:= A_\infty \end{aligned}$$

Observe that the constant added above to A_0 is finite and bounded by $|W_0^\infty|^{-1} \|v_0\|_{L^2(\mathbb{R}^+, \mathbb{C}^2)} \|S^*g\|_{L^2(\mathbb{R}^+, \mathbb{C}^2)}$. As for the proof of (5.3), by means of the short distance asymptotics (4.2) for v_0 and v_∞ we find

$$\begin{aligned} |b_\infty^{(g)}(r) v_0(r)| &\lesssim r^{-B} \int_0^r \rho^B \|g(\rho)\|_{\mathbb{C}^2} d\rho \leq \int_0^r \|g(\rho)\|_{\mathbb{C}^2} d\rho \\ &\leq r^{1/2} \|g\|_{L^2([0, r], \mathbb{C}^2)} = o(r^{1/2}) \end{aligned}$$

and

$$\begin{aligned} |b_0^{(g)}(r) v_\infty(r)| &\lesssim r^B \int_0^r \rho^{-B} \|g(\rho)\|_{\mathbb{C}^2} d\rho \leq r^B \|\rho^{-B}\|_{L^2[0, r]} \|g\|_{L^2([0, r], \mathbb{C}^2)} \\ &\lesssim r^{1/2} \|g\|_{L^2([0, r], \mathbb{C}^2)} = o(r^{1/2}), \end{aligned}$$

and (5.3) then follows.

The next preparatory step is to introduce, for later convenience, the Wronskian of any two square-integrable functions,

$$\mathbb{R}^+ \ni r \mapsto W_r(\psi, \phi) := \det \begin{pmatrix} \psi^+(r) & \phi^+(r) \\ \psi^-(r) & \phi^-(r) \end{pmatrix}, \quad \psi, \phi \in L^2(\mathbb{R}^+, \mathbb{C}^2), \quad (5.4)$$

and the boundary form for any two functions in $\mathcal{D}(S^*)$,

$$\omega(g, h) := \langle S^*g, h \rangle - \langle g, S^*h \rangle, \quad g, h \in \mathcal{D}(S^*). \quad (5.5)$$

The boundary form is antisymmetric, i.e.,

$$\omega(h, g) = -\overline{\omega(h, g)}, \quad (5.6)$$

and it is related to the Wronskian by

$$\omega(g, h) = -\lim_{r \downarrow 0} W_r(\bar{g}, h). \quad (5.7)$$

Indeed, using $\tilde{S} = \mathbf{E} \frac{d}{dr} + \mathbf{V}(r)$ from (4.5)–(4.6), one has

$$\begin{aligned} \omega(g, h) &= \int_0^{+\infty} dr \left(\langle (\tilde{S}g)(r), h(r) \rangle_{\mathbb{C}^2} - \langle g(r), (\tilde{S}h)(r) \rangle_{\mathbb{C}^2} \right) \\ &= \int_0^{+\infty} dr \left(\langle \mathbf{E}g'(r), h(r) \rangle_{\mathbb{C}^2} - \langle g(r), \mathbf{E}h'(r) \rangle_{\mathbb{C}^2} \right) \\ &= \lim_{r \downarrow 0} \left(\overline{g^-(r)} h^+(r) - \overline{g^+(r)} h^-(r) \right) = - \lim_{r \downarrow 0} W_r(\bar{g}, h). \end{aligned}$$

It is also convenient to introduce the (two-dimensional) space of solutions to the differential problem $\tilde{S}v = 0$,

$$\mathcal{L} := \{v : \mathbb{R}^+ \rightarrow \mathbb{C}^2 \mid \tilde{S}v = 0\} = \text{span}\{v_0, v_\infty\}, \quad (5.8)$$

As well known, $r \mapsto W_r(u, v)$ is constant whenever $u, v \in \mathcal{L}$, and this constant is zero if and only if u and v are linearly dependent. It will be important also to keep into account that any $v \in \mathcal{L}$ is square-integrable around $r = 0$, as determined in (4.2).

Lemma 5.2 *For given $v \in \mathcal{L}$,*

$$\begin{aligned} L_v : \mathcal{D}(S^*) &\rightarrow \mathbb{C} \\ g &\longmapsto L_v(g) := \lim_{r \downarrow 0} W_r(\bar{v}, g) \end{aligned} \quad (5.9)$$

defines a linear functional on $\mathcal{D}(S^)$ which vanishes on $\mathcal{D}(\bar{S})$.*

Proof The linearity of L_v is obvious, and the finiteness of $L_v(g)$ for $g \in \mathcal{D}(S^*)$ is checked as follows. Let us decompose $g = a_0^{(g)} v_0 + a_\infty^{(g)} v_\infty + b_\infty^{(g)} v_0 + b_0^{(g)} v_\infty$ as in (5.2) and $v = c_0 v_0 + c_\infty v_\infty$ in the basis of \mathcal{L} . Owing to (5.9), it suffices to control the finiteness of $L_{v_0}(g)$ and $L_{v_\infty}(g)$. By linearity,

$$L_{v_0}(g) = a_0^{(g)} L_{v_0}(v_0) + a_\infty^{(g)} L_{v_0}(v_\infty) + L_{v_0}(b_\infty^{(g)} v_0 + b_0^{(g)} v_\infty);$$

moreover, $L_{v_0}(v_0) = \lim_{r \downarrow 0} W_r(v_0, v_0) = 0$, $L_{v_0}(v_\infty) = W_0^\infty$, $L_{v_0}(b_\infty^{(g)} v_0) = \lim_{r \downarrow 0} W_r(v_0, b_\infty^{(g)} v_0) = \lim_{r \downarrow 0} b_\infty^{(g)}(r) W_r(v_0, v_0) = 0$, and $L_{v_0}(b_0^{(g)} v_\infty) = \lim_{r \downarrow 0} W_r(v_0, b_0^{(g)} v_\infty) = \lim_{r \downarrow 0} b_0^{(g)}(r) W_r(v_0, v_\infty) = 0$. The conclusion is $L_{v_0}(g) = a_\infty^{(g)} W_0^\infty$. Analogously, $L_{v_\infty}(g) = -a_0^{(g)} W_0^\infty$, and this establishes the finiteness of $L_v(g)$. Let us now prove now that if $f \in \mathcal{D}(\bar{S})$, then $L_v(f) = 0$. Let $\chi \in C_0^\infty([0, +\infty))$ be such that $\chi(r) = 1$ for $r \in [0, \frac{1}{2}]$ and $\chi(r) = 0$ for $r \in [1, +\infty)$. One has that $v\chi \in \mathcal{D}(S^*)$, indeed $v\chi \in L^2(\mathbb{R}^+, \mathbb{C}^2)$ and

$$\begin{aligned} \tilde{S}(v\chi) &= \left(\mathbf{E} \frac{d}{dr} + \mathbf{V}(r) \right) v\chi = \chi \left(\mathbf{E} \frac{d}{dr} + \mathbf{V}(r) \right) v + \mathbf{E}v\chi' \\ &= (\tilde{S}v)\chi + \mathbf{E}v\chi' = \mathbf{E}v\chi' \in L^2(\mathbb{R}^+, \mathbb{C}^2), \end{aligned}$$

where we used $\tilde{S} = \mathbf{E} \frac{d}{dr} + \mathbf{V}(r)$ and $\tilde{S}v = 0$. Moreover, because of the behaviour of χ around $r = 0$, the Wronskians $W_r(\overline{v\chi}, g)$ and $W_r(\overline{v}, g)$ are asymptotically equal as $r \downarrow 0$, that is, $L_{v\chi} = L_v$. As a consequence of this latter fact and of (5.7),

$$\begin{aligned} L_v(f) &= L_{v\chi}(f) = \lim_{r \downarrow 0} W_r(\overline{v\chi}, f) = -\omega(v\chi, f) \\ &= \langle v\chi, S^* f \rangle - \langle S^*(v\chi), f \rangle = \langle v\chi, \overline{S}f \rangle - \langle v\chi, \overline{S}f \rangle = 0, \end{aligned}$$

which completes the proof. \square

We come now to the characterisation of the space $\mathcal{D}(\overline{S})$ which constitutes the main result of this Section.

Proposition 5.3 *Let $f \in \mathcal{D}(S^*)$. The following conditions are equivalent:*

- (i) $f \in \mathcal{D}(\overline{S})$.
- (ii) $\omega(f, g) = 0$ for all $g \in \mathcal{D}(S^*)$.
- (iii) $L_v(f) = 0$ for all $v \in \mathcal{L}$.
- (iv) With respect to the decomposition (5.2) for f , $a_0^{(f)} = a_\infty^{(f)} = 0$.

Proof The implication (i) \Rightarrow (ii) follows at once from

$$\omega(f, g) = \langle S^* f, g \rangle - \langle f, S^* g \rangle = \langle \overline{S}f, g \rangle - \langle \overline{S}f, g \rangle = 0.$$

For the converse implication (ii) \Rightarrow (i), we observe that

$$0 = \omega(f, g) = \langle S^* f, g \rangle - \langle f, S^* g \rangle \quad \forall g \in \mathcal{D}(S^*)$$

is equivalent to $\langle S^* f, g \rangle = \langle f, S^* g \rangle \quad \forall g \in \mathcal{D}(S^*)$, which implies that $f \in \mathcal{D}(S^{**}) = \mathcal{D}(\overline{S})$.

The implication (i) \Rightarrow (iii) is given by Lemma 5.2. Conversely, let us assume that $L_v(f) = 0$ for all $v \in \mathcal{L}$, and let us prove that for such f one has $\omega(f, g) = 0$ for all $g \in \mathcal{D}(S^*)$. Since we already established the equivalence (i) \Leftrightarrow (ii), we would then conclude that $f \in \mathcal{D}(\overline{S})$, and hence (iii) \Rightarrow (i). Owing to the decomposition (5.2) for g ,

$$\omega(f, g) = a_0^{(g)} \omega(f, v_0) + a_\infty^{(g)} \omega(f, v_\infty) + \omega(f, b_\infty^{(g)} v_0) + \omega(f, b_0^{(g)} v_\infty).$$

One has

$$\overline{\omega(f, v_0)} = -\omega(v_0, f) = \lim_{r \downarrow 0} W_r(\overline{v_0}, f) = L_{v_0}(f) = 0,$$

having used (5.6) in the first step, (5.7) in the second, (5.9) in the third, and the assumption $L_v(f) = 0$ for all $v \in \mathcal{L}$ in the last step. Analogously,

$$\overline{\omega(f, v_\infty)} = -\omega(v_\infty, f) = \lim_{r \downarrow 0} W_r(\overline{v_\infty}, f) = L_{v_\infty}(f) = 0.$$

Therefore, $\omega(f, v_0) = \omega(f, v_\infty) = 0$, and one is left with

$$\begin{aligned} \overline{\omega(f, g)} &= \overline{\omega(f, b_\infty^{(g)} v_0)} + \overline{\omega(f, b_0^{(g)} v_\infty)} = -\omega(b_\infty^{(g)} v_0, f) - \omega(b_0^{(g)} v_\infty, f) \\ &= \lim_{r \downarrow 0} \left(W_r(\overline{b_\infty^{(g)} v_0}, f) + W_r(\overline{b_0^{(g)} v_\infty}, f) \right) \\ &= \lim_{r \downarrow 0} \left(\overline{b_\infty^{(g)}(r)} W_r(\overline{v_0}, f) + \overline{b_0^{(g)}(r)} W_r(\overline{v_\infty}, f) \right). \end{aligned}$$

As $r \downarrow 0$, $W_r(\overline{v_0}, f) \rightarrow L_{v_0}(f) = 0$ and $W_r(\overline{v_\infty}, f) \rightarrow L_{v_\infty}(f) = 0$, and also (as seen in Lemma 5.1) $b_\infty^{(g)}(r) \rightarrow 0$ and $b_0^{(g)}(r) \rightarrow 0$, whence $\omega(f, g) = 0$. This completes the proof of the implication (iii) \Rightarrow (i).

Last, in order to establish the equivalence (i) \Leftrightarrow (iv), let us decompose f as in (5.2), namely,

$$f = a_0^{(f)} v_0 + a_\infty^{(f)} v_\infty + b_\infty^{(f)} v_0 + b_0^{(f)} v_\infty,$$

and let us compute

$$\begin{aligned} L_{v_0}(f) &= \lim_{r \downarrow 0} W_r(\overline{v_0}, f) = a_0^{(f)} \lim_{r \downarrow 0} W_r(\overline{v_0}, v_0) + a_\infty^{(f)} \lim_{r \downarrow 0} W_r(\overline{v_0}, v_\infty) \\ &\quad + \lim_{r \downarrow 0} b_\infty^{(f)}(r) W_r(\overline{v_0}, v_0) + \lim_{r \downarrow 0} b_0^{(f)}(r) W_r(\overline{v_0}, v_\infty) \\ &= a_\infty^{(f)} W_0^\infty. \end{aligned}$$

Indeed, $W_r(\overline{v_0}, v_0) = W_r(v_0, v_0) = 0$, and $W_r(\overline{v_0}, v_\infty) = W_r(v_0, v_\infty) \rightarrow W_0^\infty$, $b_\infty^{(f)}(r) \rightarrow 0$, and $b_0^{(f)}(r) \rightarrow 0$ as $r \downarrow 0$. Similarly,

$$L_{v_\infty}(f) = -a_0^{(f)} W_0^\infty.$$

Because of the already proved equivalence (i) \Leftrightarrow (iii), we then conclude that $f \in \mathcal{D}(\overline{S})$ if and only if $L_{v_0}(f) = L_{v_\infty}(f) = 0$, which from the above computation is tantamount as $a_0^{(f)} = a_\infty^{(f)} = 0$. This completes the proof. \square

We thus see that Proposition 2.4 is therefore an immediate corollary of Lemma 5.1 and Proposition 5.3 above.

Proof of Proposition 2.4 The vanishing limit (2.14) for a generic $f \in \mathcal{D}(\overline{S})$ follows from the fact that, owing to Proposition 5.3(iv), $f = b_\infty^{(f)} v_0 + b_0^{(f)} v_\infty$, and from the asymptotics (5.3) of Lemma 5.1. The H^1 -regularity of f on any interval $[\varepsilon, +\infty)$, with $\varepsilon > 0$, follows from the fact that on such interval the r^{-1} potential is bounded and hence the closure of $\mathcal{D}(S)$ in the graph norm is in fact the closure of the smooth and compactly supported functions in the H^1 -norm. \square

6 Resolvents and spectral gap

In this Section we give the details of the derivation of a couple of relevant consequences from the general classification Theorem 2.1, which concern the invertibility of each member of the family of self-adjoint extensions and the expression of the resolvent. As an application to the Dirac-Coulomb Hamiltonian under consideration, we then prove Theorem 2.7.

In fact, Theorem 6.1 below is standard within the Kreĭn-Višik-Birman extension theory for semi-bounded operators (see, e.g., [16, Section 6]): we present for completeness the proof in the more general framework of self-adjoint extensions of a symmetric operator with a distinguished, invertible extension. An analogous argument, from a somewhat different perspective, can be found in [19, Theorems 13.8, 13.23, and 13.25].

Theorem 6.1 (Invertibility of extensions and resolvents) *Let S be a densely defined symmetric operator on a Hilbert space \mathcal{H} which admits a self-adjoint extension S_D that has everywhere defined and bounded inverse on \mathcal{H} . In terms of the parametrisation (2.1) of Theorem 2.1, let S_T be a generic self-adjoint extension of S and $P_T : \mathcal{H} \rightarrow \mathcal{H}$ be the orthogonal projection onto $\overline{\mathcal{D}(T)}$, where the operator T is the extension parameter.*

- (i) S_T is invertible on the whole \mathcal{H} if and only if T is invertible on the whole $\overline{\mathcal{D}(T)}$.
- (ii) When S_T is invertible, and so is T , because of (i), one has

$$S_T^{-1} = S_D^{-1} + P_T T^{-1} P_T. \quad (6.1)$$

- (iii) Assume further that $\dim \ker S^* = 1$, i.e., S has deficiency indices $(1,1)$. Let \widehat{S} be a self-adjoint extension of S other than the distinguished extension S_D . Let $\Phi \in \ker S^* \setminus \{0\}$ and for each $z \in \rho(\widehat{S}) \cap \mathbb{R}$ set

$$\Phi(z) := \Phi + z(S_D - zI)^{-1}\Phi \in \ker(S^* - z\mathbb{1}). \quad (6.2)$$

Then there exists an analytic function $\eta : \rho(\widehat{S}) \cap \mathbb{R} \rightarrow \mathbb{R}$ with $\eta(z) \neq 0$, such that

$$(\widehat{S} - zI)^{-1} = (S_D - zI)^{-1} + \eta(z)|\Phi(z)\rangle\langle\Phi(z)|. \quad (6.3)$$

$\eta(z)$, $\Phi(z)$ and (6.3) admit an analytic continuation to $\rho(S_D) \cap \rho(\widehat{S})$.

Proof (i) Let us show first that S_T is injective if and only if T is injective. Assume that S_T is injective and pick $v \in \mathcal{D}(T)$ such that $Tv = 0$. Then v is an element in $\mathcal{D}(S_T)$, because it is a vector of the form (2.1), namely $g = f + S_D^{-1}(Tv + w) + v$, with $f = w = 0$. Since $S_T v = 0$ by injectivity one concludes that $v = 0$. Conversely if T is injective and for some $g = f + S_D^{-1}(Tv + w) + v \in \mathcal{D}(S_T)$ one has $S_T g = 0$, then $\bar{S}f + Tv + w = 0$. Since $\bar{S}f + Tv + w \in \text{ran } \bar{S} \boxplus \text{ran } T \boxplus (\ker S^* \cap \mathcal{D}(T)^\perp)$, one must have $\bar{S}f = Tv = w = 0$. Owing to the injectivity of \bar{S} and T , $f = v = 0$ and hence $g = 0$. Next, let us show that S_T is surjective if and only if T is surjective. One has $\text{ran } S_T = \text{ran } \bar{S} \boxplus \text{ran } T \boxplus (\ker S^* \cap \mathcal{D}(T)^\perp)$ and in fact $\text{ran } \bar{S} = \overline{\text{ran } \bar{S}}$. Thus T is surjective

if and only if $\text{ran} T \boxplus (\ker S^* \cap \mathcal{D}(T)^\perp) = \overline{\text{ran} T} \oplus (\ker S^* \cap \mathcal{D}(T)^\perp) = \ker S^*$, if and only if $\text{ran} S_T = \overline{\text{ran} S} \oplus \ker S^* = \mathcal{H}$ if and only if S_T is surjective. The proof of (i) is thus completed.

(ii) (6.1) is an identity between bounded self-adjoint operators. For a generic $h \in \text{ran} S_T$ one has $h = S_T g$ for some $g = f + S_D^{-1}(Tv + w) + v = F + v$, where $f \in \mathcal{D}(S_{\min})$, $v \in \mathcal{D}(T)$, $w = \ker S^* \cap \mathcal{D}(T)$ (Theorem 2.1), and hence $F \in \mathcal{D}(S_D)$. Then

$$\langle h, S_T^{-1}h \rangle = \langle g, S_T g \rangle = \langle F, S_D F \rangle + \langle v, T v \rangle.$$

On the other hand

$$\langle F, S_D F \rangle = \langle S_D F, S_D^{-1} S_D F \rangle = \langle S_T g, S_D^{-1} S_T g \rangle = \langle h, S_D^{-1} h \rangle$$

and

$$\langle v, T v \rangle = \langle T v, T^{-1} T v \rangle = \langle P_T S_T g, T^{-1} P_T S_T g \rangle = \langle h, P_T T^{-1} P_T h \rangle$$

whence the conclusion $\langle h, S_T^{-1} h \rangle = \langle h, S_D^{-1} h \rangle + \langle h, P_T T^{-1} P_T h \rangle$.

(iii) Even without assuming for the moment unital deficiency indices, for $z \in \rho(\widehat{S}) \cap \rho(S_D)$ let $T(z)$ be the extension parameter, in the sense of KVB parametrisation (2.1) of Theorem 2.1, of the operator $\widehat{S} - z\mathbb{1}$ considered as a self-adjoint extension of the densely defined operator $S(z) = S - z\mathbb{1}$. Correspondingly, let $P(z)$ be the orthogonal projection onto $\mathcal{D}(T(z))$. Then

$$(\widehat{S} - z\mathbb{1})^{-1} = (S_D - z\mathbb{1})^{-1} + P(z) T(z)^{-1} P(z), \quad (*)$$

which follows from part (ii), due to the fact that the distinguished extension of $S - z\mathbb{1}$ is $S_D - z\mathbb{1}$. Assuming now $\dim \ker S^* = 1$, one has $\dim \ker (S^* - z\mathbb{1}) = 1$, because of the constancy of the deficiency indices. Moreover, $\widehat{S} - z\mathbb{1}$ is a self-adjoint extension of $S - z\mathbb{1}$, whose extension parameter $T(z)$, in the sense of KVB parametrisation of Theorem 2.1, acts as the multiplication by a real number $t(z)$ on the one-dimensional space $\ker(S^* - z\mathbb{1})$. The fact that $(S^* - z\mathbb{1})\Phi(z) = 0$ is obvious by construction. Moreover $\Phi(z) \neq 0$ for each admissible z : this is obviously true if $z = 0$, and if it was not true for $z \neq 0$, then $z(S_D - z\mathbb{1})^{-1}\Phi = -\Phi \neq 0$, which would contradict $\mathcal{D}(S_D - z\mathbb{1}) \cap \ker(S^* - z\mathbb{1}) = \{0\}$. Thus, $\Phi(z)$ spans $\ker(S^* - z\mathbb{1})$ and $P_T := \|\Phi(z)\|^{-2} |\Phi(z)\rangle \langle \Phi(z)| : \mathcal{H} \rightarrow \mathcal{H}$ is the orthogonal projection onto $\ker(S^* - z\mathbb{1})$. In this case, the resolvent formula (*) above takes precisely the form (6.3) where $\eta(z) := \|\Phi(z)\|^{-2} t(z)^{-1}$. Being a product of non-zero quantities, $\eta(z) \neq 0$. Moreover, $z \mapsto (\widehat{S} - z\mathbb{1})^{-1}$ and $z \mapsto (S_D - z\mathbb{1})^{-1}$ are analytic operator-valued functions on the whole $\rho(S_D) \cap \rho(\widehat{S})$ (because of the analyticity of resolvents) and so is the vector-valued function $z \mapsto \Phi(z)$ (because of the construction (6.2)). Therefore, taking the expectation of both sides of (*) on $\Phi(z)$ shows at once that $z \mapsto \eta(z)$ is analytic on $\rho(S_D) \cap \rho(\widehat{S})$, and real analytic on $\mathbb{R} \cap \rho(\widehat{S})$. \square

Proof of Theorem 2.7 Part (i) is an immediate consequence of Theorem 6.1(i), since the KVB-extension parameter in the present case is the multiplication by β . This is of course consistent with the representation formula (2.15), which clearly implies that

when $\beta = 0$ the extension $S_{\beta=0}$ has a kernel. Analogously, part (ii) is an immediate consequence of Theorem 6.1(ii), because the orthogonal projection P_T has in the present case the expression $P_T = \|\Phi\|^{-2} |\Phi\rangle\langle\Phi|$. Concerning part (iii), (2.21) is a consequence of the fact that, as stated in (2.20), the resolvent difference between the β -extension and the distinguished extension is compact. Moreover, using (2.20) we re-write $S_\beta f = Ef$ as

$$f = E S_\beta^{-1} f = E \left(S_D^{-1} + \frac{1}{\beta \|\Phi\|^2} |\Phi\rangle\langle\Phi| \right) f.$$

This equation is surely solved by $f = 0$ and, if $E \in (-E(\beta), E(\beta))$, then the operator acting on the r.h.s. is a contraction. Thus $f = 0$ is the only function which satisfies the eigenvalue equation $S_\beta f = Ef$ and therefore there cannot be eigenvalues in such a regime of E . \square

7 Concluding remarks

We would like to end our analysis with some observations on our overall approach also in comparison with the previous literature.

As documented already, the literature concerning the problem of realising self-adjointly the Dirac-Coulomb operator is vast and unfolds uninterrupted over many decades until recent times, across different disciplines such as ODEs, functional inequalities, operator theory, etc. In fact, the Dirac-Coulomb Hamiltonian is known since long not to be uniquely realised for large Coulomb couplings, with a dominant part of the literature devoted to the study of the properties of the distinguished extension in the critical regime. The perspective of the general classification of the extensions is relatively recent [21, 35], and what we found that was missing was a comprehension of the structure of the family of extensions through the Kreĭn-Višik-Birman and Grubb scheme, as opposite to the standard von Neumann scheme.

In the former framework we could establish Theorems 2.5, 2.6, and 2.7 in a form and through steps that, to our taste, in comparison with [35, Sections 3 and 4] and [21, Sections 3-7], let emerge more straightforwardly the overall extension picture and the meaning of the self-adjointness boundary condition as a multiplicative constraint between regular and singular part of the functions in the domain of the extension, the multiplicative constant giving also immediate information on the invertibility property and on the resolvent and spectral gap of the extension.

This is evident comparing our form (2.18) of the (asymptotic) boundary condition of self-adjointness with [21, Theorem 7.1, Eq. (56)], purely based on von Neumann's extension theory.

It is also worth pointing out that [35, Eq. (67)] expresses the vanishing rate of elements of what is here the space $\mathcal{D}(\bar{S})$ only as $O(r^{1/2})$, as $r \downarrow 0$, whereas we proved that the correct vanishing rate is $o(r^{1/2})$ (Proposition 2.4). We record that the $o(r^{1/2})$ -rate was mentioned, but not substantiated, already in [9, Section 2].

As far as our use of techniques from ODE theory is concerned, most of what we did is somewhat standard, but it has to be highlighted that our analysis of the space $\mathcal{D}(\bar{S})$

in Sect. 5 is very much inspired to that of the recent work [8] on the ‘twin’ problem of the (scalar) homogeneous Schrödinger operator $h = -\frac{d^2}{dr^2} + \nu r^{-2}$ on half-line. Despite the difference of goals with [8], where the Kreĭn-Višik-Birman scheme is not exploited, the resemblance of results is not surprising: in [8, Proposition 4.17] the family of self-adjoint realisations of h is qualified to be a collection $(h_\theta)_{\theta \in [0, 2\pi)}$ where $\mathcal{D}(h_\theta)$ is formed by elements that as $r \downarrow 0$ have the form

$$f + c \left(r^{\frac{1}{2}-m} \cos \theta + r^{\frac{1}{2}+m} \sin \theta \right)$$

for some $c \in \mathbb{C}$ and some function f with $f(r) \sim r^{-3/2}$ as $r \downarrow 0$, where $m := \sqrt{\nu + \frac{1}{4}}$. In fact, one would say in the present language that also in that case it is possible to identify a distinguished extension in the regime $\nu > -\frac{1}{4}$, the one with $\theta = \frac{\pi}{2}$, which has the property that $\mathcal{D}(h_{\theta=\pi/2}) \subset \mathcal{D}[r^{-2}]$.

Acknowledgements We are indebted to Naiara Arrizabalaga, Gianfausto Dell’Antonio, Marko Erceg, Diego Noja, and Giulio Ruzza for many instructive and inspirational discussions on the subject of this work.

Compliance with ethical standards

Conflict of interest All authors declare that they have no conflict of interest.

References

1. Abramowitz, M., Stegun, I.A.: Handbook of mathematical functions with formulas, graphs, and mathematical tables, vol. 55 of National Bureau of Standards Applied Mathematics Series, For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, DC (1964)
2. Alonso, A., Simon, B.: The Birman-Kreĭn-Vishik theory of selfadjoint extensions of semibounded operators. *J. Operator Theory* **4**, 251–270 (1980)
3. Arai, M.: On essential selfadjointness, distinguished selfadjoint extension and essential spectrum of Dirac operators with matrix valued potentials. *Publ. Res. Inst. Math. Sci.* **19**, 33–57 (1983)
4. Arai, M., Yamada, O.: Essential selfadjointness and invariance of the essential spectrum for Dirac operators. *Publ. Res. Inst. Math. Sci.* **18**, 973–985 (1982)
5. Arrizabalaga, N.: Distinguished self-adjoint extensions of Dirac operators via Hardy-Dirac inequalities. *J. Math. Phys.* **52**, 092301 (2011). 14
6. Arrizabalaga, N., Duoandikoetxea, J., Vega, L.: Self-adjoint extensions of Dirac operators with Coulomb type singularity. *J. Math. Phys.* **54**, 041504 (2013). 20
7. Berthier, A., Georgescu, V.: On the point spectrum of Dirac operators. *J. Funct. Anal.* **71**, 309–338 (1987)
8. Bruneau, L., Dereziński, J., Georgescu, V.: Homogeneous Schrödinger operators on half-line. *Ann. Henri Poincaré* **12**, 547–590 (2011)
9. Burnap, C., Brysk, H., Zweifel, P.F.: Dirac Hamiltonian for strong Coulomb fields. *Il Nuovo Cimento B (1971-1996)* **64**, 407–419 (1981)
10. Chernoff, P.R.: Schrödinger and Dirac operators with singular potentials and hyperbolic equations. *Pac. J. Math.* **72**, 361–382 (1977)
11. Esteban, M.J., Lewin, M., Séré, E.: Domains for Dirac-Coulomb min–max levels (2017). [arXiv:1702.04976](https://arxiv.org/abs/1702.04976)
12. Esteban, M.J., Loss, M.: Self-adjointness for Dirac operators via Hardy-Dirac inequalities. *J. Math. Phys.* **48**, 112107 (2007). 8
13. Evans, W.D.: On the unique self-adjoint extension of the Dirac operator and the existence of the Green matrix. *Proc. Lond. Math. Soc. (3)* **20**, 537–557 (1970)

14. Fall, M.M., Felli, V.: Sharp essential self-adjointness of relativistic Schrödinger operators with a singular potential. *J. Funct. Anal.* **267**, 1851–1877 (2014)
15. Gallone, M.: Self-adjoint extensions of Dirac operator with Coulomb potential. In: Dell’Antonio, G., Michelangeli, A. (eds.) *Advances in Quantum Mechanics*, INdAM-Springer series, vol. 18, pp. 169–185. Springer, Berlin (2017)
16. Gallone, M., Michelangeli, A., Ottolini, A.: Kreĭn-Višik-Birman self-adjoint extension theory revisited, SISSA preprint 25/2017/MATE (2017)
17. Georgescu, V., Măntoiu, M.: On the spectral theory of singular Dirac type Hamiltonians. *J. Operator Theory* **46**, 289–321 (2001)
18. Grubb, G.: A characterization of the non-local boundary value problems associated with an elliptic operator. *Ann. Scuola Norm. Sup. Pisa* (3) **22**, 425–513 (1968)
19. Grubb, G.: *Distributions and Operators*. Graduate Texts in Mathematics, vol. 252. Springer, New York (2009)
20. Gustafson, K.E., Rejto, P.A.: Some essentially self-adjoint Dirac operators with spherically symmetric potentials. *Israel J. Math.* **14**, 63–75 (1973)
21. Høegreve, H.: The overcritical Dirac–Coulomb operator. *J. Phys. A* **46**, 025301 (2013). 22
22. Kalf, H., Schmincke, U.-W., Walter, J., Wüst, R.: *On the Spectral Theory of Schrödinger and Dirac Operators with Strongly Singular Potentials*. Lecture Notes in Mathematics, vol. 448, pp. 182–226. Springer, Berlin (1975)
23. Kato, T.: Holomorphic families of Dirac operators. *Math. Z.* **183**, 399–406 (1983)
24. Klaus, M., Wüst, R.: Characterization and uniqueness of distinguished selfadjoint extensions of Dirac operators. *Comm. Math. Phys.* **64**, 171–176 (1978/79)
25. Landgren, J.J., Rejto, P.A.: An application of the maximum principle to the study of essential selfadjointness of Dirac operators. I. *J. Math. Phys.* **20**, 2204–2211 (1979)
26. Landgren, J.J., Rejto, P.A.: On a theorem of Jörgens and Chernoff concerning essential selfadjointness of Dirac operators. *J. Reine Angew. Math.* **322**, 1–14 (1981)
27. Landgren, J.J., Rejto, P.A., Klaus, M.: An application of the maximum principle to the study of essential selfadjointness of Dirac operators II. *J. Math. Phys.* **21**, 1210–1217 (1980)
28. Le Yaouanc, A., Oliver, L., Raynal, J.-C.: The Hamiltonian $(p^2 + m^2)^{1/2} - \alpha/r$ near the critical value $\alpha_c = 2/\pi$. *J. Math. Phys.* **38**, 3997–4012 (1997)
29. Nenciu, G.: Self-adjointness and invariance of the essential spectrum for Dirac operators defined as quadratic forms. *Commun. Math. Phys.* **48**, 235–247 (1976)
30. Reed, M., Simon, B.: *Methods of Modern Mathematical Physics. II. Fourier Analysis, Self-Adjointness*. Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London (1975)
31. Rejto, P.A.: Some essentially self-adjoint one-electron Dirac operators. (With appendix.). *Israel J. Math.* **9**, 144–171 (1971)
32. Schmincke, U.-W.: Distinguished selfadjoint extensions of Dirac operators. *Math. Z.* **129**, 335–349 (1972)
33. Schmincke, U.-W.: Essential selfadjointness of Dirac operators with a strongly singular potential. *Math. Z.* **126**, 71–81 (1972)
34. Thaller, B.: *The Dirac Equation*, Texts and Monographs in Physics. Springer, Berlin (1992)
35. Voronov, B.L., Gitman, D.M., Tyutin, I.V.: The Dirac Hamiltonian with a superstrong Coulomb field. *Teoret. Mat. Fiz.* **150**, 41–84 (2007)
36. Wasow, W.: *Asymptotic Expansions for Ordinary Differential Equations*. Dover Publications Inc, New York (1987). (Reprint of the 1976 edition)
37. Weidmann, J.: *Oszillationsmethoden für Systeme gewöhnlicher Differentialgleichungen*. *Math. Z.* **119**, 349–373 (1971)
38. Weidmann, J.: *Spectral Theory of Ordinary Differential Operators*. Lecture Notes in Mathematics. Springer, Berlin (1987)
39. Wüst, R.: Distinguished self-adjoint extensions of Dirac operators constructed by means of cut-off potentials. *Math. Z.* **141**, 93–98 (1975)
40. 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**, 259–271 (1977)
41. Xia, J.: On the contribution of the Coulomb singularity of arbitrary charge to the Dirac Hamiltonian. *Trans. Am. Math. Soc.* **351**, 1989–2023 (1999)