

Bifurcating standing waves for effective equations in gapped honeycomb structures

W. Borrelli¹, R. Carlone²

¹Centro De Giorgi, Scuola Normale Superiore, Piazza dei Cavalieri 3, I-56100, Pisa, Italy

²Università “Federico II” di Napoli, Dipartimento di Matematica e Applicazioni “R. Caccioppoli”,
MSA, via Cinthia, I-80126, Napoli, Italy

william.borrelli@sns.it, raffaele.carlone@unina.it

PACS 03.65.-w, 02.30.Rz

DOI 10.17586/2220-8054-2021-12-1-5-14

In this paper, we deal with two-dimensional cubic Dirac equations, appearing as an effective model in gapped honeycomb structures. We give a formal derivation starting from cubic Schrödinger equations and prove the existence of standing waves bifurcating from one band-edge of the linear spectrum.

Keywords: nonlinear Dirac equations, bifurcation methods, existence results, honeycomb structures.

Received: 28 December 2020

Revised: 6 January 2021

1. Introduction

1.1. Motivation and main results

In this paper, we deal with nonlinear massive Dirac equations of the form:

$$(\mathcal{D} + m\sigma_3 - \omega)\psi = h(\psi)\psi \quad \text{on } \mathbb{R}^2, \quad (1)$$

where $\omega \in (-m, m)$ is a frequency in the spectral gap of the Dirac operator $\mathcal{D} + m\sigma_3$, with $m > 0$ (see Section 2). We consider the nonlinearity in (1) of the form:

$$h(z) = \begin{pmatrix} \beta_1|z_1|^2 + 2\beta_2|z_2|^2 & 0 \\ 0 & \beta_1|z_2|^2 + 2\beta_2|z_1|^2 \end{pmatrix}, \quad z \in \mathbb{R}^2, \quad (2)$$

with given parameters $\beta_1, \beta_2 > 0$.

Equation (1) appears as an effective model of wave propagation in two-dimensional honeycomb structures. As proved in [1], if $V \in C^\infty(\mathbb{R}^2, \mathbb{R})$ is a potential having the symmetries of a honeycomb lattice, then the Schrödinger operator:

$$H = -\Delta + V(x), \quad x \in \mathbb{R}^2, \quad (3)$$

exhibits generically conical touching points in its dispersion bands called *Dirac points*. The dynamics of wave packets spectrally concentrated around Dirac points, see [2], is thus effectively described by the *massless* (i.e., $m = 0$) Dirac operator. Adding a perturbation that breaks parity induces a mass term in the effective operator, as proved in [1, Appendix].

An important model in nonlinear optics and in the description of macroscopic phenomena is given by the *nonlinear Schrödinger / Gross-Pitaevski equation* [3–5]:

$$i\partial_t u = Hu + |u|^2 u. \quad (4)$$

This equation, in the approximation described before, leads (at least formally) to the effective cubic nonlinear-ity (2). Indeed, as first computed in [6], the effective equation around Dirac points reads:

$$\begin{cases} \partial_t \Xi_1 + \bar{\lambda}(\partial_{x_1} + i\partial_{x_2})\Xi_2 = i(2\beta_2|\Xi_1|^2 + \beta_1|\Xi_2|^2)\Xi_1, \\ \partial_t \Xi_2 + \lambda(\partial_{x_1} - i\partial_{x_2})\Xi_1 = i(\beta_1|\Xi_1|^2 + 2\beta_2|\Xi_2|^2)\Xi_2, \end{cases} \quad (5)$$

where the parameters $\lambda \in \mathbb{C} \setminus \{0\}$, $\beta_1, \beta_2 > 0$ depend on the potential V in (3).

Setting $\Psi_1 := -\frac{\lambda}{|\lambda|}\Xi_2$, $\Psi_2 := \Xi_1$ and looking for stationary solutions

$$\Psi(t, x) = \psi(x),$$

we get the massless version of (1), i.e. with $m = \omega = 0$. As shown in Section 3, adding a perturbation breaking the parity of the potential V in (3) gives an additional mass term in the effective equation. This corresponds to a gap $(-m, m)$ in the linear spectrum so that we can consider stationary solutions at frequency $\omega \in (-m, m)$, leading to (1).

In [7] the validity of the effective cubic equation is studied. In Section 3 we give a formal derivation of the effective model (1) using a multiscale expansion.

Existence and qualitative properties of solutions to the massless version of (1) have been studied in [8, 9]. The massive case in (1) has been addressed in [10, 11] for a special choice of parameters in (1).

In this paper we partly generalize those results dealing with arbitrary $\beta_1, \beta_2 > 0$ and proving the existence of stationary solutions bifurcating from one edge of the spectral gap of the operator $\mathcal{D} + m\sigma_3$.

We remark that cubic Dirac equations in two dimensions are *critical* for the Sobolev embedding. Such types of equations have been studied also in different contexts. We mention, for instance, problems from conformal spin geometry, for which we refer the reader to [12–15] and references therein, and in the case of coupled systems involving the Dirac operator and critical nonlinearities related to supersymmetric models coupling gravity with fermions, see [16, 17]. The main difficulty in studying those equations comes from the underlying conformal symmetry so that looking for stationary solutions by variational methods one has to deal with the induced loss of compactness, see [10, 11]. This problem can be circumvented, for instance, using a bifurcation argument to find solutions to (1), as done in this paper following [18]. We mention that the same method has been recently used for nonlinear Dirac equations on star graphs [19].

The results given in [10, 11] correspond to the choice of parameters $\beta_1 = 2\beta_2$, so that one can assume $\beta_1 = 1$, $\beta_2 = 1/2$ by scaling. In this paper, we deal with general $\beta_1, \beta_2 > 0$, but this forces us to put restrictions on the frequency ω that will be close to the band-edge at m . More precisely, we focus on the existence of standing waves to (1) of symmetric form:

$$\psi(r, \theta) = \begin{pmatrix} v(r) \\ iu(r)e^{i\theta} \end{pmatrix}, \quad (r, \theta) \in (0, \infty) \times \mathbb{S}^1, \quad (6)$$

(r, θ) being polar coordinates in \mathbb{R}^2 , and u, v real-valued functions. Notice that (6) is the two-dimensional analogue of the *Soler/Wakano ansatz* [20–22].

Theorem 1.1. *Let $\varepsilon := m - \omega$. There exists $\varepsilon_0 > 0$ such that for $\varepsilon \in (0, \varepsilon_0)$ equation (1) admits a solution ψ_ε of the form (6), with*

$$u_\varepsilon(r) = \varepsilon(-f'(\sqrt{\varepsilon}r) + e_1(\sqrt{\varepsilon}r)), \quad v_\varepsilon(r) = \sqrt{\varepsilon}(f(\sqrt{\varepsilon}r) + e_2(\sqrt{\varepsilon}r)), \quad r > 0,$$

where $\|e_j\|_{H^1(\mathbb{R}^2)} \leq C\varepsilon$, $j = 1, 2$, and $f \in H^1(\mathbb{R}^2)$ is the positive ground state of the NLS

$$-\Delta f - f^3 + f = 0, \quad \text{on } \mathbb{R}^2.$$

Remark 1.1. Arguing as in Section 4 one can deal with the regime $-m < \omega < 0$, $\omega \rightarrow -m$. However in that case the limit equation (57) is replaced by the following *defocusing* NLS

$$-\Delta U + U^3 + U = 0, \quad \text{on } \mathbb{R}^2, \quad (7)$$

which has no non-trivial solution in $H^1(\mathbb{R}^2)$. This can be easily seen multiplying the equation by such a solution and integrating by parts.

2. The Dirac operator

The *Dirac operator* is the constant coefficients first order differential operator defined in two dimensions as:

$$\mathcal{D}_m = \mathcal{D} + m\sigma_3 := -i\sigma \cdot \nabla + m\sigma_3. \quad (8)$$

The constant $m > 0$ usually represents the mass of the particle described by the equation. We adopt the notation $\sigma \cdot \nabla := \sigma_1\partial_1 + \sigma_2\partial_2$ and the σ_k 's are the Pauli matrices:

$$\sigma_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (9)$$

The operator \mathcal{D}_m is a self-adjoint operator on $L^2(\mathbb{R}^2, \mathbb{C}^2)$, with domain $H^1(\mathbb{R}^2, \mathbb{C}^2)$ and form-domain $H^{1/2}(\mathbb{R}^2, \mathbb{C}^2)$.

Passing to the Fourier domain $p = (p_1, p_2)$ the Dirac operator (8) becomes the multiplication by the matrix:

$$\widehat{\mathcal{D}}_m(p) = \begin{pmatrix} m & p_1 - ip_2 \\ p_1 + ip_2 & m \end{pmatrix}$$

and then the spectrum is easily found to be:

$$\text{Spec}(\mathcal{D}_m) = (-\infty, -m] \cup [m, +\infty). \quad (10)$$

The above mentioned results can be found, e.g., in [23].

3. Formal derivation of the model

In this section, we give a formal derivation of equation (1) from the corresponding cubic Schrödinger equation with honeycomb potential following the exposition given in [24].

We consider a fixed triangular lattice $\Lambda := \mathbb{Z}v_1 \oplus \mathbb{Z}v_2$, where $v_1, v_2 \in \mathbb{R}^2$ are two linearly independent vectors.

3.1. Honeycomb Schrödinger operators

Consider the Schrödinger operator

$$H := -\Delta + V(x), \quad x \in \mathbb{R}^2. \quad (11)$$

Definition 3.1. *The function $V \in C^\infty(\mathbb{R}^2)$ is called honeycomb potential, see [1], if there exists $x_0 \in \mathbb{R}^2$ such that $\tilde{V}(x) = V(x - x_0)$ has the following properties:*

- (1) \tilde{V} is periodic with respect to some triangular lattice Λ , that is, $\tilde{V}(x + v) = \tilde{V}(x)$, $\forall x \in \mathbb{R}^2, \forall v \in \Lambda$;
- (2) \tilde{V} is even: $\tilde{V}(-x) = \tilde{V}(x)$, $\forall x \in \mathbb{R}^2$;
- (3) \tilde{V} is invariant by $\frac{2\pi}{3}$ counterclockwise rotation:

$$\mathcal{R}[\tilde{V}](x) := \tilde{V}(R^*x) = \tilde{V}(x) \quad \forall x \in \mathbb{R}^2,$$

where R is the corresponding rotation matrix:

$$R = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}. \quad (12)$$

Remark 3.1. (Some examples of honeycomb potentials [1])

- (1) **Atomic potentials:** Let $\mathbb{H} = (A + \Lambda) \cup (B + \Lambda)$ be a hexagonal lattice, given by the superposition of two triangular lattices. Consider a radial function $V_0 \in C^\infty(\mathbb{R}^2)$ rapidly decaying at infinity (for instance, with polynomial rate) representing the potential generated by a nucleus located on a vertex of the lattice. The potential

$$V(x) = \sum_{y \in \mathbb{H}} V_0(x - y)$$

is then given by the superposition of atomic potentials, and it is a honeycomb potential (Def. 3.1).

- (2) **Optical lattices:** The envelop ψ of the electric field of a monochromatic beam propagating in a dielectric medium can be described by a Schrödinger equation. More precisely, denoting by z the direction of propagation of the beam and assuming that the refraction index varies only in the transversal directions (x, y) , the function ψ solve the following equation:

$$i\partial_z \psi = (-\Delta + V(x, y)) \psi. \quad (13)$$

In this case, the honeycomb potential is generated using optical interference techniques [25]. A typical example is the potential of the form:

$$V(x, y) \simeq V_0(\cos(k_1 \cdot (x, y)) + \cos(k_1 \cdot (x, y)) + \cos((k_1 + k_2) \cdot x)), \quad V_0 \in \mathbb{R}, k_1, k_2 \in \mathbb{R}^2. \quad (14)$$

For any fixed $k \in \mathbb{R}^2$ consider the following eigenvalue problem with *pseudo-periodic boundary conditions* (see [1] and [26, Sec. XIII.16]):

$$\begin{cases} H\Phi(x; k) = \mu(k)\Phi(x; k), & x \in \mathbb{R}^2 \\ \Phi(x + v; k) = e^{ik \cdot v} \Phi(x; k), & v \in \Lambda. \end{cases} \quad (15)$$

Remark 3.2. The eigenfunctions $\Phi(x; k)$ are of class C^∞ by elliptic regularity theory.

Recall that, given a lattice, its (*first*) Brillouin zone \mathcal{B} is defined as the fundamental cell of the dual lattice. In the case of a honeycomb lattice, both its fundamental cell and its Brillouin zone are hexagonal [1]. An important property of \mathcal{B} is that waves propagating in a periodic medium can be described in terms of *Bloch functions*.

Given $k \in \mathcal{B}$, the resolvent of $H(k)$ is compact and then the spectrum of the operator is real and purely discrete, accumulating at $+\infty$:

$$\mu_1(k) \leq \mu_2(k) \leq \dots \leq \mu_j(k) \leq \dots \uparrow +\infty. \quad (16)$$

Fixing $n \in \mathbb{N}$, one says that $k \mapsto \mu_n(k)$ is the n -th *dispersion band* of the operator H and call n -th *Bloch wave* the function $\Phi_n(x, k)$. The spectrum may also have some gaps, and it can be obtained as union of the images of the dispersion bands of the operator

$$\text{Spec}(H) = \bigcup_{n \in \mathbb{N}} \mu_n(\mathcal{B}). \quad (17)$$

Moreover, the Bloch waves constitute a complete systems, meaning that for all $f \in L^2(\mathbb{R}^2)$:

$$f(x) - \sum_{1 \leq n \leq N} \int_{\mathcal{B}} \langle \Phi_n(\cdot, k), f(\cdot) \rangle_{L^2(\mathbb{R}^2)} \Phi_n(x; k) dk \longrightarrow 0 \quad (18)$$

in $L^2(\mathbb{R}^2)$, for $N \longrightarrow +\infty$ [1, 26].

The Cauchy problem:

$$\begin{cases} i\partial_t u(t, x) = Hu(t, x), & (t, x) \in \mathbb{R} \times \mathbb{R}^2, \\ u(0, x) = u_0(x) \in L^2(\mathbb{R}^2), \end{cases} \quad (19)$$

admits the solution:

$$e^{-iHv t} u_0 = \sum_{n \in \mathbb{N}} \int_{\mathcal{B}} e^{-i\mu_n(k)} \langle \Phi_n(\cdot, k), u_0(\cdot) \rangle_{L^2(\mathbb{R}^2)} \Phi_n(x, k) dk. \quad (20)$$

As a consequence, it is evident that the dynamics (20) are strongly influenced by the behavior of the band functions $\mu_n(\cdot)$, $n \in \mathbb{N}$. In particular, as showed in [2], there exist two bands μ_N, μ_{N+1} that meet at conical points located at the vertices of \mathcal{B} . That is, locally near such a point $K_* \in \mathcal{B}$ there holds:

$$\begin{cases} \mu_{N+1}(k) - \mu_{N+1}(K_*) = |\lambda| |k - K_*| (1 + E_+(k - K_*)), \\ \mu_N(k) - \mu_N(K_*) = -|\lambda| |k - K_*| (1 + E_-(k - K_*)), \end{cases} \quad |k - K_*| < \delta, \lambda \in \mathbb{C}, \lambda \neq 0. \quad (21)$$

Here, $E_{\pm} : U_{\delta} \rightarrow \mathbb{R}$, with $U_{\delta} := \{y \in \mathbb{R}^2 : |y| < \delta\}$, are Lipschitz functions such that $E_{\pm}(y) = O(|y|)$, for $|y| \rightarrow 0$. This means that, to first order, the dispersion relation near $k = K_*$ is a cone. This corresponds to the dispersion relation of the two-dimensional Dirac operator (8), as it can be readily seen in the Fourier domain.

Consider a wave packet $u_0(x) = u_0^{\varepsilon}(x)$ concentrated around a Dirac point K_*

$$u_0^{\varepsilon}(x) = \sqrt{\varepsilon} (\psi_{0,1}(\varepsilon x) \Phi_1(x) + \psi_{0,2}(\varepsilon x) \Phi_2(x)), \quad (22)$$

where Φ_j , $j = 1, 2$, are the Bloch functions at K_* and the functions $\psi_{0,j}$ are some (complex) amplitudes to be determined. Then, the solution of the NLS (57), with initial conditions u_0^{ε} is expected to evolve to leading order in ε still as a modulation of Bloch functions,

$$u^{\varepsilon}(t, x) \underset{\varepsilon \rightarrow 0^+}{\sim} \sqrt{\varepsilon} (\psi_1(\varepsilon t, \varepsilon x) \Phi_1(x) + \psi_2(\varepsilon t, \varepsilon x) \Phi_2(x) + \mathcal{O}(\varepsilon)), \quad t > 0, x \in \mathbb{R}^2, \quad (23)$$

and the amplitudes ψ_j solve the effective equation (5).

Given a Dirac point $K_* \in \mathcal{B}$, let $\mu_* := \mu_N(K_*) = \mu_{N+1}(K_*)$ be the frequency at which the conical crossing occurs. Consider then the NLS:

$$(-\Delta + V - \mu_*)u = |u|^2 u, \quad \mathbb{R}^2. \quad (24)$$

As in (23), one thus looks for solutions to (24) of the form:

$$u^{\varepsilon}(t, x) \underset{\varepsilon \rightarrow 0^+}{\sim} \sqrt{\varepsilon} e^{-t\mu_*} (\psi_1(\varepsilon x) \Phi_1(x) + \psi_2(\varepsilon x) \Phi_2(x) + \mathcal{O}(\varepsilon)), \quad t > 0, x \in \mathbb{R}^2. \quad (25)$$

3.2. Derivation of the massless equation

The aim of this subsection is to formally derive the effective Dirac equation for the amplitudes ψ_j appearing in (25) through a multiscale expansion (see e.g. [7, 27]).

Since the coefficients $\psi_j(\varepsilon x)$ and the Bloch functions $\Phi_j(x)$ vary on different scales, one can consider x and $y := \varepsilon x$, $0 < \varepsilon \ll 1$, as independent variables. Moreover, we look for solution to (24) as formal power series in ε , as follows:

$$u_\varepsilon = \sqrt{\varepsilon} U_\varepsilon(x, y), \quad U_\varepsilon(x, y) = U_0(x, y) + \varepsilon U_1(x, y) + \varepsilon^2 U_2(x, y) + \dots \quad (26)$$

We moreover impose K_* -pseudoperiodicity with respect to x , i.e.

$$U_\varepsilon(x + v, y) = e^{-iK_* \cdot v} U_\varepsilon(x, y), \quad \forall v \in \Lambda, x, y \in \mathbb{R}^2. \quad (27)$$

Similarly, we look for μ of the form:

$$\mu = \mu_\varepsilon = \mu_* + \varepsilon \mu_1 + \varepsilon^2 \mu_2 + \dots \quad (28)$$

Rewriting (24) in terms of U_ε and μ_ε then gives:

$$\left(-(\nabla_x + \varepsilon \nabla_y)^2 + V(x) - \mu_\varepsilon \right) U_\varepsilon(x, y) = \varepsilon |U_\varepsilon(x, y)|^2 U_\varepsilon(x, y). \quad (29)$$

Plugging (26,28) into (29) one finds a hierarchy of equations.

At order $\mathcal{O}(\varepsilon^0)$ we obtain:

$$(-\Delta_x + V - \mu_*) U_0 = 0. \quad (30)$$

Recall that $\ker_{L^2_{K_*}}(-\Delta + V - \mu_*) = \text{Span}\{\Phi_1, \Phi_2\}$, and then by (27) we have:

$$U_0(x, y) = \psi_1(y) \Phi_1(x) + \psi_2(y) \Phi_2(x), \quad (31)$$

where the amplitudes are to be determined solving the next equation in the formal expansion. Here $L^2_{K_*}$ denotes square integrable functions satisfying the pseudo-periodicity condition in (15).

The equation for $\mathcal{O}(\varepsilon)$ terms reads:

$$(-\Delta_x + V - \mu_*) U_1 = (2\nabla_x \cdot \nabla_y + \mu_1) U_0 + |U_0|^2 U_0. \quad (32)$$

By Fredholm alternative, solvability of the above equation requires its right hand side to be L^2 -orthogonal to the kernel of $(-\Delta_x + V - \mu_*)$. Then the functions ψ_j are determined imposing orthogonality to the Bloch functions Φ_k . For simplicity we deal with linear part and the cubic term in the right hand side of (32) separately.

The linear terms can be calculated using the following lemma from [1]:

Lemma 3.1. *Let $\zeta = (\zeta_1, \zeta_2) \in \mathbb{C}^2$ be a vector. Then there exists $\lambda \in \mathbb{C} \setminus \{0\}$ such that we have:*

$$\begin{aligned} \langle \Phi_k, \zeta \cdot \nabla \Phi_k \rangle_{L^2(\Omega)} &= 0, \quad k = 1, 2, \\ 2i \langle \Phi_1, \zeta \cdot \nabla \Phi_2 \rangle_{L^2(\Omega)} &= \overline{2i \langle \Phi_2, \zeta \cdot \nabla \Phi_1 \rangle_{L^2(\Omega)}} = -\bar{\lambda} (\zeta_1 + i\zeta_2), \\ 2i \langle \Phi_2, \zeta \cdot \nabla \Phi_1 \rangle_{L^2(\Omega)} &= -\lambda (\zeta_1 - i\zeta_2). \end{aligned} \quad (33)$$

Notice that: $(\nabla_x \cdot \nabla_y) U_0 = \sum_{j=1}^2 \nabla_y \psi_j \cdot \nabla_x \Phi_j$ and then applying Lemma 3.1 with $\zeta = \nabla_y \Phi_j$, $j = 1, 2$ we get:

$$\begin{aligned} 2i \langle \Phi_1, \nabla_y \psi_2 \cdot \nabla \Phi_2 \rangle_{L^2(\Omega)} &= \overline{2i \langle \Phi_2, \nabla_y \psi_2 \cdot \nabla \Phi_1 \rangle_{L^2(\Omega)}} = -\bar{\lambda} (\partial_{y_1} + i\partial_{y_2}) \psi_2, \\ 2i \langle \Phi_2, \nabla_y \psi_1 \cdot \nabla \Phi_1 \rangle_{L^2(\Omega)} &= -\lambda (\partial_{y_1} - i\partial_{y_2}) \psi_1. \end{aligned} \quad (34)$$

Thus, we see that taking the $L^2(\Omega)$ scalar product of the linear part in the right hand side of (32) with the Bloch functions Φ_j gives the linear part of (5). We now want to show that the cubic nonlinearity in (5) is obtained calculating the same product for the cubic term in (32). By symmetry taking this projection many terms vanish. The cubic term reads:

$$|U_0|^2 U_0 = \sum_{1 \leq j, k, l \leq 2} \psi_j \psi_k \bar{\psi}_k \Phi_j \Phi_k \bar{\Phi}_l. \quad (35)$$

Let us consider, for instance, the term: $\psi_1 \psi_1 \overline{\psi_2} \Phi_1 \Phi_1 \overline{\Phi_2}$ and then project it onto Φ_1 . We compute

$$\begin{aligned}
\langle \Phi_1, \Phi_1 \Phi_1 \overline{\Phi_2} \rangle_{L^2(\Omega)} &= \int_{\Omega} \overline{\Phi_1(x)} \Phi_1(x) \Phi_1(x) \overline{\Phi_2(x)} dx \\
&= \int_{R\Omega} \overline{\Phi_1(R^*y)} \Phi_1(R^*y) \Phi_1(R^*y) \overline{\Phi_2(R^*y)} dy \\
&= \int_{R\Omega} \tau \overline{\Phi_1(y)} \tau \Phi_1(y) \tau \Phi_1(y) \overline{\tau \Phi_2(y)} dy \\
&= \tau^2 \int_{\Omega} \overline{\Phi_1(x)} \Phi_1(x) \Phi_1(x) \overline{\Phi_2(x)} dx = \tau^2 \langle \Phi_1, \Phi_1 \Phi_1 \overline{\Phi_2} \rangle_{L^2(\Omega)},
\end{aligned} \tag{36}$$

where R is the rotation matrix (12), and we used that $\mathcal{R}\Phi_1 = \tau\Phi_1$ and $\mathcal{R}\Phi_2 = \overline{\tau}\Phi_2$ with $\tau = \exp(2i\pi/3)$, see [1]. From (36) we get

$$(1 - \tau^2) \langle \Phi_1, \Phi_1 \Phi_1 \overline{\Phi_2} \rangle_{L^2(\Omega)} = 0,$$

and thus

$$\langle \Phi_1, \Phi_1 \Phi_1 \overline{\Phi_2} \rangle_{L^2(\Omega)} = 0.$$

Iterating this calculations one can check that:

$$\begin{cases} \langle \Phi_1, |U_0|^2 U_0 \rangle_{L^2(\Omega)} = (2\beta_2 |\psi_1|^2 + \beta_1 |\psi_2|^2) \psi_1 \\ \langle \Phi_2, |U_0|^2 U_0 \rangle_{L^2(\Omega)} = (\beta_1 |\psi_1|^2 + 2\beta_2 |\psi_2|^2) \psi_2, \end{cases} \tag{37}$$

thus recovering the cubic term in (5), with:

$$\beta_1 := \int_{\Omega} |\Phi_1|^4 dx = \int_{\Omega} |\Phi_2|^4 dx, \quad \beta_2 := \int_{\Omega} |\Phi_1|^2 |\Phi_2|^2 dx. \tag{38}$$

It is then easy to see that the stationary version of (5) (i.e. $\partial_t \Xi_1 = \partial_t \Xi_2 = 0$) appears as compatibility condition for the solvability of (32), combining (34, 37) and taking $\mu_1 = 0$ in (32).

3.3. Derivation of the effective mass term

The same multiscale argument as in the previous Section allows to derive the mass term in (5), induced by a suitable perturbation.

As shown in [1, Appendix], breaking the \mathcal{PT} symmetry lifts the conical degeneracy in the dispersion relation of a honeycomb Schrödinger operator $(-\Delta + V)$ admitting Dirac points. Let us consider the following equation:

$$(-\Delta + V + \varepsilon W - \mu_*) u = |u|^2 u, \tag{39}$$

that is, we consider a potential perturbation of (24) where we add a linear term W breaking parity. More precisely, we assume that W is *odd*:

$$W(-x) = -W(x), \quad \forall x \in \mathbb{R}^2. \tag{40}$$

In this case, compared to the analysis in the previous Section, we get an additional term at order $\mathcal{O}(\varepsilon)$ corresponding to the potential εW in (39). Then, we have to compute the projections:

$$\langle W U_0, \Phi_k \rangle_{L^2(\Omega)} = \sum_{j=1}^2 \psi_j \langle W \Phi_j, \Phi_k \rangle_{L^2(\Omega)}, \quad k = 1, 2. \tag{41}$$

Recall that

$$\Phi_2(x) = \overline{\Phi_1(-x)}, \tag{42}$$

and this relation allows us to compute:

$$\begin{aligned}
\langle W \Phi_2, \Phi_1 \rangle_{L^2(\Omega)} &= \int_{\Omega} (W \Phi_2)(x) \overline{\Phi_1(x)} dx = \int_{\Omega} W(x) \overline{\Phi_1(-x)} \overline{\Phi_1(x)} dx \\
&= \int_{-\Omega} W(-y) \overline{\Phi_1(y)} \overline{\Phi_1(-y)} dy = - \int_{\Omega} (W \Phi_2)(y) \overline{\Phi_1(y)} dy \\
&= - \langle W \Phi_2, \Phi_1 \rangle_{L^2(\Omega)},
\end{aligned} \tag{43}$$

where we have also used (40). We thus obtain:

$$\langle W\Phi_2, \Phi_1 \rangle_{L^2(\Omega)} = \langle W\Phi_1, \Phi_2 \rangle_{L^2(\Omega)} = 0. \quad (44)$$

Moreover, arguing as in (43) one easily finds:

$$\langle W\Phi_1, \Phi_1 \rangle_{L^2(\Omega)} = -\langle W\Phi_2, \Phi_2 \rangle_{L^2(\Omega)}, \quad (45)$$

and then:

$$\begin{aligned} \sum_{j=1}^2 \psi_j \langle W\Phi_j, \Phi_1 \rangle_{L^2(\Omega)} &= \psi_1 \langle W\Phi_1, \Phi_1 \rangle_{L^2(\Omega)} \\ \sum_{j=1}^2 \psi_j \langle W\Phi_j, \Phi_2 \rangle_{L^2(\Omega)} &= -\psi_2 \langle W\Phi_1, \Phi_1 \rangle_{L^2(\Omega)}. \end{aligned} \quad (46)$$

Assuming that $m := \langle W\Phi_1, \Phi_1 \rangle_{L^2(\Omega)} > 0$, we obtain the mass term in (5).

4. Proof of the main result

In this section, we prove the existence of branches of bound states for (1) that bifurcate from the trivial solution at the positive band-edge of the spectrum of \mathcal{D} . Those solutions are constructed from bound states of a suitable nonlinear Schrödinger equation (57), which (after scaling) gives the asymptotic profile as $\omega \rightarrow m$.

We start by rewriting (1) componentwise. Setting $\psi = (\psi_1, \psi_2)^T$, equation (1) becomes the system:

$$\begin{cases} -i(\partial_1 - i\partial_2)\psi_2 = (\beta_1|\psi_2|^2 + 2\beta_2|\psi_1|^2)\psi_1 - (m - \omega)\psi_1, \\ -i(\partial_1 + i\partial_2)\psi_1 = -(2\beta_2|\psi_2|^2 + \beta_1|\psi_1|^2)\psi_2 - (m + \omega)\psi_2, \end{cases} \quad (47)$$

that can be regarded as a functional equation of the form:

$$\mathcal{H}(\psi_1, \psi_2) = 0,$$

where $\mathcal{H} : X \times X \rightarrow L^2(\mathbb{R}^2, \mathbb{C}^2)$ is the map defined by:

$$\mathcal{H}(\psi_1, \psi_2) = \begin{pmatrix} -i(\partial_1 - i\partial_2)\psi_2 - (\beta_1|\psi_2|^2 + 2\beta_2|\psi_1|^2)\psi_1 + (m - \omega)\psi_1 \\ -i(\partial_1 + i\partial_2)\psi_1 + (2\beta_2|\psi_2|^2 + \beta_1|\psi_1|^2)\psi_2 + (m + \omega)\psi_2 \end{pmatrix},$$

with

$$X := H^1(\mathbb{R}^2, \mathbb{C}). \quad (48)$$

In what follows, we shall consider the subspace

$$X_r \subset X \times X \quad (49)$$

given by functions of the form (6). For simplicity they will be denoted by (u, v) , where those functions are the radial factors in (6).

4.1. Rescaling the equation

Plugging the ansatz (6) in (47) leads to the follows system for the real valued functions u, v :

$$\begin{cases} -u' + \frac{u}{r} = (\beta_1 u^2 + 2\beta_2 v^2)v - (m - \omega)v, \\ v' = -(2\beta_2 u^2 + \beta_1 v^2)u - (m + \omega)u. \end{cases} \quad (50)$$

Now set $\varepsilon := (m - \omega)$, and consider the following rescaling

$$u_\varepsilon(r) = \varepsilon f(\sqrt{\varepsilon}r), \quad v_\varepsilon(r) = \sqrt{\varepsilon}g(\sqrt{\varepsilon}r), \quad r > 0, \quad (51)$$

so that, by (50), after some straightforward computations we find the equations for $f(\rho), g(\rho)$:

$$\begin{cases} f' + \frac{f}{\rho} = (\varepsilon\beta_1 f^2 + 2\beta_2 g^2)g - g, \\ g' = -(2\beta_2 \varepsilon^2 f^2 + \beta_1 \varepsilon g^2)f - (2m - \varepsilon)f, \end{cases} \quad (52)$$

where we also used the fact that $\varepsilon := m - \omega$ and then $m + \omega = 2m - \varepsilon$.

Remark 4.1. The branch point of the solutions is given by $\varepsilon = 0$. The equivalence between (50) and (52) is valid only for $\varepsilon > 0$, while (52) makes sense for arbitrary $\varepsilon \in \mathbb{R}$.

4.2. Solutions of the rescaled problem

Our goal is to apply the implicit function theorem to prove the existence of a local branch of solutions to (52).

To this aim, we rewrite reformulate the problem as follows. We define the map:

$$\mathcal{F} : \mathbb{R} \times X_r \longrightarrow L^2(\mathbb{R}) \times L^2(\mathbb{R}),$$

with X_r in (49), and acting as:

$$\mathcal{F}(\varepsilon, u_\varepsilon, v_\varepsilon) := \begin{pmatrix} f' + \frac{f}{\rho} - (\varepsilon\beta_1 f^2 + 2\beta_2 g^2)g + g \\ g' + (2\beta_2 \varepsilon^2 f^2 + \beta_1 \varepsilon g^2)f + (2m - \varepsilon)f \end{pmatrix}. \quad (53)$$

Therefore, the original problem is equivalent to the following:

$$\begin{cases} (\varepsilon, u_\varepsilon, v_\varepsilon) \in \mathbb{R} \times X_r, \\ u_\varepsilon, v_\varepsilon \neq 0, \\ \mathcal{F}(\varepsilon, u_\varepsilon, v_\varepsilon) = 0. \end{cases} \quad \varepsilon > 0. \quad (54)$$

Remark 4.2. In order to simplify the notation, without loss of generality, we take $m = 1/2$ and $\beta_2 = 1/2$.

Proposition 4.1. *There exists $\varepsilon_0 > 0$ such that (54) admits a solution for $\varepsilon \in (-\varepsilon_0, \varepsilon)$.*

Remark 4.3. The above proposition is equivalent to the main result stated in Theorem 1.1. Then we equivalently prove the former.

4.2.1. *Solutions for $\varepsilon = 0$.* Take $\varepsilon = 0$. Looking for non-trivial solutions of (54) in X_r we get:

$$\begin{cases} f' + \frac{f}{\rho} = g^3 - g, \\ g' = -f, \end{cases} \quad (55)$$

see Remark 4.2. Then (f, g) solves the following nonlinear Schrödinger equation:

$$\begin{cases} -g'' - \frac{1}{\rho}g' - g^3 + g = 0, \\ f = -g'. \end{cases} \quad (56)$$

Since $\frac{\partial^2}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial}{\partial \rho}$ is the radial part of the two-dimensional Laplacian, one immediately recognizes that g must be a radial solution of the following elliptic equation:

$$-\Delta U - U^3 + U = 0 \quad \text{on } \mathbb{R}^2. \quad (57)$$

It is well known that such equation admits a unique positive radial *ground state solution* U , which is smooth and exponentially decaying at infinity [28, Thm. 8.1.5]. Given such function, we shall consider the solution to (56) given by:

$$(u_0, v_0) = (U, V), \quad V = -U'. \quad (58)$$

4.2.2. *Solutions for small ε .* In order to prove existence of solutions of (54) for small values of ε we have to check the assumptions of the implicit function theorem.

It is not hard to verify that the map \mathcal{F} is of class C^1 , and then we need to prove the following:

Lemma 4.1. *The differential of \mathcal{F} with respect to (u, v) -variables, $D_{(u,v)}\mathcal{F}$, evaluated at $(0, u_0, v_0)$ is an isomorphism.*

The proof of this lemma requires the results that are contained in the following lemmas.

Lemma 4.2. *The operator $D_{(X,Y)}\mathcal{F}(0, u_0, v_0) : X_r \rightarrow L^2(\mathbb{R}^2, \mathbb{R}^2)$ is injective.*

Proof. We need to prove that $\ker\{D_{(u,v)}\mathcal{F}(0, u_0, v_0)\}$ is trivial. For this reason let us consider the linearization of (55) at $(0, u_0, v_0)$ and $(h, k) \in \ker D_{(u,v)}\mathcal{F}(0, u_0, v_0)$, so that

$$D_{(u,v)}\mathcal{F}(0, u_0, v_0)[h, k] := \begin{pmatrix} k' + \frac{k}{\rho} + h - u_0^2 h \\ h' + k \end{pmatrix} = 0. \quad (59)$$

Then h solves

$$-h'' - \frac{1}{\rho}h' - u_0^2h + h = 0$$

that is, h lies in the kernel of the linearization of (57) at the ground state solution U . By known results [29], such kernel is empty and thus $h \equiv 0$ and $k \equiv 0$, proving the Lemma. \square

Now we want to prove that $D_{(u,v)}\mathcal{F}(0, u_0, v_0)$ is surjective, using the Fredholm alternative [30, Thm. 6.6]. Namely, using classical arguments from perturbation theory of linear operators, the claim follows showing that $D_{(u,v)}\mathcal{F}(0, u_0, v_0)$ is given by the sum of an isomorphism and a compact operator.

By (59), let:

$$D_{(u,v)}\mathcal{F}(0, u_0, v_0) = J + K(u_0), \quad (60)$$

where $J, K(u_0) : X_r \rightarrow L^2(\mathbb{R}^2, \mathbb{R}^2)$ are defined as

$$J[h, k] := \left(k' + \frac{k}{\rho} + h, h' + k \right)^T, \quad (61)$$

and

$$K(U)[h, k] := (u_0h, 0)^T. \quad (62)$$

Lemma 4.3. *The operator $J : X_r \rightarrow L^2(\mathbb{R}^2, \mathbb{R}^2)$ is an isomorphism.*

Proof. The operator J is clearly continuous, so that we only need to prove injectivity and surjectivity.

Step (i): J is injective. Assume $(h, k)^T \in X_r$ solves $J[h, k] = 0$. The argument in the proof of Lemma 4.2 gives:

$$-h'' + \frac{1}{\rho}h' + h = 0,$$

i.e.

$$-\Delta h + h = 0,$$

so that multiplying by h and integrating by parts one immediately sees $h \equiv 0$. By (61) we also get $k \equiv 0$.

Step (ii): J is surjective. Let $a, b \in L^2(\mathbb{R}^2)$. We want to prove that there exists $(h, k)^T \in X_r$ such that

$$J[h, k] = (a, b)^T, \quad (63)$$

that is, such that:

$$\begin{cases} k' + \frac{k}{\rho} + h = a, \\ h' + k = b. \end{cases}$$

Assuming that $b \equiv 0$, arguing as in *Step (i)*, we have to find a weak solution $h_1 \in H^1(\mathbb{R}^2)$ of

$$-\Delta h_1 + h_1 = a. \quad (64)$$

The existence of such function is an immediate application of the Lax–Milgram Lemma [30, Cor. 5.8]. Then taking $k_1 = -h_1'$, the pair $(h_1, k_1)^T \in X_r$ solves (63) with $b \equiv 0$. The same argument, exchanging k and h , allows to find a solution $(h_2, k_2)^T \in X_r$ of (63) with $a \equiv 0$. By linearity of J we thus get the claim. \square

Lemma 4.4. *The operator $K(u_0) : X_r \rightarrow L^2(\mathbb{R}^2, \mathbb{R}^2)$ is compact.*

Proof. Let $((h_n, k_n)^T)_n \subset X_r$ be a bounded sequence.

Note that, up to subsequences,

$$(h_n, k_n) \longrightarrow (h, k) \quad \text{in } L^2_{loc}(\mathbb{R}^2, \mathbb{R}^2). \quad (65)$$

On the other hand, since the soliton u_0 tends to zero at infinity, for all $\eta > 0$ there exists $M_\eta > 0$ such that $u_0(x) < \eta$ if $|x| > M_\eta$. Thus,

$$\begin{aligned} \|K(u_0)[h_n, k_n] - K(u_0)[h, k]\|_{L^2(\mathbb{R}^2)} &= \|u_0^2(h_n - h)\|_{L^2(\mathbb{R}^2 \setminus B_{M_\eta})} \\ &\quad + \|u_0^2(h_n - h)\|_{L^2(B_{M_\eta})} \leq C\eta^2 + o(1), \quad \text{as } n \rightarrow \infty. \end{aligned}$$

where $B_{M_\eta} := \{|x| \leq M_\eta\} \subset \mathbb{R}^2$. Then

$$\lim_{n \rightarrow \infty} \|K(u_0)[h_n, k_n] - K(u_0)[h, k]\|_{L^2(\mathbb{R}^2)} \leq C\eta^2, \quad \forall \eta > 0,$$

so that the statement follows. \square

Now, we can combine all the previous results to prove Proposition 4.1.

Proof of Proposition 4.1. Notice that $D_{(X,Y)}\mathcal{F}(0, u_0, v_0)$ is clearly continuous, and it is injective by Lemma 4.2. On the other hand, see (60), Lemmata 4.3 and 4.4 show that:

$$D_{(X,Y)}\mathcal{F}(0, u_0, v_0) = J + K$$

is the sum of an isomorphism and of a compact operator. Then:

$$D_{(X,Y)}\mathcal{F}(0, u_0, v_0) = J + K = J(I + J^{-1}K).$$

Since J is an isomorphism, the map $(I + J^{-1}K)$, which is of the form identity plus compact, is also injective. Then the claim follows by Fredholm's alternative [30, Thm. 6.6]. \square

The proof of Theorem 1.1 immediately follows, as we can now prove Proposition 4.1.

Proof of Proposition 4.1. There holds $\mathcal{F}(0, u_0, v_0) = 0$ and by Proposition 4.1, the differential $D_{(u,v)}\mathcal{F}(0, u_0, v_0)$ is an isomorphism. Then the claim follows by the implicit function theorem. \square

References

- [1] Fefferman C.L., Weinstein M.I. Honeycomb lattice potentials and dirac points. *J. Amer. Math. Soc.*, 2012, **25**, P. 1169–1220.
- [2] Fefferman C.L., Weinstein M.I. Wave packets in honeycomb structures and two-dimensional Dirac equations. *Comm. Math. Phys.*, 2014, **326**, P. 251–286.
- [3] Erdős L., Schlein B., Yau H.-T. Derivation of the cubic non-linear Schrödinger equation from quantum dynamics of many-body systems. *Invent. Math.*, 2007, **167**, P. 515–614.
- [4] Moloney J., Newell A. *Nonlinear optics*. Westview Press, Advanced Book Program, Boulder, CO, 2004.
- [5] Pitaevskii L., Stringari S. Bose-Einstein condensation. *International Series of Monographs on Physics*, The Clarendon Press, Oxford University Press, Oxford, 2003, **116**.
- [6] Fefferman C.L., Weinstein M.I. Waves in honeycomb structures. *Journ'ees 'equations aux d'eriv'ees partielles*, 2012, 12, 12 p.
- [7] Arbunich J., Sparber C. Rigorous derivation of nonlinear Dirac equations for wave propagation in honeycomb structures. *J. Math. Phys.*, 2018, **59**, 011509.
- [8] Borrelli W. Weakly localized states for nonlinear Dirac equations. *Calc. Var. Partial Differential Equations*, 2018, **57**, P. 57–155.
- [9] Borrelli W., Frank R.L. Sharp decay estimates for critical Dirac equations. *Trans. Amer. Math. Soc.*, 2020, **373**, P. 2045–2070.
- [10] Borrelli W. Stationary solutions for the 2D critical Dirac equation with Kerr nonlinearity. *J. Differential Equations*, 2017, **263**, P. 7941–7964.
- [11] Borrelli W. Symmetric solutions for a 2D critical Dirac equation. ArXiv e-prints, 2020, arXiv:2010.04630.
- [12] Ammann B., Grosjean J.-F., Humbert E., Morel B. A spinorial analogue of Aubin's inequality. *Math. Z.*, 2008, **260**, P. 127–151.
- [13] Grosse N. On a conformal invariant of the Dirac operator on noncompact manifolds. *Ann. Global Anal. Geom.*, 2006, **30**, P. 407–416.
- [14] Isobe T. Nonlinear Dirac equations with critical nonlinearities on compact Spin manifolds. *J. Funct. Anal.*, 2011, **260**, P. 253–307.
- [15] Maalaoui A. Infinitely many solutions for the spinorial Yamabe problem on the round sphere. *Nonlinear Differential Equations Appl. NoDEA*, 2016, **23**, 25.
- [16] Borrelli W., Maalaoui A. Some properties of Dirac-Einstein bubbles. *J. Geometric Analysis*, 2020, <https://doi.org/10.1007/s12220-020-00503-1>.
- [17] Maalaoui A., Martino V. Characterization of the Palais-Smale sequences for the conformal Dirac-Einstein problem and applications. *J. Differential Equations*, 2019, **266**, P. 2493–2541.
- [18] Ounaies H. Perturbation method for a class of nonlinear Dirac equations. *Differential Integral Equations*, 2000, **13**, P. 707–720.
- [19] Borrelli W., Carlone R., Tentarelli L. On the nonlinear Dirac equation on noncompact metric graphs. *Journal of Differential Equations*, 2021, **278**, P. 326–357.
- [20] Cazenave T., V´azquez L. Existence of localized solutions for a classical nonlinear Dirac field, *Comm. Math. Phys.*, 1986, **105**, P. 35–47.
- [21] Cuevas-Maraver J., Kevrekidis P.G., et al. Stability of solitary waves and vortices in a 2D nonlinear Dirac model. *Phys. Rev. Lett.*, 2016, **116**, 214101.
- [22] Esteban M.J., S´er´e E. Stationary states of the nonlinear Dirac equation: a variational approach. *Comm. Math. Phys.*, 1995, **171**, P. 323–350.
- [23] Thaller B. *The Dirac equation*, Texts and Monographs in Physics, Springer-Verlag, Berlin, 1992.
- [24] Borrelli W. L'equation de Dirac en physique du solide et en optique non lin´eaire. PhD thesis, Universit´e Paris-Dauphine PSL, (2018).
- [25] Peleg O., Bartal G., et al. Conical diffraction and gap solitons in honeycomb photonic lattices. *Phys. Rev. Lett.*, 2007, **98**, 103901.
- [26] Reed M., Simon B. *Methods of modern mathematical physics. IV. Analysis of operators*, Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1978.
- [27] Ilan B., Weinstein M.I. Band-edge solitons, nonlinear Schrödinger/Gross–Pitaevskii equations, and effective media. *Multiscale Model. Simul.*, 2010, **8**, P. 1055–1101.
- [28] Cazenave T. *Semilinear Schrödinger equations*. Courant Lecture Notes in Mathematics, New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2003, **10**.
- [29] Chang S.-M., Gustafson S., Nakanishi K., Tsai T.-P. Spectra of linearized operators for NLS solitary waves. *SIAM J. Math. Anal.*, 2007–2008, **39**, P. 1070–1111.
- [30] Brezis H. *Functional analysis, Sobolev spaces and partial differential equations*. Universitat, Springer, New York, 2011.