Coupled Mode Equations and Gap Solitons for the 2D Gross-Pitaevskii equation with a non-separable periodic potential

Tomáš Dohnal[1] and Hannes Uecker[2]

[1] Institut für Angewandte und Numerische Mathematik 2, Universität Karlsruhe, Germany

[2] Institut für Mathematik, Carl von Ossietzky Universität Oldenburg, Germany

August 30, 2010

Abstract

Gap solitons near a band edge of a spatially periodic nonlinear PDE can be formally approximated by solutions of Coupled Mode Equations (CMEs). Here we study this approximation for the case of the 2D Periodic Nonlinear Schrödinger / Gross-Pitaevskii Equation with a non-separable potential of finite contrast. We show that unlike in the case of separable potentials [T. Dohnal, D. Pelinovsky, and G. Schneider, J. Nonlin. Sci. **19**, 95–131 (2009)] the CME derivation has to be carried out in Bloch rather than physical coordinates. Using the Lyapunov-Schmidt reduction we then give a rigorous justification of the CMEs as an asymptotic model for reversible non-degenerate gap solitons and even potentials and provide H^s estimates for this approximation. The results are confirmed by numerical examples including some new families of CMEs and gap solitons absent for separable potentials.

1 Introduction

Coherent structures, like gap solitons, in nonlinear periodic wave propagation problems are important both theoretically and in applications. Typical examples include optical waves in photonic crystals and matter waves in Bose-Einstein condensates loaded onto optical lattices. A standard model in these contexts is the Nonlinear Schrödinger/Gross-Pitaevskii equation with a periodic potential, which applies in Kerr-nonlinear photonic crystals [37, 15, 23, 25, 16] as well as in Bose-Einstein condensates loaded onto an optical lattice [20, 30, 26]. Here we consider the case of two spatial dimensions and without loss of generality take the potential 2π–periodic in both directions and, hence, consider

$$\mathrm{i}E_t = -\Delta E + V(x)E + \sigma|E|^2E, \quad V(x_1+2\pi, x_2) = V(x_1, x_2+2\pi) = V(x),\ x \in \mathbb{R}^2,\ t \in \mathbb{R}, \tag{1.1}$$

with $E = E(x,t) \in \mathbb{C}$, $\sigma = \pm 1$ and $V \in H^m_{\mathrm{loc}}(\mathbb{R}^2), m > 1, m \in \mathbb{R}$.

We are interested in stationary gap solitons (GSs) $E(x,t) = \phi(x)e^{-\mathrm{i}\omega t}$. Thus ϕ solves

$$(-\Delta + V(x) - \omega)\phi + \sigma|\phi|^2\phi = 0, \tag{1.2}$$

where soliton is understood in the sense of a solitary wave, which means that $|\phi(x)| \to 0$ exponentially as $|x| \to 0$. Necessarily, then ω has to lie in a gap of the essential spectrum of the operator $L := -\Delta + V(x)$, hence the name "gap soliton." From the phenomenological and experimental point of view multidimensional GSs have been widely studied in the context of both photonic crystals [6, 25, 15, 17, 12, 16] and Bose-Einstein condensates [1, 26, 7].

The essential spectrum of L is given by the so called band structure, and for our analysis we choose ω close to a band edge, i.e., $\omega = \omega_* + \varepsilon^2\Omega,\ 0 < \varepsilon \ll 1$, where ω_* is an edge of a band gap and Ω has a sign chosen so that ω lies inside the gap. Using a multiple scales expansion one may formally derive coupled mode equations (CMEs) to approximate envelopes of the gap solitons near gap edges. CMEs are a constant coefficient problem formulated in slowly varying variables. They are, therefore, typically more amiable to analysis and also cheaper for numerical approximations compared to the original system (1.2). The multiple scales approach has been used both for the Gross-Pitaevskii and Maxwell equations with infinitesimal, i.e. $\mathcal{O}(\varepsilon)$, contrast in the periodicity $V(x)$ [2, 6, 34, 3, 5, 12] as well as with finite contrast [11, 33, 13]. The main difference in the asymptotic approximation of the two cases is that for infinitesimal contrasts the expansion modes are Fourier waves while for finite contrast they are Bloch waves. However, in dimension two and higher sufficiently large (finite) contrast is necessary to generate band gaps due to overlapping of bands in the corresponding homogeneous medium. The only exception is the semi-infinite gap of the band structure of L of the Gross-Pitaevskii equation. As a result, gap solitons in finite gaps of the Gross-Pitaevskii equation and in any gap of Maxwell systems in dimensions two and higher can only exist for finite contrast structures.

Localized solutions of CMEs formally yield gap solitons of the original system. However, the formal derivation of the CMEs, discarding some error at higher order in ε, does not imply that all localized solutions of the CMEs yield gap solitons. For this we need to estimate the error in some function space and to show the persistence of the CME solitons under perturbation of the CMEs. A famous result concerning non–persistence is the non-existence of breathers in perturbations of the sine–Gordon equation, e.g., [10]. On the other hand, GSs are known to exist in every band gap of L, see, e.g. [36, 27]. The proofs, however, are based on variational methods and do not relate GSs to solutions of the CMEs.

A rigorous justification of the CMEs has been given for (1.2) in 1D in [28], and in 2D in [13], but only for the case of a separable potential

$$V(x_1, x_2) = W_1(x_1) + W_2(x_2).$$

Here we transfer this result to not necessarily separable potentials, where we need some minimal smoothness, namely $V \in H^m_{\text{loc}}(\mathbb{R}^2), m > 1$. As an example we choose

$$V(x_1, x_2) = 1 + (\eta - 1)W(x_1)W(x_2), \quad W(s) = \frac{1}{2}\left[\tanh\left(7\left(s - \frac{2\pi}{5}\right)\right) + \tanh\left(7\left(\frac{8\pi}{5} - s\right)\right)\right], \quad (1.3)$$

which represents a square geometry with smoothed-out edges (Fig. 1). We choose the contrast η in V so that two finite gaps appear in the band structure of the corresponding linear eigenvalue problem. One main difference between the separable and non-separable case lies in the fact that for the non-separable case band edges may be attained at wavenumbers not within the set of vertices of the first irreducible Brillouin zone. Then the CME derivation and justification is impossible to carry out in

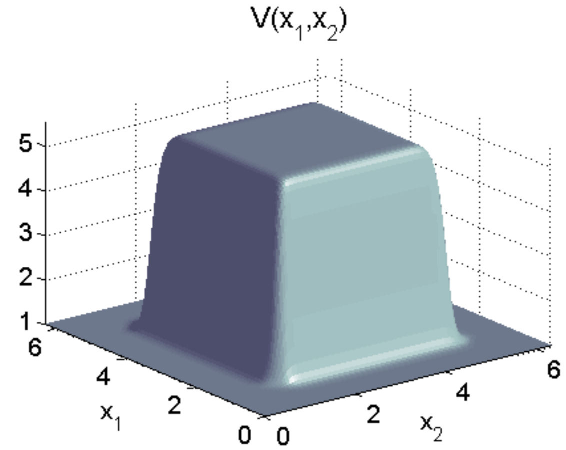

Figure 1: The periodic potential V in (1.3) over the Wigner-Seitz cell.

physical variables and has to be performed in Bloch variables. This case occurs at least at one band edge of the potential (1.3), and the presented CMEs corresponding to this edge have, to our knowledge, not been studied before. Similarly, the GSs which we show to bifurcate from this edge are new.

In §2 we discuss in detail the band structure for (1.3) and the associated Bloch eigenfunctions, together with their symmetries. Then in §3 we first give the formal derivation of the CME in physical space, reporting a failure in one case where the band edge is attained simultaneously at four wave numbers outside the set of vertices of the first Brillouin zone, and present a general CME derivation in Bloch variables. The existence of gap solitons is proved in §4 based on the existence of special (namely reversible and non-degenerate, see below) localized solutions of the CMEs, in the following sense.

Function spaces and notation. For $m \in \mathbb{N}$, the Sobolev spaces $H^m(\mathbb{R}^2)$ are classically defined as $H^m(\mathbb{R}^2) := \{u \in L^2(\mathbb{R}^d) : \partial_x^\alpha u \in L^2(\mathbb{R}^2) \text{ for } |\alpha| \le m\}$, with norm $\|u\|_{H^s} = \left(\sum_{|\alpha|\le m} \|\partial_x^\alpha u\|_{L^2}^2\right)^{1/2}$, where $\partial_x^\alpha u$ denotes the distributional derivative, see, e.g., [4]. Then, for $s = m \in \mathbb{N}$, Fourier transform

$$\hat{\phi}(k) := (\mathcal{F}\phi)(k) := \frac{1}{(2\pi)^2}\int_{\mathbb{R}^2} \phi(x)e^{-\mathrm{i}k\cdot x}\,\mathrm{d}x, \quad \phi(x) = (\mathcal{F}^{-1}\hat{\phi})(k) := \int_{\mathbb{R}^2} \hat{\phi}(k)e^{\mathrm{i}k\cdot x}\,\mathrm{d}k, \tag{1.4}$$

is an isomorphism from $H^s(\mathbb{R}^2)$ to

$$L_s^2(\mathbb{R}^2) := \{\hat{\phi} \in L^2(\mathbb{R}^2) : \|\hat{\phi}\|_{L_s^2} := \|(1+|k|)^s\hat{\phi}\|_{L^2} < \infty\}, \quad \text{i.e. } C_1\|\hat{\phi}\|_{L_s^2} \le \|\phi\|_{H^s} \le C_2\|\hat{\phi}\|_{L_s^2}. \tag{1.5}$$

From the applied point of view we could restrict to integer s. However, since our analysis is strongly based on Fourier transform, it is conceptually cleaner to use a definition of Sobolev spaces based on L_s^2 with arbitrary $s \ge 0$. Thus, henceforth we use $H^s(\mathbb{R}^2) := \mathcal{F}^{-1}L_s^2(\mathbb{R}^2)$ as definition for $0 \le s \in \mathbb{R}$. This also gives a very simple proof of the Sobolev embedding theorem $\|\phi\|_{C^k} \le C\|\phi\|_{H^s}$ for $k < s-1$, see Lemma 4.2 below.

Main result. Let $s \ge 2$ and $V \in H_{\text{loc}}^{\lceil s\rceil-1+\delta}(\mathbb{R}^2), \delta > 0$, where $\lceil s\rceil$ is the smallest integer larger than or equal to s, and V even in x_1, x_2. Let $\mathbf{A} = (A_1, \ldots, A_N)$ be a reversible non-degenerate localized solution of the CMEs with $\mathbf{A} \in [H^q(\mathbb{R}^2)]^N$ for all $q \ge 0$. Then for $\omega = \omega_* + \varepsilon^2\Omega$ with ε^2 sufficiently small there exists a GS ϕ_{GS} for (1.2), such that $\phi_{GS} \in H^s(\mathbb{R}^2)$, and ϕ_{GS} can be approximated by

$$\varepsilon\phi^{(0)} = \varepsilon \sum_{j=1}^{N} A_j(\varepsilon x) u_{n_j}(k^{(j)}; x), \tag{1.6}$$

where $u_{n_j}(k^{(j)}; x) \in H^{\lceil s \rceil + 1 + \delta}_{\text{loc}}(\mathbb{R}^2)$ are the pertinent Bloch waves, $j = 1, \ldots, N$. In detail, we prove

$$\|\phi_{GS} - \varepsilon\phi^{(0)}\|_{H^s(\mathbb{R}^2)} \leq C\varepsilon^{2/3}, \tag{1.7}$$

where the estimate can be improved in special cases, see below.

Note that $\|\varepsilon\phi^{(0)}\|_{L^\infty(\mathbb{R}^2)} = \mathcal{O}(\varepsilon)$ but $\|\varepsilon\phi^{(0)}\|_{L^2(\mathbb{R}^2)} = \mathcal{O}(1)$ such that the error in (1.7) is indeed smaller than the approximation. The proof is based on a Lyapunov–Schmidt reduction and analysis of suitable extended CMEs. In §5 we give some numerical illustrations and verify convergence of the asymptotic coupled mode approximation.

Remark 1.1 The (apparent) lack of an estimate $\|\phi_{GS} - \varepsilon\phi^{(0)}\|_{L^\infty(\mathbb{R}^2)} \leq C\varepsilon^{1+\beta}$ with $\beta > 0$ is a disadvantage of our analysis. It is due to the fact that we work in L^2–spaces in Fourier resp. Bloch variables, while a direct L^∞ estimate in physical variables would require working in L^1–spaces in Fourier resp. Bloch variables. This is not possible due to a technical obstacle, see [13, §8]. On the other hand, Hilbert spaces L^2 are also more natural spaces to work in since they allow direct transition from physical to Bloch variables and back. Moreover, localization in x in the sense of decay to 0 for $|x| \to \infty$ follows directly in spaces of integrable functions. Note also that based on the formal asymptotics, instead of $\varepsilon^{2/3}$ one can expect the convergence rate ε^1 in H^s in (1.7) which is the approximate rate observed in our numerical examples. Finally, in Remark 4.12 we explain how the long wave modulational form of the formal asymptotics allows to obtain an $\mathcal{O}(\varepsilon^{1+\beta})$ convergence of the error in L^∞ from the $\mathcal{O}(\varepsilon^\beta)$ convergence in H^s. However, a completely rigorous calculation is lengthy and therefore here we content ourselves with (1.7). $\rfloor$

Remark 1.2 Time-dependent CMEs have been justified in 1D for infinitesimal [19, 32] and finite [8] contrast, and in 2D for finite contrast under the condition of a separable potential in [13, §7]. Here justification means that non–stationary solutions of (1.1) can be approximated by CME dynamics over long but finite intervals. Given the analysis below, this result of [13] can be immediately transfered to our non–separable case. $\rfloor$

2 Band structure and Bloch functions

Let $\omega_n(k), n \in \mathbb{N}$, denote the spectral bands and $u_n(k;x)$ the corresponding Bloch functions of the operator $L := -\Delta + V(x)$, where k runs through the first Brillouin zone $\mathbb{T}^2 = (-1/2, 1/2]^2$. This means that $(\omega_n(k), u_n(k;x))$ is an eigenpair of the quasiperiodic eigenvalue problem

$$\begin{aligned} &Lu_n(k;x) = \omega_n(k)u_n(k;x), \;\; x \in \mathbb{P}^2 := [0, 2\pi)^2, \\ &u_n(k;(2\pi, x_2)) = \mathrm{e}^{\mathrm{i}2\pi k_1} u_n(k;(0,x_2)), \;\; u_n(k;(x_1, 2\pi)) = \mathrm{e}^{\mathrm{i}2\pi k_2} u_n(k;(x_1, 0)). \end{aligned} \tag{2.1}$$

The Bloch functions $u_n(k;x)$ can be rewritten as

$$u_n(k;x) = \mathrm{e}^{\mathrm{i}k\cdot x} p_n(k;x), \quad \text{where } p_n \text{ is } 2\pi\text{-periodic in both } x_1 \text{ and } x_2, \text{ and fulfills} \tag{2.2}$$

$$\tilde{L}(k;x)p_n(k;x) := [(\mathrm{i}\partial_{x_1} - k_1)^2 + (\mathrm{i}\partial_{x_2} - k_2)^2 + V(x)]p_n(k;x) = \omega_n(k)p_n(k;x). \tag{2.3}$$

For each $k \in \mathbb{T}^2$ the operator $\tilde{L}(k;\cdot)$ is elliptic and self adjoint in $L^2(\mathbb{P}^2)$, which immediately yields the existence of infinitely many real eigenvalues $\omega_n(k), n \in \mathbb{N}$ with $\omega_n(k) \to \infty$ as $n \to \infty$. The spectrum of L equals $\bigcup_{n\in\mathbb{N},k\in\mathbb{T}^2} \omega_n(k)$, see Theorem 6.5.1 in [14]. Moreover, if V satisfies $V(-x_1,x_2) = V(x_1,-x_2) = V(x)$ and $V(x) = V(x_2,x_1)\ \forall x \in \mathbb{R}^2$, the $\omega_n(k)$ can be recovered from their values in the irreducible Brillouin zone B_0, see Fig. 2.

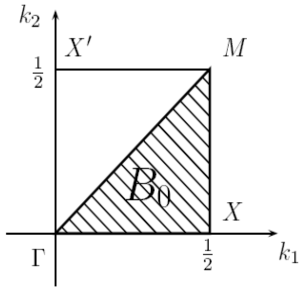

Figure 2: The first irreducible Brillouin zone B_0 for the two-dimensional potential V.

From (2.2) we also note that for $x \in \mathbb{P}^2$ and $n \in \mathbb{Z}^2$ we have

$$u_n(k;(x_1 + 2n_1\pi, x_2 + 2n_2\pi)) = \mathrm{e}^{2\pi\mathrm{i}n\cdot k} u_n(k;x). \tag{2.4}$$

Gaps in the spectrum of L have to be confined by extrema of bands. Unlike in the case of the separable potential $V(x_1,x_2) = W_1(x_1) + W_2(x_2)$ the extrema of ω_n within B_0 do not have to occur only at $k = \Gamma, X$ and M but may occur anywhere throughout B_0. Thus we need to solve (2.1) for all $k \in B_0$.

In the example (1.3) we choose the contrast η so that two finite band gaps are open. Our computations show that this happens, for instance, at $\eta = 5.35$, which we select. The band structure of L is computed in a 4th order centered finite-difference discretization. For reasons of tradition we plot in Fig. 3 the band structure along ∂B_0. In Fig. 4 we plot the first few bands over B. Though not true in general [21], in our case the extrema of the first 6 bands fall on ∂B_0. The dots in Figs. 3 and 4 label those band edge extrema which also mark gap edges. One of these extrema in Fig. 3 (corresponding to 4 extrema in Fig. 4) falls out of the vertex set $\{\Gamma, X, M\}$. We also label in Fig. 3 the first 7 bands $\omega_1, \ldots, \omega_7$ and the gap edges $s_1, s_2, \ldots, s_5$. The edge values with six converged decimal places are

$$s_1 \approx 1.502064,\ \ s_2 \approx 1.702299,\ \ s_3 \approx 2.034433,\ \ s_4 \approx 3.807113,\ \text{ and } s_5 \approx 3.832442.$$

For any corresponding value of k each gap edge eigenvalue of (2.1) is simple because none of the edge-defining extrema belongs to more than one band. We now combine this with symmetries of the problem to find symmetries of the Bloch functions, which will be needed in the derivation of the CME. In the rest of this section we assume $\|u_n(k;\cdot)\|_{L^2(\mathbb{P}^2)} = 1$, where we are, of course, still free to multiply

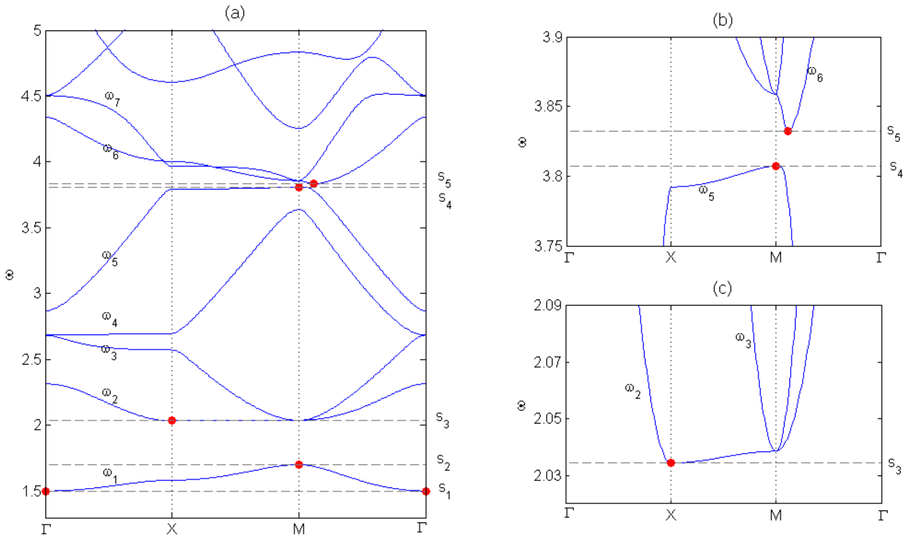

Figure 3: (a) Band structure of L with $\eta = 5.35$ along ∂B_0. Red dots label band extrema at gap edges $s_1, \ldots, s_5$. (b) Detail in the second finite gap. (c) Detail near the edge s_3 showing that ω_2 is not flat for k between X and M.

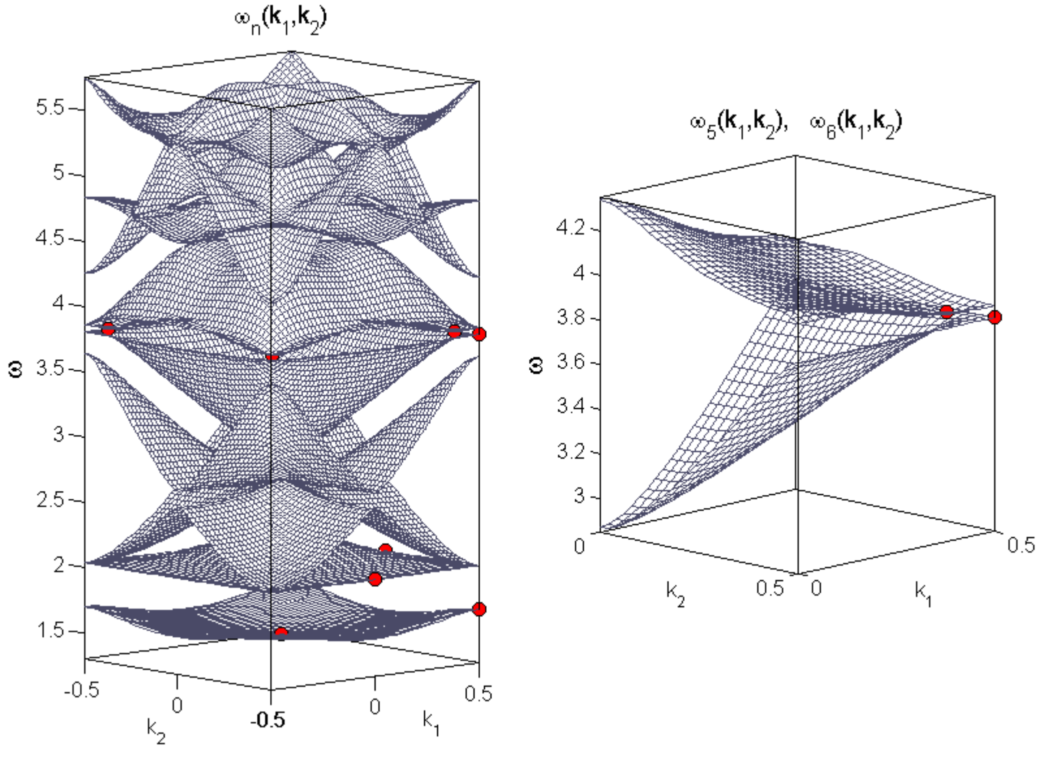

Figure 4: Band structure of L with $\eta = 5.35$. On the right a detail of ω_5 and ω_6 near the second finite gap.

any mode u_n by a phase factor e^{ia}, $a \in \mathbb{R}$. See also Remark 2.1.

First, due to evenness of $V(x)$ in (1.3) in both variables we have

$$
\begin{aligned}
u_n((-k_1,k_2);(x_1,x_2)) &= \mathrm{e}^{\mathrm{i}a_1} u_n((k_1,k_2);(2\pi - x_1, x_2)),\\
u_n((k_1,-k_2);(x_1,x_2)) &= \mathrm{e}^{\mathrm{i}a_2} u_n((k_1,k_2);(x_1, 2\pi - x_2)),\\
\omega_n(-k_1,k_2) = \omega_n(k_1,&-k_2) = \omega_n(k).
\end{aligned}
\tag{2.5}
$$

for some $a_1, a_2 \in \mathbb{R}$. Note that when $(-k_1,k_2) \doteq (k_1,k_2)$, where $k \doteq l$ reads "k congruent to l" and means $k = l + m$ for some $m \in \mathbb{Z}^2$, a renormalization of the phase cannot be used in general to obtain $a_1 = 0$ because $u_n((k_1,k_2);(\pi,x_2)) = 0 \ \forall x_2 \in \mathbb{P}$ is possible. Similarly, when $(k_1,-k_2) \doteq (k_1,k_2)$, we cannot generally achieve $a_2 = 0$ because $u_n((k_1,k_2);(x_1,\pi)) = 0 \ \forall x_1 \in \mathbb{P}$ is possible.

Next, the symmetry $V(x_1,x_2) = V(x_2,x_1)$ implies

$$
u_n((k_1,k_2);(x_1,x_2)) = \mathrm{e}^{\mathrm{i}a} u_n((k_2,k_1);(x_2,x_1)), \qquad \omega_n(k_1,k_2) = \omega_n(k_2,k_1). \tag{2.6}
$$

for some $a \in \mathbb{R}$. Similarly to the case of symmetry (2.5), when $k_1 \doteq k_2$, one cannot, in general, apply renormalization to achieve $a = 0$ because $u_n((k_1,k_1);(x_1,x_1)) = 0 \ \forall x_1 \in \mathbb{P}$ is possible.

Finally, since L is real, $\overline{u_n(k;x)}$ satisfies (2.1) with the factors in the boundary conditions replaced by $\mathrm{e}^{-\mathrm{i}2\pi k_1}$ and $\mathrm{e}^{-\mathrm{i}2\pi k_2}$. Thus

$$
u_n(-k;x) = \overline{u_n(k;x)}, \qquad \omega_n(-k) = \omega_n(k). \tag{2.7}
$$

Note that unlike in (2.5) and (2.6) no exponential factor appears in (2.7). This is because for the conjugation symmetry (2.7) such a factor $\mathrm{e}^{\mathrm{i}a}$ can be easily removed via multiplication by $\mathrm{e}^{-\mathrm{i}a/2}$.

Remark 2.1 If, e.g., $(-k_1,k_2)$ is not congruent to (k_1,k_2), we can, for instance, multiply $u_n((-k_1,k_2);\cdot)$ by $\mathrm{e}^{\mathrm{i}a_1}$ and obtain $u_n((-k_1,k_2);(x_1,x_2)) = u_n(k;(2\pi - x_1, x_2))$. However, one will generally not be able to simultaneously ensure also $u_n((k_1,k_2);(x_1,x_2)) = u_n((k_2,k_1);(x_2,x_1))$ in (2.6) and therefore we stick to the factors in (2.5) and (2.6). ⌋

Let us consider implications of the above three symmetries (2.5), (2.6) and (2.7) for our example (1.3) and plot the gap edge Bloch functions in Fig. 5. Each edge s_1, s_2 and s_4 is attained only at a single extremum within B, namely at $k = \Gamma, M$ and M respectively. The corresponding Bloch functions are $u_1((0,0);x)$, $u_1((1/2,1/2);x)$ and $u_5((1/2,1/2);x)$ respectively, which are all real due to (2.7). The edge s_3 is attained by extrema at $k = X$ and X' with the Bloch functions $u_2((1/2,0);x)$ and $u_2((0,1/2);x)$. Referring to (2.6) only $u_2((1/2,0);(x_1,x_2))$ is plotted, which is again real due to (2.7). Finally, the edge s_5 is attained by 4 extrema, namely at $k = (k_c,k_c), (-k_c,k_c), (-k_c,-k_c)$ and $(k_c,-k_c)$, where the numerically computed value, converged to 6 decimal places, is $k_c \approx 0.439028$. The corresponding Bloch functions are $u_6((k_c,k_c);x)$, $u_6((-k_c,k_c);x)$, $u_6((-k_c,-k_c);x)$ and $u_6((k_c,-k_c);x)$. Due to (2.5) and (2.7) and because $k_c \notin \{0, 1/2\}$, we can normalize the Bloch functions so that $u_6((-k_c,k_c);(x_1,x_2)) = u_6((k_c,k_c);(2\pi - x_1, x_2))$, $u_6((k_c,-k_c);(x_1,x_2)) = u_6((k_c,k_c);(x_1, 2\pi - x_2))$, $u_6((-k_c,-k_c);(x_1,x_2)) = u_6((k_c,k_c);(2\pi - x_1, 2\pi - x_2)) = \overline{u_6((k_c,k_c);(x_1,x_2))}$. Thus it suffices to plot only $u_6((k_c,k_c);(x_1,x_2))$. In addition, (2.6) and the fact that $u_6((k_c,k_c);(x_1,x_1))$ is not identically zero imply $u_6((k_c,k_c);(x_1,x_2)) = u_6((k_c,k_c);(x_2,x_1))$. The Bloch waves $u_6((k_c,k_c);x)$ and $u_6((-k_c,-k_c);x)$ are, therefore, symmetric about the diagonal $x_1 = x_2$.

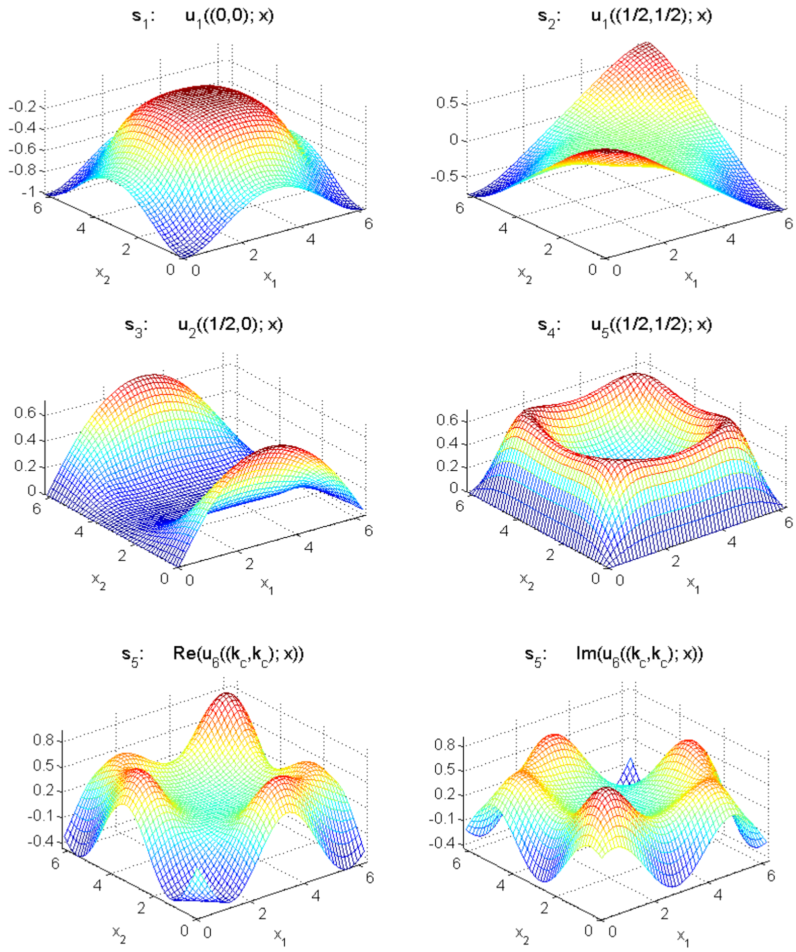

Figure 5: Bloch functions at gap edges $s_1, s_2 \dots, s_5$.

Fig. 5 shows that all the Bloch functions at $s_1, s_2, \dots, s_4$ are either even or odd in each variable. This actually follows from (2.5) and the fact that these gap edges occur at $k \in \Sigma = \{\Gamma, X, X', M\}$. As each coordinate of any $k \in \Sigma$ is either 0 or $\frac{1}{2}$, the eigenvalue problem (2.1) is real and we can choose $a_1, a_2 \in \{0, \pi\}$ in (2.5). The choice $a_1 = a_2 = 0$ is, however, in general impossible as explained after (2.5). Taking, for instance, $k_1 = \frac{1}{2}$, we have

$$\begin{aligned} u_n((1/2, k_2); x) &= \pm u_n((-1/2, k_2); (2\pi - x_1, x_2)) = \pm u_n((1/2, k_2); (2\pi - x_1, x_2)) \\ &= \pm e^{\mathrm{i}2\pi\frac{1}{2}} u_n((1/2, k_2); (-x_1, x_2)) = \mp u_n((1/2, k_2); (-x_1, x_2)), \end{aligned}$$

where the second equality follows from $1-$periodicity of u_n in each $k-$coordinate and the third equality from the quasi-periodic boundary conditions in (2.1). Similarly, we get $u_n((0, k_2); (x_1, x_2)) = \pm u_n((0, k_2); (-x_1, x_2))$. Therefore, we have the following

Lemma 2.2 *Suppose $V(x)$ is even in the variable x_j for some $j \in \{1, 2\}$. If $k_j \in \{0, 1/2\}$ and $\omega_n(k)$, as an eigenvalue of* (2.1) *has geometric multiplicity 1, then $u_n(k; x)$ is either even or odd in x_j.*

3 Formal asymptotic derivation of Coupled Mode Equations

Gap solitons in the vicinity of a given band edge are expected to be approximated by the Bloch waves at the band edge modulated by slowly varying spatially localized envelopes. The governing equations for the envelopes, called Coupled Mode Equations (CMEs), can be derived by a formal asymptotic procedure. Here we are interested in gap solitons $E(x,t) = \phi(x)e^{-\mathrm{i}\omega t}$ with $\omega = \omega_* + \varepsilon^2\Omega,\ 0 < \varepsilon \ll 1$, where ω_* is an edge of a given band gap of a fixed ($\mathcal{O}(1)$) width, and Ω has a sign chosen so that ω lies inside the gap. The leading order term in the asymptotic expansion of the spatial profile ϕ is expected to be

$$\phi(x) \sim \varepsilon \sum_{j=1}^{N} A_j(\varepsilon x) u_{n_j}(k^{(j)}; x), \tag{3.1}$$

where $\{u_{n_j}(k^{(j)}; x)\}_{j=1}^{N}$ are the Bloch waves at $\omega = \omega_*$ and (2.4) is used for $x \notin \mathbb{P}^2$. We assume:
Assumption A.1 The band structure defined by (2.1) has a gap with an edge (lower/upper) defined by $0 < N < \infty$ extrema (maxima/minima) of the bands $\omega_n(k)$. The extrema occur for bands $\omega_{n_j}(k), j = 1, \dots N$ at the corresponding points $k^{(j)} \in B$, where $k \mapsto \omega_{n_j}(k)$ is analytic in k locally near $k^{(j)}$.
Assumption A.2 The quadratic form $\partial_{k_1}^2\omega_{n_j}(k^{(j)})x^2 + 2\partial_{k_1}\partial_{k_2}\omega_{n_j}(k^{(j)})xy + \partial_{k_2}^2\omega_{n_j}(k^{(j)})y^2$ defined by the Hessian of ω_{n_j} at $k = k^{(j)}$ is (positive or negative) definite.

Remark 3.1 a) Analyticity of the n_j-th band near $k^{(j)}$ holds if $\omega_{n_j}(k^{(j)})$ is simple, see [39].
b) The definiteness in A.2 ensures that the extremum of ω_{n_j} at $k = k^{(j)}$ is quadratic and that the resulting CMEs are of second order. Unlike in the separable case [13] it is possible that $\partial_{k_1}\partial_{k_2}\omega_{n_j}(k^{(j)}) \neq 0$, which then leads to CMEs with mixed second order derivatives.
c) The Bloch waves $u_{n_j}(k^{(j)}; \cdot), j{=}1, \dots, N$ defined by the extrema are called "resonant".
d) Assumptions A.1 and A.2 are satisfied by the potential (1.3) with $\eta = 5.35$ at all the gap edges $s_1, \dots, s_5$. ⌋

Remark 3.2 The approximation (3.1) with the same ε-scaling applies also to gap solitons in an $\mathcal{O}(\varepsilon^2)$-wide gap which closes at $\omega = \omega_*$ as $\varepsilon \to 0$ in such a way that the plane $\omega = \omega_*$ at $\varepsilon = 0$ is not intersected by any band but is tangent to bands at N extremal points. $u_{n_1}, \dots, u_{n_N}$ are then the resonant Bloch waves at $\omega = \omega_*$ at $\varepsilon = 0$. Such a case was studied in [13] for a separable periodic potential.

The above discussion is not limited to the case of the Gross-Pitaevskii equation but applies to general differential equations with periodic coefficients, as it depends only on the band structure. A typical example is Maxwell's equations with spatially periodic coefficients. ⌋

We now give the derivation of CMEs under the assumptions A.1 and A.2. For the example (1.3) with $\eta = 5.35$ we first review the derivation in physical variables $\phi(x)$ near $\omega = s_3$, then comment on an obstacle for this calculus near $\omega = s_5$, and therefore present a derivation in the general case in the so called Bloch variables which avoids this obstacle. Finally, we apply this general procedure to all the five gap edges of the example (1.3).

3.1 CME derivation in Physical Variables $\phi(x)$

The ansatz in physical variables is

$$\begin{aligned}\phi(x) &= \varepsilon\phi^{(0)}(x)+\varepsilon^2\phi^{(1)}(x)+\varepsilon^3\phi^{(2)}(x)+\mathcal{O}(\varepsilon^4),\\ \varepsilon\phi^{(0)}(x) &= \varepsilon\sum_{j=1}^{N}A_j(y)u_{n_j}(k^{(j)};x), \qquad \omega=\omega_*+\varepsilon^2\Omega, \qquad y=\varepsilon x, \qquad 0<\varepsilon\ll 1.\end{aligned} \tag{3.2}$$

To review the derivation of the CMEs we choose $\omega_* = s_3$ for the example (1.3) with $\eta = 5.35$.

3.1.1 CMEs near the gap edge $\omega = s_3$

At the edge s_3 we have $N = 2, n_1 = n_2 = 2, k^{(1)} = X$ and $k^{(2)} = X'$, i.e. the two resonant Bloch waves are $v_1(x) := u_2(X;x)$ and $v_2(x) := u_2(X';x)$. Using (2.4), Lemma 2.2 and (2.7), we have that v_1 is odd and $2\pi-$antiperiodic in x_1 and even and $2\pi-$periodic in x_2. Opposite symmetries hold for v_2. Moreover, (2.7) implies that v_1 and v_2 are real. We normalize the Bloch functions $v_{1,2}$ over their common period $[-2\pi, 2\pi]^2$ so that $\|v_j\|_{L^2([-2\pi,2\pi]^2)} = 1,\ j = 1, 2$.

Substituting (3.2) in (1.2) leads to a hierarchy of problems at distinct powers of ε, each of which we try to solve within the space of functions 4π-periodic in both x_1 and x_2, invoking the Fredholm alternative (see e.g. chapter 3.4 of [35]) where necessary. At $\mathcal{O}(\varepsilon)$ we have the linear eigenvalue problem $[L - s_3]v_j(x) = 0,\ j = 1, 2$. At $\mathcal{O}(\varepsilon^2)$ we have

$$[L-s_3]\phi^{(1)} = 2\left(\partial_{y_1}A_1\partial_{x_1}v_1+\partial_{y_1}A_2\partial_{x_1}v_2+\partial_{y_2}A_1\partial_{x_2}v_1+\partial_{y_2}A_2\partial_{x_2}v_2\right).$$

By differentiating the eigenvalue problem (2.1) with respect to $k_j, j \in \{1, 2\}$ and evaluating at $n = 2, k = X = (1/2, 0)$, we find that

$$[L-s_3]v_1^{(x_j)}(x) = 2\partial_{x_j}v_1, \tag{3.3}$$

and similarly $[L - s_3]v_2^{(x_j)}(x) = 2\partial_{x_j}v_2$, where

$$v_1^{(x_j)}(x) = -\mathrm{i}(\partial_{k_j}p_2(X;x))e^{\mathrm{i}X\cdot x} \text{ and } v_2^{(x_j)}(x) = -\mathrm{i}(\partial_{k_j}p_2(X';x))e^{\mathrm{i}X'\cdot x}$$

are called generalized Bloch functions [29]. Thus $\phi^{(1)} = \partial_{y_1}A_1v_1^{(x_1)}+\partial_{y_1}A_2v_2^{(x_1)}+\partial_{y_2}A_1v_1^{(x_2)}+\partial_{y_2}A_2v_2^{(x_2)}$. (3.3) implies that $v_n^{(x_j)}(x)$ is odd/even in x_j if $v_n(x)$ is even/odd in x_j respectively.

At $\mathcal{O}(\varepsilon^3)$ we obtain the CMEs. We have

$$\begin{aligned}[L-s_3]\phi^{(2)} = \;& \Omega(A_1v_1+A_2v_2)+\Delta_{y_1,y_2}A_1v_1+\Delta_{y_1,y_2}A_2v_2\\ &+2\Big[\partial_{y_1}^2A_1\partial_{x_1}v_1^{(x_1)}+\partial_{y_1}^2A_2\partial_{x_1}v_2^{(x_1)}+\partial_{y_2}^2A_1\partial_{x_2}v_1^{(x_2)}+\partial_{y_2}^2A_2\partial_{x_2}v_2^{(x_2)}\\ &\quad+\partial_{y_1}\partial_{y_2}A_1\partial_{x_1}v_1^{(x_2)}+\partial_{y_1}\partial_{y_2}A_2\partial_{x_1}v_2^{(x_2)}+\partial_{y_1}\partial_{y_2}A_1\partial_{x_2}v_1^{(x_1)}+\partial_{y_1}\partial_{y_2}A_2\partial_{x_2}v_2^{(x_1)}\Big]\\ &-\sigma\Big[\textstyle\sum_{j=1}^2|A_j|^2A_jv_j^3+2|A_1|^2A_2v_1^2v_2+2|A_2|^2A_1v_2^2v_1+A_1^2\bar{A}_2v_1^2v_2+A_2^2\bar{A}_1v_2^2v_1\Big],\end{aligned}$$

and the Fredholm alternative requires the right hand side to be $L^2(-2\pi, 2\pi]^2$-orthogonal to v_1 and v_2, the two generators of $\mathrm{Ker}(L^* - s_3)$. Taking the inner product, we see that the terms $\langle v_1, v_2\rangle$ and $\langle v_2, v_1\rangle$ in the inner product vanish due to orthogonality of Bloch waves. Many additional terms vanish due to odd or 2π-antiperiodic integrands (in at least one variable). Namely, in the inner product of the right hand side with v_1 the integrals $\langle \partial_{x_1} v_2^{(x_1)}, v_1\rangle, \langle \partial_{x_2} v_2^{(x_2)}, v_1\rangle, \langle \partial_{x_1} v_1^{(x_2)}, v_1\rangle$ and $\langle \partial_{x_2} v_1^{(x_1)}, v_1\rangle$ vanish due to odd integrands and the integrals $\langle v_2^3, v_1\rangle, \langle v_1^2 v_2, v_1\rangle, \langle v_1^2 v_2, v_1\rangle$ $\langle \partial_{x_1} v_2^{(x_2)}, v_1\rangle$ and $\langle \partial_{x_2} v_2^{(x_1)}, v_1\rangle$ due to $2\pi-$antiperiodic integrands. An analogous discussion applies for the orthogonality with the respect to v_2. The remaining terms have to be set to zero, which leads to the CMEs for the envelopes A_1 and A_2:

$$\begin{aligned}\Omega A_1 + \alpha_1 \partial_{y_1}^2 A_1 + \alpha_2 \partial_{y_2}^2 A_1 - \sigma\left[\gamma_1 |A_1|^2 A_1 + \gamma_2(2|A_2|^2 A_1 + A_2^2 \bar{A}_1)\right] &= 0,\\ \Omega A_2 + \alpha_2 \partial_{y_1}^2 A_2 + \alpha_1 \partial_{y_2}^2 A_2 - \sigma\left[\gamma_1 |A_2|^2 A_2 + \gamma_2(2|A_1|^2 A_2 + A_1^2 \bar{A}_2)\right] &= 0,\end{aligned} \tag{3.4}$$

$$\begin{aligned}\alpha_1 &= 1 + 2\int_{-2\pi}^{2\pi}\int_{-2\pi}^{2\pi} v_1 \partial_{x_1} v_1^{(x_1)} dx, \qquad & \alpha_2 &= 1 + 2\int_{-2\pi}^{2\pi}\int_{-2\pi}^{2\pi} v_1 \partial_{x_2} v_1^{(x_2)} dx,\\ \gamma_1 &= \int_{-2\pi}^{2\pi}\int_{-2\pi}^{2\pi} v_1^4 dx \qquad \text{and} & \gamma_2 &= \int_{-2\pi}^{2\pi}\int_{-2\pi}^{2\pi} v_1^2 v_2^2 dx.\end{aligned}$$

3.1.2 CMEs near the gap edge s_5

At $\omega_* = s_5$ we have $N = 4$. The resonant Bloch waves are $v_1 := u_6((k_c, k_c); x)$, $v_2 := u_6((-k_c, k_c); x)$, $v_3 := u_6((-k_c, -k_c); x)$ and $v_4 := u_6((k_c, -k_c); x)$. Analogously to §3.1.1 the asymptotic expansion needs to be carried out in the space of functions periodic over the common period of $v_1, \ldots, v_4$. The Bloch functions are then pairwise orthogonal over this domain. However, if k_c is not rational then the Bloch waves are not periodic but only quasi-periodic. Therefore, unlike in the case of a separable $V(x)$ [13], where always $k_c \in \{0, 1/2\}$, in the non-separable case in general the derivation in physical variables is impossible.

3.2 CME Derivation in Bloch Variables $\tilde{\phi}(k; x)$

An alternative to the derivation in §3.1 is to transform the problem to Bloch variables. The advantage is that the linear eigenfunctions are then all $2\pi-$periodic in each $x-$coordinate. The orthogonalization domain is, therefore, always $\mathbb{P}^2$.

3.2.1 General Case

The Bloch transform $\mathcal{T}$ is formally defined by

$$\tilde{\phi}(k; x) = (\mathcal{T}\phi)(k; x) = \sum_{m\in\mathbb{Z}^2} e^{\mathrm{i}m\cdot x}\hat{\phi}(k+m), \quad \phi(x) = (\mathcal{T}^{-1}\tilde{\phi})(x) = \int_{\mathbb{T}^2} e^{\mathrm{i}k\cdot x}\tilde{\phi}(k; x)dk, \tag{3.5}$$

where $\hat{\phi}(k)$ denotes the Fourier transform. $\mathcal{T}$ is an isomorphism from $H^s(\mathbb{R}^2, \mathbb{C})$ to $L^2(\mathbb{T}^2, H^s(\mathbb{P}^2, \mathbb{C}))$, $\|\tilde{\phi}\|^2_{L^2(\mathbb{T}^2, H^s(\mathbb{P}^2,\mathbb{C}))} = \int_{\mathbb{T}^2} \|\tilde{\phi}(k; \cdot)\|^2_{H^s(\mathbb{P}^2)} dk$, cf., e.g.,[31], and by construction we have

$$\tilde{\phi}(k; (x_1 + 2\pi, x_2)) = \tilde{\phi}(k; (x_1, x_2 + 2\pi)) = \tilde{\phi}(k; x), \tag{3.6}$$

$$\tilde{\phi}((k_1 + 1, k_2); x) = e^{-\mathrm{i}x_1}\tilde{\phi}(k; x), \quad \tilde{\phi}((k_1, k_2 + 1); x) = e^{-\mathrm{i}x_2}\tilde{\phi}(k; x). \tag{3.7}$$

Multiplication in physical space corresponds to convolution in Bloch space, i.e.,

$$(\mathcal{T}(\phi\psi))(k;x) = \int_{\mathbb{T}^2} \tilde{\phi}(k-l;x)\tilde{\psi}(l;x)dl =: (\tilde{\phi} *_B \tilde{\psi})(k;x), \tag{3.8}$$

where (3.7) is used if $k-l \notin \mathbb{T}^2$. However, if g is $2\pi-$periodic in both x_1 and x_2, then $(\mathcal{T}(gu))(k;x) = g(x)(\mathcal{T}u)(k;x)$.

In order to choose a suitable asymptotic ansatz for $\tilde{\phi}(k;x)$, note first that the Bloch transform $\mathcal{T}$ of the ansatz (3.2) for $\varepsilon\phi^{(0)}(x)$ is

$$\varepsilon\tilde{\phi}^{(0)}(k;x) = \frac{1}{\varepsilon}\sum_{j=1}^{N} p_{n_j}(k^{(j)};x) \sum_{m\in\mathbb{Z}^2} \hat{A}_j\left(\frac{k-k^{(j)}+m}{\varepsilon}\right)e^{\mathrm{i}m\cdot x} \tag{3.9}$$

with $k \in \mathbb{T}^2, x \in \mathbb{P}^2$. As $\hat{A}_j(p)$ is localized near $p=0$, we approximate $\hat{A}_j\left(\frac{k-k^{(j)}+m}{\varepsilon}\right)$ by $\chi_{D_j}(k+m)\hat{A}_j\left(\frac{k-k^{(j)}+m}{\varepsilon}\right)$, where $\chi_{D_j}(k)$ is the characteristic function of the set

$$D_j = \{k \in \mathbb{R}^2 : |k-k^{(j)}| < \varepsilon^r\} \tag{3.10}$$

and

$$0 < r < \frac{2}{3}. \tag{3.11}$$

The reason for (3.11) will be explained in §4.2.

Below we will also use periodically wrapped versions $\tilde{D}_j$ of these neighborhoods, i.e.

$$\tilde{D}_j := \{k \in \mathbb{T}^2 : |k-k^{(j)}| < \varepsilon^r \text{ modulo } \doteq\} \tag{3.12}$$

where 'modulo $\doteq$' means equal modulo 1 in each component, see Fig. 6 for an example.

Note that $k+m \in D_j$ with $k \in \mathbb{T}^2$ is possible only for $m \in \{m \in \mathbb{Z}^2 : 0 \le m_1, m_2 \le 1\}$. We define the set of $m-$values for which $k+m \in D_j$ for some $k \in \mathbb{T}^2$ by $M_j := \{m \in \mathbb{Z}^2 : k+m \in D_j \text{ for some } k \in \mathbb{T}^2\}$. In fact, for small ε only the following cases occur: $M_j = \{\left(\begin{smallmatrix}0\\0\end{smallmatrix}\right), \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)\}$ if $k_1^{(j)} = 1/2$ and $k^{(j)} \neq (1/2,1/2)$, $M_j = \{\left(\begin{smallmatrix}0\\0\end{smallmatrix}\right), \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)\}$ if $k_2^{(j)} = 1/2$ and $k^{(j)} \neq (1/2,1/2)$, $M_j = \{\left(\begin{smallmatrix}0\\0\end{smallmatrix}\right), \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right), \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right), \left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)\}$ if $k^{(j)} = (1/2,1/2)$, and $M_j = \{\left(\begin{smallmatrix}0\\0\end{smallmatrix}\right)\}$ if $k^{(j)} \in \mathrm{int}(\mathbb{T}^2)$.

Thus we are lead to the following asymptotic ansatz in Bloch variables

$$\begin{aligned}
\tilde{\phi}(k;x) &= \frac{1}{\varepsilon}\tilde{\psi}^{(0)}(k;x) + \tilde{\psi}^{(1)}(k;x) + \varepsilon\tilde{\psi}^{(2)}(k;x) + \mathcal{O}(\varepsilon^2),\\
\tilde{\psi}^{(0)}(k;x) &= \sum_{j=1}^{N} p_{n_j}(k^{(j)};x) \sum_{m\in M_j} \chi_{D_j}(k+m)\hat{A}_j\left(\frac{k-k^{(j)}+m}{\varepsilon}\right)e^{\mathrm{i}m\cdot x},\\
\omega &= \omega_* + \Omega\varepsilon^2, \qquad 0 < \varepsilon \ll 1.
\end{aligned} \tag{3.13}$$

The periodic part p_{n_j} of the Bloch functions u_{n_j} is normalized so that $\|p_{n_j}(k;\cdot)\|_{L^2(\mathbb{P}^2)} = 1$.

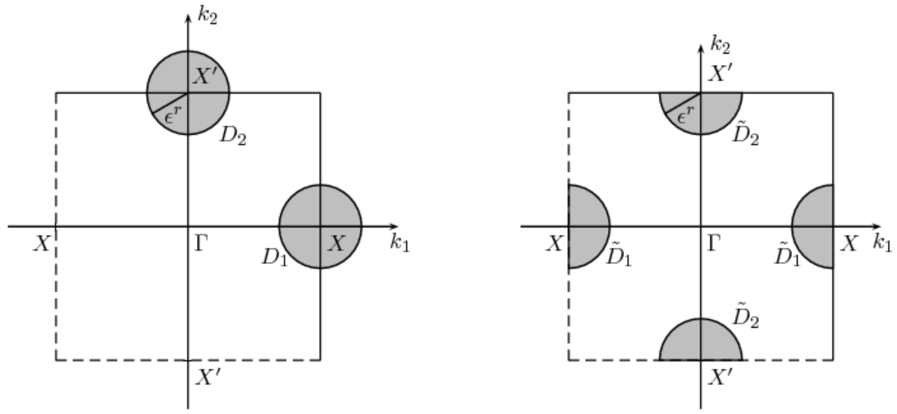

Figure 6: Sets D_j and $\tilde{D}_j$ for $k^{(j)} = X$ and $k^{(j)} = X'$ (as in the example (1.3) with $\eta = 5.35$ at $\omega_* = s_3$).

The difference between the leading order terms in (3.9) and in (3.13) is

$$\frac{1}{\varepsilon}\tilde{\psi}^{(0)}(k;x) - \varepsilon\tilde{\phi}^{(0)}(k;x) =: \sum_{j=1}^{N} \tilde{h}_j(k;x) \tag{3.14}$$

with

$$\begin{aligned} \tilde{h}_j(k;x) = \frac{1}{\varepsilon} p_{n_j}(k^{(j)};x) \Bigg[& \sum_{m\in M_j} \left(1 - \chi_{D_j}(k+m)\right) \hat{A}_j\left(\tfrac{k-k^{(j)}+m}{\varepsilon}\right) e^{\mathrm{i}m\cdot x} \\ & + \sum_{m\in\mathbb{Z}^2\setminus M_j} \hat{A}_j\left(\tfrac{k-k^{(j)}+m}{\varepsilon}\right) e^{\mathrm{i}m\cdot x} \Bigg]. \end{aligned} \tag{3.15}$$

We now estimate $\|\tilde{h}_j\|_{L^2(\mathbb{T}^2,H^s(\mathbb{P}^2))}$. In the first sum in (3.15) we have $|k+m-k^{(j)}| \geq \varepsilon^r$ while in the second sum $|k+m-k^{(j)}| \geq 1$ because $k+m \notin D_j$ for all $k \in \mathbb{T}^2$ if $m \in \mathbb{Z}^2 \setminus M_j$. By the triangle inequality and the substitution $p = (k-k^{(j)}+m)/\varepsilon$ we obtain

$$\begin{aligned} \|\tilde{h}_j\|^2_{L^2(\mathbb{T}^2,H^s(\mathbb{P}^2))} \leq & \sum_{m\in M_j} \|p_{n_j}(k^{(j)};\cdot)e^{\mathrm{i}m\cdot\cdot}\|^2_{H^s(\mathbb{P}^2)} \int_{\substack{|p|>\varepsilon^{r-1} \\ p\in(\mathbb{T}^2-k^{(j)}+m)/\varepsilon}} |\hat{A}_j(p)|^2 \,\mathrm{d}p \\ & + \sum_{m\in\mathbb{Z}^2\setminus M_j} \|p_{n_j}(k^{(j)};\cdot)e^{\mathrm{i}m\cdot\cdot}\|^2_{H^s(\mathbb{P}^2)} \int_{\substack{|p|>c\varepsilon^{-1} \\ p\in(\mathbb{T}^2-k^{(j)}+m)/\varepsilon}} |\hat{A}_j(p)|^2 \,\mathrm{d}p \\ \leq & C\left[\int_{|p|>\varepsilon^{r-1}} |\hat{A}_j(p)|^2 \,\mathrm{d}p + \int_{|p|>\varepsilon^{-1}} |\hat{A}_j(p)|^2 \,\mathrm{d}p\right], \end{aligned}$$

where the H^s regularity of $p_{n_j}(k^{(j)};\cdot)$ is guaranteed if $V \in H^{s-2}_{\mathrm{loc}}(\mathbb{R}^2)$. By rewriting the right hand side as $C\left[\int_{|p|>\varepsilon^{r-1}} |\hat{A}_j(p)|^2 \frac{(1+|p|)^{2s}}{(1+|p|)^{2s}} \,\mathrm{d}p + \int_{|p|>\varepsilon^{-1}} |\hat{A}_j(p)|^2 \frac{(1+|p|)^{2s}}{(1+|p|)^{2s}} \,\mathrm{d}p\right]$ and taking the supremum of $(1+|p|)^{-2s}$ out of the integrals, we have

$$\|\tilde{h}_j\|_{L^2(\mathbb{T}^2,H^s(\mathbb{P}^2))} \leq C(\varepsilon^{s(1-r)} + \varepsilon^s)\|\hat{A}_j\|_{L^2_s(\mathbb{R}^2)} \leq C\varepsilon^{s(1-r)}\|\hat{A}_j\|_{L^2_s(\mathbb{R}^2)}. \tag{3.16}$$

For $r < 1$ we thus have that $\varepsilon^{-1}\psi^{(0)}(x)$ approximates $\varepsilon\phi^{(0)}(x)$ up to $\mathcal{O}(\varepsilon^{s(1-r)})$ in the $H^s(\mathbb{R}^2)$ norm. Because $\|\varepsilon\phi^{(0)}\|_{H^s(\mathbb{R}^2)} = \mathcal{O}(1)$, this approximation is satisfactory.

Applying next $\mathcal{T}$ to (1.2) yields

$$\left[\tilde{L} - \omega\right]\tilde{\phi} + \sigma\,\tilde{\phi} *_B \bar{\tilde{\phi}} *_B \tilde{\phi} = 0, \tag{3.17}$$

on $(k;x) \in \mathbb{T}^2 \times \mathbb{P}^2$, where we recall from (2.3) that $\tilde{L}(k;x) = (\mathrm{i}\partial_{x_1} - k_1)^2 + (\mathrm{i}\partial_{x_2} - k_2)^2 + V(x)$.

Setting $p^{(j,m)} := \frac{k+m-k^{(j)}}{\varepsilon}$, we have

$$\begin{aligned}\tilde{L}(k;x) &= \tilde{L}(k^{(j)} - m + \varepsilon p^{(j,m)};x)\\ &= \tilde{L}(k^{(j)} - m;x) - 2\varepsilon\left[(\mathrm{i}\partial_{x_1} - k_1^{(j)} + m_1)p_1^{(j,m)} + (\mathrm{i}\partial_{x_2} - k_2^{(j)} + m_2)p_2^{(j,m)}\right] + \varepsilon^2\left[p_1^{(j,m)^2} + p_2^{(j,m)^2}\right].\end{aligned} \tag{3.18}$$

Substituting (3.13) in (3.17) and using (3.18), we obtain a hierarchy of equations on $x \in \mathbb{P}^2, k \in \mathbb{T}^2$ such that $k + m \in D_j, j \in \{1, \ldots, N\}$. Note that the combination of $k \in \mathbb{T}^2$ and $k + m \in D_j$ implies $m \in M_j$. The following hierarchy is thus for each $(j,m) \in \{1, \ldots, N\} \times M_j$.

$$\mathcal{O}(\varepsilon^{-\mathbf{1}}): \quad \hat{A}_j(p^{(j,m)})\left[\tilde{L}(k^{(j)} - m;x) - \omega_*\right](p_{n_j}(k^{(j)};x)e^{\mathrm{i}m\cdot x}) = 0,$$

which is equivalent to $\hat{A}_j(p^{(j,m)})e^{\mathrm{i}m\cdot x}\left[\tilde{L}(k^{(j)};x) - \omega_*\right]p_{n_j}(k^{(j)};x) = 0$ and thus holds by definition of $\omega_* = \omega_{n_j}(k^{(j)})$.

$$\begin{aligned}\mathcal{O}(\mathbf{1}): \quad &\left[\tilde{L}(k^{(j)} - m;x) - \omega_*\right]\tilde{\psi}^{(1)}(k;x)\\ &= 2\hat{A}_j(p^{(j,m)})\left[p_1^{(j,m)}(\mathrm{i}\partial_{x_1} - k_1^{(j)} + m_1) + p_2^{(j,m)}(\mathrm{i}\partial_{x_2} - k_2^{(j)} + m_2)\right](p_{n_j}(k^{(j)};x)e^{\mathrm{i}m\cdot x})\\ &= 2\hat{A}_j(p^{(j,m)})e^{\mathrm{i}m\cdot x}\left[p_1^{(j,m)}(\mathrm{i}\partial_{x_1} - k_1^{(j)}) + p_2^{(j,m)}(\mathrm{i}\partial_{x_2} - k_2^{(j)})\right]p_{n_j}(k^{(j)};x)\end{aligned}$$

for $k+m \in D_j, j \in \{1, \ldots, N\}$. To solve this, we note that by differentiating $[\tilde{L}(k;x) - \omega_{n_j}(k)]p_{n_j}(k;x) = 0$ with respect to $k_l, l \in \{1,2\}$ and evaluating at $k = k^{(j)} - m$, we obtain

$$\left[\tilde{L}(k^{(j)} - m;x) - \omega_*\right]\partial_{k_l}p_{n_j}(k^{(j)} - m;x) = 2(\mathrm{i}\partial_{x_l} - k_l^{(j)} + m_l)p_{n_j}(k^{(j)} - m;x). \tag{3.19}$$

Since $p_n(k - m;x) = e^{\mathrm{i}m\cdot x}p_n(k;x)$ due to (3.7), we get for $k + m \in D_j$

$$\tilde{\psi}^{(1)}(k;x) = \sum_{l=1}^{2} p_l^{(j,m)}\hat{A}_j(p^{(j,m)})e^{\mathrm{i}m\cdot x}\partial_{k_l}p_{n_j}(k^{(j)};x). \tag{3.20}$$

$\mathcal{O}(\varepsilon)$: We have

$$
\begin{aligned}
&\left[\tilde{L}(k^{(j)}-m;x)-\omega_*\right]\tilde{\psi}^{(2)}(k;x)\\
&=\Omega\hat{A}_j(p^{(j,m)})p_{n_j}(k^{(j)};x)e^{\mathrm{i}m\cdot x}+2\left[p_1^{(j,m)}(\mathrm{i}\partial_{x_1}-k_1^{(j)}+m_1)+p_2^{(j,m)}(\mathrm{i}\partial_{x_2}-k_2^{(j)}+m_2)\right]\tilde{\psi}^{(1)}(k;x)\\
&\quad-\left(p_1^{(j,m)^2}+p_2^{(j,m)^2}\right)\hat{A}_j(p^{(j,m)})p_{n_j}(k^{(j)};x)e^{\mathrm{i}m\cdot x}-\frac{\sigma}{\varepsilon^4}\chi_{D_j}(k+m)(\tilde{\psi}^{(0)}*_B\tilde{\psi}^{(0)}*_B\tilde{\bar{\psi}}^{(0)})(k;x)\\
&=\Omega\hat{A}_j(p^{(j,m)})p_{n_j}(k^{(j)};x)e^{\mathrm{i}m\cdot x}\\
&\quad-e^{\mathrm{i}m\cdot x}\sum_{l=1}^{2}\left[p_{n_j}(k^{(j)};x)-2(\mathrm{i}\partial_{x_l}-k_l^{(j)})\partial_{k_l}p_{n_j}(k^{(j)};x)\right]p_l^{(j,m)^2}\hat{A}_j(p^{(j,m)})\\
&\quad+2e^{\mathrm{i}m\cdot x}\left[(\mathrm{i}\partial_{x_1}-k_1^{(j)})\partial_{k_2}p_{n_j}(k^{(j)};x)+(\mathrm{i}\partial_{x_2}-k_2^{(j)})\partial_{k_1}p_{n_j}(k^{(j)};x)\right]p_1^{(j,m)}p_2^{(j,m)}\hat{A}_j(p^{(j,m)})\\
&\quad-\frac{\sigma}{\varepsilon^4}\chi_{D_j}(k+m)(\tilde{\psi}^{(0)}*_B\tilde{\psi}^{(0)}*_B\tilde{\bar{\psi}}^{(0)})(k;x)
\end{aligned}
\tag{3.21}
$$

using $\tilde{\psi}^{(1)}$ from (3.20).

The nonlinear term has the form

$$
\begin{aligned}
G_j(k;x):=&\frac{\sigma}{\varepsilon^4}\chi_{D_j}(k+m)(\tilde{\psi}^{(0)}*_B\tilde{\psi}^{(0)}*_B\tilde{\bar{\psi}}^{(0)})(k;x)=\frac{\sigma}{\varepsilon^4}\chi_{D_j}(k+m)\Bigg[\sum_{\alpha=1}^{N}\xi_\alpha*_B\xi_\alpha*_B\xi_\alpha^c\\
&+2\sum_{\substack{\alpha,\beta=1\\ \alpha\neq\beta}}^{N}\xi_\alpha*_B\xi_\beta*_B\xi_\alpha^c+\sum_{\substack{\alpha,\beta=1\\ \alpha\neq\beta}}^{N}\xi_\alpha*_B\xi_\alpha*_B\xi_\beta^c+\sum_{\substack{\alpha,\beta,\gamma=1\\ \alpha\neq\beta,\alpha\neq\gamma,\beta\neq\gamma}}^{N}\xi_\alpha*_B\xi_\beta*_B\xi_\gamma^c\Bigg],
\end{aligned}
\tag{3.22}
$$

where $\xi_\alpha=\xi_\alpha(k;x):=p_{n_\alpha}(k^{(\alpha)};x)\sum_{m\in M_\alpha}\chi_{D_\alpha}(k+m)\hat{A}_\alpha\left(\frac{k+m-k^{(\alpha)}}{\varepsilon}\right)e^{\mathrm{i}m\cdot x}$ and $\xi_\alpha^c=\xi_\alpha^c(k;x):=\overline{p_{n_\alpha}}(k^{(\alpha)};x)\sum_{m\in M_\alpha}\chi_{-D_\alpha}(k-m)\hat{\bar{A}}_\alpha\left(\frac{k-m+k^{(\alpha)}}{\varepsilon}\right)e^{-\mathrm{i}m\cdot x}$. The last sum or the three last sums in (3.22) are absent if $N=2$ or $N=1$ respectively. $\xi_\alpha*_B\xi_\beta*_B\xi_\gamma^c$ consists of terms of the type

$$
\begin{aligned}
g_{noq}(k;x)=&e^{\mathrm{i}(n+o-q)\cdot x}p_{n_\alpha}(k^{(\alpha)};x)p_{n_\beta}(k^{(\beta)};x)\overline{p_{n_\gamma}}(k^{(\gamma)};x)\int_{\mathbb{T}^2}\int_{\mathbb{T}^2}\chi_{D_\alpha}(k-r+n)\hat{A}_\alpha\left(\frac{k-r+n-k^{(\alpha)}}{\varepsilon}\right)\times\\
&\times\ \chi_{D_\beta}(r-s+o)\hat{A}_\beta\left(\frac{r-s+o-k^{(\beta)}}{\varepsilon}\right)\chi_{-D_\gamma}(s-q)\hat{\bar{A}}_\gamma\left(\frac{s-q+k^{(\gamma)}}{\varepsilon}\right)\mathrm{d}s\,\mathrm{d}r
\end{aligned}
\tag{3.23}
$$

with $n\in M_\alpha, o\in M_\beta$ and $q\in M_\gamma$. Clearly, the integration domains can be reduced to $r\in D_{2\varepsilon^r}(k^{(\beta)}-k^{(\gamma)}-o+q)$ and $s\in D_{\varepsilon^r}(-k^{(\gamma)}+q)$. The changes of variables $\tilde{s}:=(s+k^{(\gamma)}-q)/\varepsilon$, and $\tilde{r}:=(r-k^{(\beta)}+k^{(\gamma)}+o-q)/\varepsilon$ yield

$$
\begin{aligned}
g_{noq}(k;x)=&\varepsilon^4e^{\mathrm{i}(n+o-q)\cdot x}p_{n_\alpha}(k^{(\alpha)};x)p_{n_\beta}(k^{(\beta)};x)\overline{p_{n_\gamma}}(k^{(\gamma)};x)\times\\
&\int_{D_{2\varepsilon^{r-1}}\cap\frac{\mathbb{T}^2-k^{(\beta)}+k^{(\gamma)}+o-q}{\varepsilon}}\ \int_{D_{\varepsilon^{r-1}}\cap\frac{\mathbb{T}^2+k^{(\gamma)}-q}{\varepsilon}}\chi_{D_{\varepsilon^{r-1}}}\left(\frac{k-(k^{(\alpha)}+k^{(\beta)}-k^{(\gamma)})+n+o-q}{\varepsilon}-\tilde{r}\right)\times\\
&\hat{A}_\alpha\left(\frac{k-(k^{(\alpha)}+k^{(\beta)}-k^{(\gamma)})+n+o-q}{\varepsilon}-\tilde{r}\right)\chi_{D_{\varepsilon^{r-1}}}(\tilde{r}-\tilde{s})\hat{A}_\beta(\tilde{r}-\tilde{s})\chi_{D_{\varepsilon^{r-1}}}(\tilde{s})\hat{\bar{A}}_\gamma(\tilde{s})\,\mathrm{d}\tilde{s}\,\mathrm{d}\tilde{r},
\end{aligned}
\tag{3.24}
$$

where $D_{\varepsilon^{r-1}} = \{p \in \mathbb{R}^2 : |p| < \varepsilon^{r-1}\}$.

Only those combinations of (n, o, q) which produce nonzero values of all the three characteristic functions in (3.23) for some $k, r, s \in \mathbb{T}^2$ are of relevance. Due to $\chi_{-D_\gamma}(s-q)$ we, therefore, require $q - k^{(\gamma)} \in \overline{\mathbb{T}^2} = [-1/2, 1/2]^2$, which ensures that $s - q \in -D_\gamma$ is satisfied by some $s \in \mathbb{T}^2$ for any $\varepsilon > 0$. The first condition is, thus,

$$s_0 := q - k^{(\gamma)} \in \overline{\mathbb{T}^2}. \tag{3.25}$$

Due to $\chi_{D_\beta}(r-s+o)$ we get the condition $s_0 - o + k^{(\beta)} \in \overline{\mathbb{T}^2}$, i.e.,

$$r_0 := s_0 - o + k^{(\beta)} \in \overline{\mathbb{T}^2}. \tag{3.26}$$

Finally, $\chi_{D_\alpha}(k-r+n)$ enforces $r_0 - n + k^{(\alpha)} \in \overline{\mathbb{T}^2}$, i.e.,

$$k_0 := r_0 - n + k^{(\alpha)} \in \overline{\mathbb{T}^2}. \tag{3.27}$$

Statements (3.25), (3.26), and (3.27) form the necessary condition

$$s_0 := q - k^{(\gamma)} \in \overline{\mathbb{T}^2}, \qquad r_0 := s_0 - o + k^{(\beta)} \in \overline{\mathbb{T}^2}, \quad \text{and} \quad k_0 := r_0 - n + k^{(\alpha)} \in \overline{\mathbb{T}^2} \tag{3.28}$$

for (3.23) (and thus (3.24)) not to vanish.

Another condition on (n, o, q) appears due to the factor $\chi_{D_j}(k + m)$ in G_j. From (3.24) it is clear that g_{noq} is supported on $k \in D_{\varepsilon^r}(k^{(\alpha)} + k^{(\beta)} - k^{(\gamma)} - n - o + q)$. The factor $\chi_{D_j}(k+m)$ thus annihilates all terms g_{noq} except those for which

$$k^{(\alpha)} + k^{(\beta)} - k^{(\gamma)} - n - o + q = k^{(j)} - m. \tag{3.29}$$

If (3.29) is satisfied, (3.24) becomes

$$\begin{aligned} g_{noq}(k; x) = \varepsilon^4 e^{\mathrm{i}(n+o-q)\cdot x} p_{n_\alpha}(k^{(\alpha)}; x) p_{n_\beta}(k^{(\beta)}; x) \overline{p_{n_\gamma}}(k^{(\gamma)}; x) \times \\ \int\limits_{D_{2\varepsilon^{r-1}} \cap \frac{\mathbb{T}^2 - k^{(\beta)} + k^{(\gamma)} + o - q}{\varepsilon}} \int\limits_{D_{\varepsilon^{r-1}} \cap \frac{\mathbb{T}^2 + k^{(\gamma)} - q}{\varepsilon}} \chi_{D_{\varepsilon^{r-1}}}\left(\tfrac{k - k^{(j)} + m}{\varepsilon} - \tilde{r}\right) \times \\ \hat{A}_\alpha\left(\tfrac{k - k^{(j)} + m}{\varepsilon} - \tilde{r}\right) \chi_{D_{\varepsilon^{r-1}}}(\tilde{r} - \tilde{s}) \hat{A}_\beta(\tilde{r} - \tilde{s}) \chi_{D_{\varepsilon^{r-1}}}(\tilde{s}) \hat{\bar{A}}_\gamma(\tilde{s}) \, \mathrm{d}\tilde{s} \, \mathrm{d}\tilde{r}. \end{aligned} \tag{3.30}$$

As a result, the term $A_\alpha A_\beta \bar{A}_\gamma$ will enter the $j-$th equation of the coupled mode system provided there exist $n \in M_\alpha, o \in M_\beta$ and $q \in M_\gamma$ such that (3.28) holds and such that (3.29) holds for some $m \in M_j$. Let us denote the set of (n, o, q) that satisfy (3.28) and (3.29) by $\mathcal{A}_{\alpha,\beta,\gamma,j,m}$.

The sum of the terms (3.30) over $(n, o, q) \in \mathcal{A}_{\alpha,\beta,\gamma,j,m}$ yields a double convolution integral over the full discs $\tilde r \in D_{2\varepsilon^{r-1}}$ and $\tilde s \in D_{\varepsilon^{r-1}}$, i.e.,

$$(\xi_\alpha *_B \xi_\beta *_B \xi^c_\gamma)(k;x) = \varepsilon^4 e^{\mathrm{i}(k^{(\alpha)}+k^{(\beta)}-k^{(\gamma)}-k^{(j)}+m)\cdot x} p_{n_\alpha}(k^{(\alpha)};x)p_{n_\beta}(k^{(\beta)};x)\overline{p_{n_\gamma}}(k^{(\gamma)};x)\times$$
$$\int_{D_{2\varepsilon^{r-1}}}\int_{D_{\varepsilon^{r-1}}} \chi_{D_{\varepsilon^{r-1}}}\left(\tfrac{k-k^{(j)}+m}{\varepsilon}-\tilde r\right)\hat A_\alpha\left(\tfrac{k-k^{(j)}+m}{\varepsilon}-\tilde r\right)\chi_{D_{\varepsilon^{r-1}}}(\tilde r-\tilde s)\hat A_\beta(\tilde r-\tilde s)\chi_{D_{\varepsilon^{r-1}}}(\tilde s)\hat{\bar A}_\gamma(\tilde s)\,\mathrm{d}\tilde s\,\mathrm{d}\tilde r, \tag{3.31}$$

where $e^{\mathrm{i}(n+o-q)\cdot x}$ was replaced by $e^{\mathrm{i}(k^{(\alpha)}+k^{(\beta)}-k^{(\gamma)}-k^{(j)}+m)\cdot x}$ due to (3.29).

We return now to equation (3.21) for $\tilde\psi^{(2)}$ on $k \in (D_j - m)\cap \mathbb{T}^2$. Its solvability condition is $L^2(\mathbb{P}^2)$-orthogonality to $\mathrm{Ker}(\tilde L(k^{(j)} - m;x) - \omega_*) = \mathrm{span}\{\cup_l p_{n_l}(k^{(j)};x)e^{\mathrm{i}m\cdot x} \text{ s.t. } \omega_{n_l}(k^{(j)}) = \omega_*\}$. Clearly, the dimension of the kernel is at most N. The value N is attained if $k^{(1)} = \ldots = k^{(N)}$.

In the linear terms in (3.21) the factor $e^{\mathrm{i}m\cdot x}$ is canceled in the inner product with $p_{n_l}(k^{(j)};x)e^{\mathrm{i}m\cdot x}$ so that the same solvability condition holds for all m. The range of $p^{(j,m)}$ is a different section of the disc $D_{\varepsilon^{r-1}}$ for each m. The section is an $(1/|M_j|)$-th of the full disc so that these $|M_j|$ conditions build one equation in $p \in D_{\varepsilon^{r-1}}$.

The resulting N equations are CMEs in Fourier variables $p \in D_{\varepsilon^{r-1}}$:

$$\Omega\hat A_j - \left(\frac12\partial^2_{k_1}\omega_{n_j}(k^{(j)})p_1^2 + \frac12\partial^2_{k_2}\omega_{n_j}(k^{(j)})p_2^2 + \partial_{k_1}\partial_{k_2}\omega_{n_j}(k^{(j)})p_1p_2\right)\hat A_j - \hat{\mathcal N}_j = 0, \tag{3.32}$$

$j \in \{1,\ldots,N\}$, where $\hat{\mathcal N}_j(p^{(j,m)}) = \langle G_j(\varepsilon p^{(j,m)} + k^{(j)} - m;\cdot), p_{n_j}(k^{(j)};\cdot)e^{\mathrm{i}m\cdot\cdot}\rangle_{L^2(\mathbb{P}^2)}$.

For sufficiently smooth A_j we can neglect the contribution to $\hat A_j$ from $p \in \mathbb{R}^2 \setminus D_{\varepsilon^{r-1}}$ or, for simplicity, assume that the $\hat A_j$ satisfy (3.32) also there. Equation (3.32) is then posed on $p \in \mathbb{R}^2$. Performing the inverse Fourier transform yields the CMEs

$$\Omega A_j + \left(\frac12\partial^2_{k_1}\omega_{n_j}(k^{(j)})\partial^2_{y_1} + \frac12\partial^2_{k_2}\omega_{n_j}(k^{(j)})\partial^2_{y_2} + \partial_{k_1}\partial_{k_2}\omega_{n_j}(k^{(j)})\partial_{y_1}\partial_{y_2}\right)A_j - \mathcal N_j = 0. \tag{3.33}$$

The structure and coefficients in $\mathcal N_j$ for our example (1.3) will be discussed in §3.2.2.

In order to make the discussion of the asymptotic hierarchy complete, we need to mention the part of the $k-$domain outside the neighborhoods of $k^{(j)}$. For $k \in \mathbb{T}^2$ such that $k+m \in \mathbb{T}^2 \setminus D_j$ for all $m \in M_j$ we have $\big[\tilde L(k^{(j)} - m;x) - \omega_*\big]\tilde\psi^{(n)}(k;x) = 0$ for $n = 1,2$ so that $\tilde\psi^{(0)}(k;\cdot) \equiv \tilde\psi^{(1)}(k;\cdot) \equiv 0$ for such k.

The appearance of second derivatives of the bands ω_{n_j} in (3.32) is due to the following

Lemma 3.3 *For any $l, m \in \{1,2\}$*

$$\partial_{k_l}\partial_{k_m}\omega_{n_j}(k^{(j)}) = 2\delta_{lm} - 2\langle(\mathrm{i}\partial_{x_m} - k^{(j)}_m)\partial_{k_l}p_{n_j}(k^{(j)};\cdot) + (\mathrm{i}\partial_{x_l} - k^{(j)}_l)\partial_{k_m}p_{n_j}(k^{(j)};\cdot), p_{n_j}(k^{(j)};\cdot)\rangle_{L^2(\mathbb{P}^2)},$$

where δ_{lm} is the Kronecker delta.

Proof. This follows from differentiation of $[\tilde L(k;x) - \omega_{n_j}(k)]p_{n_j}(k;x) = 0$ w.r.t. k. □

As the next lemma shows, for even potentials $V(x)$ the mixed derivatives of ω_{n_j} are zero whenever $\omega_{n_j}(k^{(j)})$ has geometric multiplicity one and the extremal point $k^{(j)}$ coincides with one of the vertices of the first irreducible Brillouin zone or of its reflection.

Lemma 3.4 *Suppose $V(x)$ is even in x_1 as well as in x_2. Then $\partial_{k_1}\partial_{k_2}\omega_{n_j}(k^{(j)}) = \partial_{k_2}\partial_{k_1}\omega_{n_j}(k^{(j)}) = 0$ if $k^{(j)} \in \Sigma = \{\Gamma, X, X', M\}$ and provided $\omega_{n_j}(k^{(j)})$ has geometric multiplicity 1 as an eigenvalue of* (2.1).

Proof. Take $l, m \in \{1,2\}, l \neq m$. As $\mathrm{i}\partial_{x_m} - k_m^{(j)}$ is self-adjoint, we have

$$\langle (\mathrm{i}\partial_{x_m} - k_m^{(j)})\partial_{k_l} p_{n_j}(k^{(j)};\cdot), p_{n_j}(k^{(j)};\cdot)\rangle_{L^2(\mathbb{P}^2)} = \langle \partial_{k_l} p_{n_j}(k^{(j)};\cdot), (\mathrm{i}\partial_{x_m} - k_m^{(j)}) p_{n_j}(k^{(j)};\cdot)\rangle_{L^2(\mathbb{P}^2)}. \quad (3.34)$$

Based on (2.2) we have $(\mathrm{i}\partial_{x_m} - k_m^{(j)}) p_{n_j}(k^{(j)};x) = -\mathrm{i}\mathrm{e}^{-\mathrm{i}k^{(j)}\cdot x}\partial_{x_m} u_{n_j}(k^{(j)};x)$. Next, $\partial_{k_l} p_{n_j}(k^{(j)};x) = \mathrm{i}\mathrm{e}^{-\mathrm{i}k^{(j)}\cdot x} v_{n_j}^{(x_l)}(k^{(j)};x)$, where $v_{n_j}^{(x_l)}(k^{(j)};x)$ is the generalized Bloch function [29] solving

$$[L-\omega_*]u = 2\partial_{x_l} u_{n_j}(k^{(j)};x), \qquad u(2\pi, x_2) = \mathrm{e}^{\mathrm{i}2\pi k_1^{(j)}} u(0,x_2), \quad u(x_1, 2\pi) = \mathrm{e}^{\mathrm{i}2\pi k_2^{(j)}} u(x_1, 0), \quad (3.35)$$

analogously to (3.3). The inner product in (3.34) thus becomes $\langle -v_{n_j}^{(x_l)}(k^{(j)};\cdot), \partial_{x_m} u_{n_j}(k^{(j)};\cdot)\rangle_{L^2(\mathbb{P}^2)}$. Because $k^{(j)} \in \Sigma$, u_{n_j} is even or odd in x_l (Lemma 2.2). From (3.35) it is clear that $v_{n_j}^{(x_l)}(k^{(j)};x)$ has the opposite symmetry (odd or even respectively) in x_l. Thus, the integrand is odd in x_l and the integral vanishes upon shifting the integration domain to $[-\pi,\pi]^2$. □

3.2.2 CMEs for the Example (1.3)

We now calculate the explicit form of the CMEs (3.33) in the vicinity of the five gap edges in the example (1.3) with $\eta = 5.35$. It turns out that only few terms are nonzero in the nonlinearity $\mathcal{N}_j$ for this case. Of special importance is the edge $\omega_* = s_5$, where $k^{(j)} \notin \Sigma$ and, indeed, $\partial_{k_1}\partial_{k_2}\omega_{n_j}(k^{(j)}) \neq 0$.

In order to numerically evaluate the coefficients $\partial_{k_l}\partial_{k_m}\omega_{n_j}(k^{(j)})$ given in Lemma 3.3, the functions $\partial_{k_l} p_{n_j}(k^{(j)};x)$ have to be computed. They are solutions of the singular system (3.19) but as the right-hand side is orthogonal to the kernel of $\tilde{L}(k^{(j)};x) - \omega_*$, the BiCG algorithm can be used as long as the initial guess is orthogonal to the kernel. We work in a 4th order finite difference discretization and use an incomplete LU preconditioning for BiCG.

CMEs near $\omega_* = s_1$: Only one extremum defines the edge $\omega_* = s_1$, namely the minimum of the band ω_1 at $k = \Gamma$. Therefore, $N = 1, n_1 = 1$ and $k^{(1)} = \Gamma$. Because $k^{(1)} \in \mathrm{int}(\mathbb{T}^2)$, we get $M_1 = \{(0,0)^T\}$. Thus

$$\left[\Omega + \alpha(\partial_{y_1}^2 + \partial_{y_2}^2)\right] A - \sigma\gamma |A|^2 A = 0, \quad (3.36)$$

where $\alpha = \frac{1}{2}\partial_{k_1}^2\omega_1(\Gamma) = \frac{1}{2}\partial_{k_2}^2\omega_1(\Gamma)$ and $\gamma = \langle p_1(\Gamma;\cdot)^2, p_1(\Gamma;\cdot)^2\rangle_{L^2(\mathbb{P}^2)} = \|p_1(\Gamma;\cdot)\|_{L^4(\mathbb{P}^2)}^4$. The identity in α holds due to (2.6). The numerically obtained values are $\alpha \approx 0.62272$ and $\gamma \approx 0.048029$.

CMEs near $\omega_* = s_2$: Here the linear problem is characterized by $N = 1, n_1 = 1$ and $k^{(1)} = M = (1/2, 1/2)$ and we get $M_1 = \{(0,0)^T, (1,0)^T, (0,1)^T, (1,1)^T\}$. The resulting CMEs have the form (3.36). We determine next the coefficient of the nonlinearity $|A|^2 A$. In (3.24) we have $\alpha = \beta = \gamma = 1$ and $k^{(\alpha)} + k^{(\beta)} - k^{(\gamma)} - k^{(j)} = (0,0)^T$. We carry out a straightforward sweep through all the possible combinations (n, o, q, m) (performed using a Matlab script) to determine those that satisfy (3.28) and (3.29). As a result we have

- $m = M_1(:,1) = \left(\begin{smallmatrix}0\\0\end{smallmatrix}\right)$: $(n,o,q)^T \in \left\{ \begin{pmatrix} 0&0\\0&0\\0&0 \end{pmatrix}, \begin{pmatrix} 0&0\\1&0\\1&0 \end{pmatrix}, \begin{pmatrix} 0&0\\0&1\\0&1 \end{pmatrix}, \begin{pmatrix} 0&0\\1&1\\1&1 \end{pmatrix} \right\}$
- $m = M_1(:,2) = \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$: $(n,o,q)^T \in \left\{ \begin{pmatrix} 1&0\\0&0\\0&0 \end{pmatrix}, \begin{pmatrix} 1&0\\1&0\\1&0 \end{pmatrix}, \begin{pmatrix} 1&0\\0&1\\0&1 \end{pmatrix}, \begin{pmatrix} 1&0\\1&1\\1&1 \end{pmatrix} \right\}$
- $m = M_1(:,3) = \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$: $(n,o,q)^T \in \left\{ \begin{pmatrix} 0&1\\0&0\\0&0 \end{pmatrix}, \begin{pmatrix} 0&1\\1&0\\1&0 \end{pmatrix}, \begin{pmatrix} 0&1\\0&1\\0&1 \end{pmatrix}, \begin{pmatrix} 0&1\\1&1\\1&1 \end{pmatrix} \right\}$
- $m = M_1(:,4) = \left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)$: $(n,o,q)^T \in \left\{ \begin{pmatrix} 1&1\\0&0\\0&0 \end{pmatrix}, \begin{pmatrix} 1&1\\1&0\\1&0 \end{pmatrix}, \begin{pmatrix} 1&1\\0&1\\0&1 \end{pmatrix}, \begin{pmatrix} 1&1\\1&1\\1&1 \end{pmatrix} \right\}$,

where we have used $M_j(:,l)$ to denote the $l-$th vector in M_j.

The CME coefficients are thus $\alpha = \frac{1}{2}\partial_{k_1}^2 \omega_1(M) = \frac{1}{2}\partial_{k_2}^2 \omega_1(M)$ and $\gamma = \langle p_1(M;\cdot)^2, p_1(M;\cdot)^2 \rangle_{L^2(\mathbb{P}^2)} = \|p_1(M;\cdot)\|_{L^4(\mathbb{P}^2)}^4$. The identity in α holds due to (2.6). Numerically, $\alpha \approx -1.971217$ and $\gamma \approx 0.076442$.

CMEs near $\omega_* = s_3$: Here $N = 2, n_1 = n_2 = 2, k^{(1)} = X$ and $k^{(2)} = X'$. We have thus $M_1 = \{(0,0)^T, (1,0)^T\}$ and $M_2 = \{(0,0)^T, (0,1)^T\}$.

For $\mathcal{N}_j$ we sweep again through all the possible combinations (n,o,q,m) for both $j=1$ and $j=2$. The results are summarized in Table 1.

term in $\mathcal{N}_j$	$\left(\begin{smallmatrix}\alpha\\\beta\\\gamma\end{smallmatrix}\right)$	j	$k^{(\alpha)} + k^{(\beta)} - k^{(\gamma)} - k^{(j)}$	$(n,o,q)^T$ satisfying (3.28) and (3.29): $m = M_j(:,1)$	$m = M_j(:,2)$	coefficient of the term in $\sigma\mathcal{N}_j$
$\lvert A_1\rvert^2 A_1$	$\left(\begin{smallmatrix}1\\1\\1\end{smallmatrix}\right)$	1	$\left(\begin{smallmatrix}0\\0\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&0\\0&0\\0&0\end{smallmatrix}\right), \left(\begin{smallmatrix}0&0\\1&0\\1&0\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}1&0\\0&0\\0&0\end{smallmatrix}\right), \left(\begin{smallmatrix}1&0\\1&0\\1&0\end{smallmatrix}\right)$	$\langle p_2(X,\cdot)^2, p_2(X,\cdot)^2\rangle$
		2	$\left(\begin{smallmatrix}1/2\\-1/2\end{smallmatrix}\right)$	/	/	0
$\lvert A_2\rvert^2 A_2$	$\left(\begin{smallmatrix}2\\2\\2\end{smallmatrix}\right)$	1	$\left(\begin{smallmatrix}-1/2\\1/2\end{smallmatrix}\right)$	/	/	0
		2	$\left(\begin{smallmatrix}0\\0\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&0\\0&0\\0&0\end{smallmatrix}\right), \left(\begin{smallmatrix}0&0\\0&1\\0&1\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&1\\0&0\\0&0\end{smallmatrix}\right), \left(\begin{smallmatrix}0&1\\0&1\\0&1\end{smallmatrix}\right)$	$\langle p_2(X',\cdot)^2, p_2(X',\cdot)^2\rangle$
$\lvert A_1\rvert^2 A_2$	$\left(\begin{smallmatrix}1\\2\\1\end{smallmatrix}\right)$,	1	$\left(\begin{smallmatrix}-1/2\\1/2\end{smallmatrix}\right)$	/	/	0
	$\left(\begin{smallmatrix}2\\1\\1\end{smallmatrix}\right)$	2	$\left(\begin{smallmatrix}0\\0\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&0\\0&0\\0&0\end{smallmatrix}\right), \left(\begin{smallmatrix}1&0\\0&0\\1&0\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&0\\0&1\\0&0\end{smallmatrix}\right), \left(\begin{smallmatrix}1&0\\0&1\\1&0\end{smallmatrix}\right)$	$2\langle \lvert p_2(X,\cdot)\rvert^2, \lvert p_2(X',\cdot)\rvert^2\rangle$
$\lvert A_2\rvert^2 A_1$	$\left(\begin{smallmatrix}1\\2\\2\end{smallmatrix}\right)$,	1	$\left(\begin{smallmatrix}0\\0\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&0\\0&0\\0&0\end{smallmatrix}\right), \left(\begin{smallmatrix}0&1\\0&0\\0&1\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&0\\1&0\\0&0\end{smallmatrix}\right), \left(\begin{smallmatrix}0&1\\1&0\\0&1\end{smallmatrix}\right)$	$2\langle \lvert p_2(X,\cdot)\rvert^2, \lvert p_2(X',\cdot)\rvert^2\rangle$
	$\left(\begin{smallmatrix}2\\1\\2\end{smallmatrix}\right)$	2	$\left(\begin{smallmatrix}1/2\\-1/2\end{smallmatrix}\right)$	/	/	0
$A_1^2 A_2^*$	$\left(\begin{smallmatrix}1\\1\\2\end{smallmatrix}\right)$	1	$\left(\begin{smallmatrix}1/2\\-1/2\end{smallmatrix}\right)$	/	/	0
		2	$\left(\begin{smallmatrix}1\\-1\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&0\\1&0\\0&1\end{smallmatrix}\right), \left(\begin{smallmatrix}1&0\\0&0\\0&1\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&0\\1&0\\0&0\end{smallmatrix}\right), \left(\begin{smallmatrix}1&0\\0&0\\0&0\end{smallmatrix}\right)$	$\langle e^{\mathrm{i}(1,-1)^T\cdot\cdot} p_2(X,\cdot)^2, p_2(X',\cdot)^2\rangle$
$A_2^2 A_1^*$	$\left(\begin{smallmatrix}2\\2\\1\end{smallmatrix}\right)$	1	$\left(\begin{smallmatrix}-1\\1\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&0\\0&1\\1&0\end{smallmatrix}\right), \left(\begin{smallmatrix}0&1\\0&0\\1&0\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}0&0\\0&1\\0&0\end{smallmatrix}\right), \left(\begin{smallmatrix}0&1\\0&0\\0&0\end{smallmatrix}\right)$	$\langle e^{\mathrm{i}(-1,1)^T\cdot\cdot} p_2(X',\cdot)^2, p_2(X,\cdot)^2\rangle$
		2	$\left(\begin{smallmatrix}-1/2\\1/2\end{smallmatrix}\right)$	/	/	0

Table 1: Calculation of the nonlinearity terms for the CME near $\omega_* = s_3$.

The resulting CMEs are

$$\begin{aligned} \left[\Omega + \alpha_1\partial_{y_1}^2 + \alpha_2\partial_{y_2}^2\right] A_1 - \sigma\left[\gamma_1|A_1|^2A_1 + \gamma_2(2|A_2|^2A_1 + A_2^2\bar{A}_1)\right] &= 0, \\ \left[\Omega + \alpha_2\partial_{y_1}^2 + \alpha_1\partial_{y_2}^2\right] A_2 - \sigma\left[\gamma_1|A_2|^2A_2 + \gamma_2(2|A_1|^2A_2 + A_1^2\bar{A}_2)\right] &= 0, \end{aligned} \tag{3.37}$$

where

$$\begin{aligned}
\alpha_1 &= \frac{1}{2}\partial_{k_1}^2\omega_2(X) = \frac{1}{2}\partial_{k_2}^2\omega_2(X'), \quad \alpha_2 = \frac{1}{2}\partial_{k_2}^2\omega_2(X) = \frac{1}{2}\partial_{k_1}^2\omega_2(X'),\\
\gamma_1 &= \langle p_2(X;\cdot)^2, p_2(X;\cdot)^2\rangle_{L^2(\mathbb{P}^2)} = \langle p_2(X';\cdot)^2, p_2(X';\cdot)^2\rangle_{L^2(\mathbb{P}^2)}\\
&= \|p_2(X;\cdot)\|_{L^4(\mathbb{P}^2)}^4 = \|p_2(X';\cdot)\|_{L^4(\mathbb{P}^2)}^4,\\
\gamma_2 &= \langle e^{\mathrm{i}(1,-1)^T\cdot\cdot}p_2(X;\cdot)^2, p_2(X';\cdot)^2\rangle_{L^2(\mathbb{P}^2)} = \langle e^{\mathrm{i}(-1,1)^T\cdot\cdot}p_2(X';\cdot)^2, p_2(X;\cdot)^2\rangle_{L^2(\mathbb{P}^2)}\\
&= \langle |p_2(X;\cdot)|^2, |p_2(X';\cdot)|^2\rangle_{L^2(\mathbb{P}^2)}.
\end{aligned}$$

The identities in α_1, α_2 and γ_1 hold due to (2.6). The equalities in γ_2 yield $\gamma_2 \in \mathbb{R}$ and follow from the fact that $u_2(X,x) = e^{\mathrm{i}x_1/2}p_2(X;x)$ and $u_2(X',x) = e^{\mathrm{i}x_2/2}p_2(X';x)$ are real. In detail

$$\begin{aligned}
&\int_{\mathbb{P}^2} e^{\mathrm{i}x_1}p_2(X;x)^2 e^{-\mathrm{i}x_2}\overline{p_2(X';x)}^2\,\mathrm{d}x = \int_{\mathbb{P}^2} u_2(X;x)^2\overline{u_2(X';x)}^2\,\mathrm{d}x\\
= &\int_{\mathbb{P}^2} u_2(X;x)^2 u_2(X';x)^2\,\mathrm{d}x = \int_{\mathbb{P}^2} u_2(X';x)^2\overline{u_2(X;x)}^2\,\mathrm{d}x = \int_{\mathbb{P}^2} e^{\mathrm{i}x_2}p_2(X';x)^2 e^{-\mathrm{i}x_1}\overline{p_2(X;x)}^2\,\mathrm{d}x.
\end{aligned}$$

The CMEs (3.37) are thus identical to those in Sec. 3.1.1 derived in physical variables. Numerically, $\alpha_1 \approx 2.599391, \alpha_2 \approx 0.040561, \gamma_1 \approx 0.090082,$ and $\gamma_2 \approx 0.003032$.

CMEs near $\omega_* = s_4$: Here $N = 1, n_1 = 5$ and $k^{(1)} = M$. This case is completely analogous to $\omega_* = s_2$. The CMEs are (3.36) with $\alpha = \frac{1}{2}\partial_{k_1}^2\omega_5(M) = \frac{1}{2}\partial_{k_2}^2\omega_5(M) \approx -0.300655$ and $\gamma = \langle p_5(M;\cdot)^2, p_5(M;\cdot)^2\rangle_{L^2(\mathbb{P}^2)} = \|p_5(M;\cdot)\|_{L^4(\mathbb{P}^2)}^4 \approx 0.039755$.

CMEs near $\omega_* = s_5$: Here $N = 4, n_1 = n_2 = n_3 = n_4 = 6, k^{(1)} = (k_c, k_c), k^{(2)} = (-k_c, k_c), k^{(3)} = (-k_c, -k_c)$ and $k^{(4)} = (k_c, -k_c)$, where $k_c \approx 0.439028$. This is an important case in our example because $k^{(j)} \notin \Sigma$ here. Note that because $k^{(j)} \in \mathrm{int}(\mathbb{T}^2)$ for all $j \in \{1,\ldots,N\}$, we have $M_1 = \ldots = M_4 = \{(0,0)^T\}$.

We start with the last two sums of G (see (3.23)). Terms of the type $\xi_l *_B \xi_l *_B \xi_m^c$ (the third sum in G) do not contribute to the CMEs because $2k^{(l)} - k^{(m)}$ is not congruent to any $k^{(j)}, j \in \{1,\ldots,4\}$ for any choice of $l, m \in \{1,\ldots,4\}, l \neq m$. For example, $2k^{(1)} - k^{(2)} = (3k_c, k_c)$, which is not congruent to any $k^{(j)}$ since $k_c \notin \{0, 1/2\}$. Only four terms of the type $\xi_l *_B \xi_m *_B \xi_n^c$ (the last sum in G) contribute to the CMEs, namely $\xi_2 *_B \xi_4 *_B \xi_3^c$ to the equation for $k \in D_1$, $\xi_1 *_B \xi_3 *_B \xi_4^c$ to the equation for $k \in D_2$, $\xi_2 *_B \xi_4 *_B \xi_1^c$ to the equation for $k \in D_3$ and $\xi_1 *_B \xi_3 *_B \xi_2^c$ to the equation for $k \in D_4$. This is because $k^{(2)} + k^{(4)} - k^{(3)} = k^{(1)}, k^{(1)} + k^{(3)} - k^{(4)} = k^{(2)}, k^{(2)} + k^{(4)} - k^{(1)} = k^{(3)}$ and $k^{(1)} + k^{(3)} - k^{(2)} = k^{(4)}$. The other terms in the last sum in G do not contribute. As an example, $k^{(1)} + k^{(2)} - k^{(3)} = (k_c, 3k_c)$.

Another consequence of $k^{(j)} \notin \Sigma$ is that Lemma 3.4 does not apply and mixed derivatives of A_j may appear. The system of CMEs thus becomes

$$
\begin{aligned}
0 &= \left[\Omega + \alpha_1(\partial_{y_1}^2 + \partial_{y_2}^2) + \alpha_2 \partial_{y_1}\partial_{y_2}\right] A_1 \\
&\quad - \sigma\Big[\gamma_1 |A_1|^2 A_1 + 2\big(\gamma_2 \left(|A_2|^2 + |A_4|^2\right) A_1 + \tilde{\gamma}_1 |A_3|^2 A_1 + \tilde{\gamma}_2 A_2 A_4 \bar{A}_3\big)\Big], \\
0 &= \left[\Omega + \alpha_1(\partial_{y_1}^2 + \partial_{y_2}^2) - \alpha_2 \partial_{y_1}\partial_{y_2}\right] A_2 \\
&\quad - \sigma\Big[\gamma_1 |A_2|^2 A_2 + 2\big(\gamma_2 \left(|A_1|^2 + |A_3|^2\right) A_2 + \tilde{\gamma}_1 |A_4|^2 A_2 + \bar{\tilde{\gamma}}_2 A_1 A_3 \bar{A}_4\big)\Big], \\
0 &= \left[\Omega + \alpha_1(\partial_{y_1}^2 + \partial_{y_2}^2) + \alpha_2 \partial_{y_1}\partial_{y_2}\right] A_3 \\
&\quad - \sigma\Big[\gamma_1 |A_3|^2 A_3 + 2\big(\gamma_2 \left(|A_2|^2 + |A_4|^2\right) A_3 + \tilde{\gamma}_1 |A_1|^2 A_3 + \tilde{\gamma}_2 A_2 A_4 \bar{A}_1\big)\Big], \\
0 &= \left[\Omega + \alpha_1(\partial_{y_1}^2 + \partial_{y_2}^2) - \alpha_2 \partial_{y_1}\partial_{y_2}\right] A_4 \\
&\quad - \sigma\Big[\gamma_1 |A_4|^2 A_4 + 2\big(\gamma_2 \left(|A_1|^2 + |A_3|^2\right) A_4 + \tilde{\gamma}_1 |A_2|^2 A_4 + \bar{\tilde{\gamma}}_2 A_1 A_3 \bar{A}_2\big)\Big],
\end{aligned}
\tag{3.38}
$$

where

$$
\begin{aligned}
\alpha_1 = &\ \tfrac{1}{2}\partial_{k_1}^2 \omega_6(k_c((-1)^m, (-1)^n)) = \tfrac{1}{2}\partial_{k_2}^2 \omega_6(k_c((-1)^p, (-1)^q)) \quad \text{for any} \quad m, n, p, q \in \{0, 1\}, \\
\alpha_2 = &\ \partial_{k_1}\partial_{k_2}\omega_6(k_c, k_c) = \partial_{k_1}\partial_{k_2}\omega_6(-k_c, -k_c) = -\partial_{k_1}\partial_{k_2}\omega_6(-k_c, k_c) = -\partial_{k_1}\partial_{k_2}\omega_6(k_c, -k_c), \\
\gamma_1 = &\ \langle |p_6(k_c((-1)^m, (-1)^n); \cdot)|^2, |p_6((k_c((-1)^m, (-1)^n); \cdot)|^2 \rangle_{L^2(\mathbb{P}^2)} = \|p_6((k_c, k_c); \cdot)\|_{L^4(\mathbb{P}^2)}^4 \\
&\ \text{for any } m, n \in \{0, 1\}, \\
\gamma_2 = &\ \langle |p_6((-k_c, k_c); \cdot)|^2, |p_6((k_c, k_c); \cdot)|^2 \rangle_{L^2(\mathbb{P}^2)} = \langle |p_6((k_c, -k_c); \cdot)|^2, |p_6((k_c, k_c); \cdot)|^2 \rangle_{L^2(\mathbb{P}^2)} \\
= &\ \langle |p_6((-k_c, -k_c); \cdot)|^2, |p_6((-k_c, k_c); \cdot)|^2 \rangle_{L^2(\mathbb{P}^2)} = \langle |p_6((k_c, -k_c); \cdot)|^2, |p_6((-k_c, -k_c); \cdot)|^2 \rangle_{L^2(\mathbb{P}^2)}, \\
\tilde{\gamma}_1 = &\ \langle |p_6((-k_c, -k_c); \cdot)|^2, |p_6((k_c, k_c); \cdot)|^2 \rangle_{L^2(\mathbb{P}^2)} = \langle |p_6((k_c, -k_c); \cdot)|^2, |p_6((-k_c, k_c); \cdot)|^2 \rangle_{L^2(\mathbb{P}^2)}, \\
\tilde{\gamma}_2 = &\ \langle p_6((-k_c, k_c); \cdot) p_6((k_c, -k_c); \cdot), p_6((-k_c, -k_c); \cdot) p_6((k_c, k_c); \cdot) \rangle_{L^2(\mathbb{P}^2)}.
\end{aligned}
$$

The identities in α_1, α_2 and γ_1 are due to (2.5) and the identities in γ_2 due to (2.5) and (2.6).

Moreover, $\gamma_1 = \tilde{\gamma}_1$ and $\gamma_2 = \tilde{\gamma}_2$ due to (2.7). This also implies $\tilde{\gamma}_2 = \bar{\tilde{\gamma}}_2$. Using these identities, we arrive at the system

$$
\begin{aligned}
\left[\Omega + \alpha_1(\partial_{y_1}^2 + \partial_{y_2}^2) + \alpha_2 \partial_{y_1}\partial_{y_2}\right] A_1 - \sigma\left[\gamma_1 \left(|A_1|^2 + 2|A_3|^2\right) A_1 + 2\gamma_2 \left((|A_2|^2 + |A_4|^2) A_1 + A_2 A_4 \bar{A}_3\right)\right] &= 0, \\
\left[\Omega + \alpha_1(\partial_{y_1}^2 + \partial_{y_2}^2) - \alpha_2 \partial_{y_1}\partial_{y_2}\right] A_2 - \sigma\left[\gamma_1 \left(|A_2|^2 + 2|A_4|^2\right) A_2 + 2\gamma_2 \left((|A_1|^2 + |A_3|^2) A_2 + A_1 A_3 \bar{A}_4\right)\right] &= 0, \\
\left[\Omega + \alpha_1(\partial_{y_1}^2 + \partial_{y_2}^2) + \alpha_2 \partial_{y_1}\partial_{y_2}\right] A_3 - \sigma\left[\gamma_1 \left(|A_3|^2 + 2|A_1|^2\right) A_3 + 2\gamma_2 \left((|A_2|^2 + |A_4|^2) A_3 + A_2 A_4 \bar{A}_1\right)\right] &= 0, \\
\left[\Omega + \alpha_1(\partial_{y_1}^2 + \partial_{y_2}^2) - \alpha_2 \partial_{y_1}\partial_{y_2}\right] A_4 - \sigma\left[\gamma_1 \left(|A_4|^2 + 2|A_2|^2\right) A_4 + 2\gamma_2 \left((|A_1|^2 + |A_3|^2) A_4 + A_1 A_3 \bar{A}_2\right)\right] &= 0.
\end{aligned}
\tag{3.39}
$$

The numerical values of the coefficients are $\alpha_1 \approx 6.051248, \alpha_2 \approx 0.096394, \gamma_1 \approx 0.039118$ and $\gamma_2 \approx 0.029926$.

4 Justification of the Coupled Mode Equations

If $\omega = \omega_* + \varepsilon^2\Omega$ is in the band gap, then families of solitons, i.e., of smooth exponentially localized solitary wave solutions, are known for many classes of CMEs [33]. However, as already noted in the introduction, the formal derivation of the CMEs in §3, discarding some error at higher order in ε, does not imply that localized solutions of the CMEs yield gap solitons of (1.2). For this we need to estimate the error in some function space and show persistence of the CME solitons under perturbation of the CME including the error. We proceed similarly to [13]. However, as function space we choose $H^s(\mathbb{R}^2)$ with $s > 1$, in contrast to $\mathcal{F}^{-1}L^1_s(\mathbb{R}^2)$ in [13]. The latter is possible in the separable case but there is the technical obstacle of the extension of [13, (3.7)] to the nonseparable case. On one hand, $L^1_s(\mathbb{R}^2)$ in Fourier space gives a direct pointwise estimate in physical space via $\|\phi\|_{C^s} \leq C\|\hat{\phi}\|_{L^1_s}$. On the other hand, working in Hilbert spaces H^s is conceptually simpler since it allows for going back and forth between physical space (for the nonlinearity) and Bloch space (for the linear part). Moreover, localization of the resulting gap solitons follows directly in H^s spaces. By the embedding $\|\phi\|_{C^k} \leq C\|\phi\|_{H^s}$ for $k < s-1$ pointwise estimates are still obtainable although these are typically far from optimal.

4.1 Preliminaries

We have the asymptotic distribution

$$C_1 n \leq \omega_n(k) \leq C_2 n \quad \forall n \in \mathbb{N}, \forall k \in \mathbb{T}^2 \tag{4.1}$$

of bands $\omega_n(k)$, with some constants C_1, C_2 independent of k and n. This follows from the asymptotic "density of states"(see p. 55 of [22]) $N(\lambda;k) = a\lambda + \mathcal{O}(\lambda^{\frac{1}{2}})$ as $\lambda \to \infty$, where $N(\lambda;k)$ is the number of eigenvalues $\omega_n(k)$ smaller than or equal to λ.

We introduce the diagonalization operator

$$\mathcal{D}(k)_{k\in\mathbb{T}^2} : \widetilde{\phi}(k;x) \to \vec{\widetilde{\phi}}(k),\ \widetilde{\phi}_n(k) = \left\langle \widetilde{\phi}(k;\cdot), p_n(k;\cdot) \right\rangle_{L^2(\mathbb{P}^2)}.$$

Based on (4.1) we may estimate $\mathcal{D}$. Similarly to [8, Lemma 3.3] we find that $\mathcal{D}(k)$ for all k is an isomorphism between $H^s(\mathbb{P}^2)$ and ℓ^2_s, $s \geq 0$, where

$$\ell^2_s := \{\vec{\widetilde{\phi}} : \|\vec{\widetilde{\phi}}\|^2_{\ell^2_s} := \sum_{j\in\mathbb{N}} |\widetilde{\phi}_j|^2(1+j)^s < \infty\}.$$

Moreover, $\|\mathcal{D}(k)\|, \|\mathcal{D}(k)^{-1}\| \leq C$ independent of k. Thus, $\tilde{\phi} \mapsto \vec{\tilde{\phi}}$ is an isomorphism between $L^2(\mathbb{T}^2, H^s(\mathbb{P}^2))$ and $L^2(\mathbb{T}^2, \ell^2_s)$ and therefore we have

Lemma 4.1 *For $s \geq 0$ the map $\phi \mapsto \vec{\tilde{\phi}} = \mathcal{D}\mathcal{T}\phi$ is an isomorphism between $H^s(\mathbb{R}^2)$ and $\mathcal{X}^s := L^2(\mathbb{T}^2, \ell^2_s)$, $\|\vec{\tilde{\phi}}\|^2_{\mathcal{X}^s} = \int_{\mathbb{T}^2} \sum_{n\in\mathbb{N}} |\tilde{\phi}_n(k)|^2(1+n)^s\,\mathrm{d}k$, i.e., we have the equivalence of norms*

$$C_1\|\phi\|_{H^s} \leq \|\vec{\tilde{\phi}}\|_{\mathcal{X}^s} \leq C_2\|\phi\|_{H^s}. \tag{4.2}$$

For the sake of convenience, and since the idea of the proof is needed below, we also include a simple version of a Sobolev embedding theorem.

Lemma 4.2 *For $\phi, \psi \in H^s(\mathbb{R}^d)$, $s > d/2$, we have $\phi, \psi \in C^0$, $\|\phi\|_{C^0} \le C\|\phi\|_{H^s}$, and $\|\phi\psi\|_{H^s} \le C\|\phi\|_{H^s}\|\psi\|_{H^s}$.*

Proof. Let $w(k) = (1+|k|)^s$. Then $w^{-1} \in L^2$ for $s > d/2$ and thus

$$\|\phi\|_{C^0} \le C\|\hat{\phi}\|_{L^1} = C\|w^{-1}\hat{\phi}w\|_{L^1} \le C\|w^{-1}\|_{L^2}\|\hat{\phi}\|_{L^2_s} \le C\|\hat{\phi}\|_{L^2_s}. \tag{4.3}$$

Next $w(k) \le w(k-\ell) + w(\ell)$ and thus

$$|w(k)(\hat{\phi} * \hat{\psi})(k)| \le \int |w(k-\ell)\hat{\phi}(k-\ell)\hat{\psi}(\ell)|\,\mathrm{d}\ell + \int |w(\ell)\hat{\phi}(k-\ell)\hat{\psi}(\ell)|\,\mathrm{d}\ell. \tag{4.4}$$

Therefore, using Young's inequality $\|\hat{\phi} * \hat{\psi}\|_{L^r} \le C\|\hat{\phi}\|_{L^p}\|\hat{\psi}\|_{L^q}$, $1/p + 1/q = 1/r + 1$, with $r = 2, p = 2, q = 1$ we have

$$\begin{aligned}\|\phi\psi\|_{H^s} &\le C\|\hat{\phi} * \hat{\psi}\|_{L^2_s} \le C\left(\||w\hat{\phi}| * |\hat{\psi}|\|_{L^2} + \||\hat{\phi}| * |w\hat{\psi}|\|_{L^2}\right)\\ &\le C\left(\|w\hat{\phi}\|_{L^2}\|\hat{\psi}\|_{L^1} + \|w\hat{\psi}\|_{L^2}\|\hat{\phi}\|_{L^1}\right) \le 2C\|\hat{\phi}\|_{L^2_s}\|\hat{\psi}\|_{L^2_s} \le C\|\phi\|_{H^s}\|\psi\|_{H^s}.\end{aligned} \tag{4.5}$$

□

By Lebesgue dominated convergence, $\phi \in H^s$ with $s > 1$ also implies $\phi(x) \to 0$ pointwise as $|x| \to \infty$, which shows that solutions in H^s are indeed localized. Moreover, using the product rule, Lemma 4.2 can be generalized to $\|\phi\|_{C^k} \le C\|\phi\|_{H^s}$, $k \in \mathbb{N}$, $k < s - d/2$. Finally, by Lemma 4.1 these statements directly transfer to $\mathcal{X}^s$.

Lemma 4.3 *Let $s > 1$ and $\vec{\tilde{\phi}}, \vec{\tilde{\psi}} \in \mathcal{X}^s$. Then $\overrightarrow{\widetilde{\phi\psi}} \in \mathcal{X}^s$, $\phi \in C_b^k(\mathbb{R}^2)$ for $k < s-1$, and $\phi(x) \to 0$ as $|x| \to \infty$.*

4.2 Justification Step I: The extended Coupled Mode Equations

To justify the general stationary CMEs (3.32) as an asymptotic model for stationary gap solitons near an edge $\omega = \omega_*$, we again consider (3.17), i.e.

$$\left[\tilde{L}(k;x) - \omega_* - \varepsilon^2\Omega\right]\tilde{\phi}(k;x) = -\sigma\,\left(\tilde{\phi} *_B \tilde{\bar{\phi}} *_B \tilde{\phi}\right)(k;x). \tag{4.6}$$

In contrast to the formal derivation of the CME in §3 we now want to keep track of higher order remainders. We first apply the diagonalization operator $\mathcal{D} : \tilde{\phi}(k;x) \to \vec{\tilde{\phi}}(k)$, $\tilde{\phi}_n(k) = \langle\tilde{\phi}(k;\cdot), p_n(k,\cdot)\rangle_{L^2(\mathbb{P}^2)}$ to get

$$\left[\omega_n(k) - \omega_* - \varepsilon^2\Omega\right]\tilde{\phi}_n(k) = -\sigma\tilde{g}_n(k) \tag{4.7}$$

with $\tilde{g}_n(k) = \langle(\tilde{\phi} *_B \tilde{\bar{\phi}} *_B \tilde{\phi})(k;\cdot), p_n(k;\cdot)\rangle_{L^2(\mathbb{P}^2)}$.

Lemma 4.3 allows us to consider (4.7) in the space $\mathcal{X}^s$, $s > 1$. The multiplication operator $\omega_n(k) - \omega_*$ is not invertible since it vanishes at N points $(n,k) = (n_1, k^{(1)}), \ldots, (n_N, k^{(N)})$. This dictates our generalized Lyapunov-Schmidt decomposition of the solution $\vec{\tilde{\phi}}$, namely

$$\vec{\tilde{\phi}}(k) = \varepsilon^{-1}\vec{\tilde{\eta}}^{(0)}_{\text{LS}}(k) + \vec{\tilde{\psi}}(k), \text{ where } \varepsilon^{-1}\vec{\tilde{\eta}}^{(0)}_{\text{LS}}(k) = \varepsilon^{-1}\sum_{j=1}^{N} e_{n_j} \sum_{m\in M_j} \hat{B}_j\left(\frac{k+m-k^{(j)}}{\varepsilon}\right) \tag{4.8}$$

with $\operatorname{supp}\hat{B}_j \subset D_{\varepsilon^{r-1}}$, $\widetilde{\psi}_{n_j}(k) = 0$ for $k \in K_c := \cup\{\widetilde{D}_l : l \in \{1,\ldots,N\}, n_l = n_j\}$, where $\widetilde{D}_l$ is defined in (3.12) and where e_{n_j} is the unit vector in the n_j direction in $\mathbb{R}^{\mathbb{N}}$.

Remark 4.4 Note that in general $\varepsilon^{-1}\vec{\tilde{\eta}}^{(0)}_{\text{LS}}(k)$ is not the diagonalization of the leading order term in an ansatz of the form (3.13) (with $\chi_{D_j}(k+m)\hat{A}_j\left(\frac{k+m-k^{(j)}}{\varepsilon}\right)$ replaced by $\hat{B}_j\left(\frac{k+m-k^{(j)}}{\varepsilon}\right)$), except at $k = k^{(j)}$, since in $\varepsilon^{-1}\tilde{\eta}^{(0)}_{\text{LS}}(k;x)$ we have $p_{n_j}(k;\cdot)$ instead of $p_{n_j}(k^{(j)},\cdot)$.

However, we have

$$\varepsilon^{-1}\tilde{\eta}^{(0)}_{\text{LS}}(k,x) = \varepsilon^{-1}\tilde{\psi}^{(0)}_{\text{LS}}(k,x) + \tilde{\rho}(k,x) \tag{4.9}$$

with

$$\varepsilon^{-1}\tilde{\psi}^{(0)}_{\text{LS}}(k,x) = \varepsilon^{-1}\sum_{j=1}^{N} p_{n_j}(k^{(j)},x)\sum_{m\in M_j}\hat{B}_j\left(\frac{k+m-k^{(j)}}{\varepsilon}\right)\mathrm{e}^{\mathrm{i}m\cdot x}, \tag{4.10}$$

and where for $\hat{B}_j \in L^2_s$ with $s \geq 1$

$$\|\tilde{\rho}\|_{L^2(\mathbb{T}^2, H^s(\mathbb{P}^2))} \leq C\varepsilon\sum_{j=1}^{N}\|\hat{B}_j\|_{L^2_s}. \tag{4.11}$$

This follows from writing $k = k^{(j)} - m + (k + m - k^{(j)})$, expanding

$$p_{n_j}(k,x) = p_{n_j}(k^{(j)} - m, x) + \nabla_k p_{n_j}(k^{(j)}_\star, x)\cdot(k + m - k^{(j)})$$

with $k^{(j)}_{\star,l} \in [\min(k^{(j)}_l - m_l, k_l), \max(k^{(j)}_l - m_l, k_l)]$, $l = 1,2$, and using $p_{n_j}(k^{(j)} - m, x) = p_{n_j}(k^{(j)}, x)\mathrm{e}^{\mathrm{i}m\cdot x}$, which yields

$$\tilde{\rho}(k,x) = \varepsilon^{-1}\sum_{j=1}^{N}\sum_{m\in M_j}\hat{B}_j\left(\tfrac{k+m-k^{(j)}}{\varepsilon}\right)(k+m-k^{(j)})\cdot\nabla_k p_{n_j}(k^{(j)}_\star, x).$$

To prove (4.11), we may fix some (of the finitely many) j, m. Since $\omega_{n_j}(k)$ are simple for $k \in \tilde{D}_j$, we have $\sup_{k\in\tilde{D}_j}\|\nabla_k p_{n_j}(k,\cdot)\|_{H^s(\mathbb{P}^2)} \leq C$, and it remains to estimate

$$\begin{aligned}\left\|\varepsilon^{-1}\hat{B}_j\left(\tfrac{k+m-k^{(j)}}{\varepsilon}\right)|k+m-k^{(j)}|\right\|^2_{L^2(\mathbb{T}^2)} &= \varepsilon^{-2}\int_{k\in\mathbb{T}^2}\left|\hat{B}_j\left(\tfrac{k+m-k^{(j)}}{\varepsilon}\right)\right|^2|k+m-k^{(j)}|^2\,\mathrm{d}k\\ &\leq C\varepsilon^2\int_{K\in\mathbb{R}^2}|\hat{B}_j(K)|^2|K|^2\,\mathrm{d}K \leq C\varepsilon^2\|\hat{B}_j\|^2_{L^2_s}\end{aligned} \tag{4.12}$$

for $s \geq 1$, and this proves (4.11). $\rfloor$

With the decomposition (4.8) we obtain the Lyapunov–Schmidt equations

$$\frac{1}{\varepsilon}(\omega_{n_j}(k) - \omega_* - \varepsilon^2\Omega)\hat{B}_j\left(\frac{k+m-k^{(j)}}{\varepsilon}\right) = -\sigma\chi_{\widetilde{D}_j}(k)\tilde{g}_{n_j}(k), \quad j = 1,\ldots,N,\ m \in M_j, \tag{4.13}$$

$$(\omega_n(k) - \omega_* - \varepsilon^2\Omega)\widetilde{\psi}_n(k) = -\sigma\left(1 - \sum_{j=1}^{N}\chi_{\widetilde{D}_j}(k)\delta_{n,n_j}\right)\tilde{g}_n(k), \quad n \in \mathbb{N}. \tag{4.14}$$

The goal is to solve (4.14) for the correction $\vec{\tilde{\psi}}$ as a function of $\hat{\mathbf{B}} = (\hat{B}_j)_{j=1}^N \in L^2(D_{\varepsilon^{r-1}}, \mathbb{C}^N)$ and plug this into (4.13). It turns out that the right norm for $\hat{B}_j$ is $\|\hat{B}_j\|_{L^2_s(D_{\varepsilon^{r-1}})}$, where we recall

$$\|\hat{B}_j\|_{L^2_s(D_{\varepsilon^{r-1}})} = \|(1+|p|)^s\hat{B}_j\|_{L^2(D_{\varepsilon^{r-1}})}.$$

Note that $L^2(D_{\varepsilon^{r-1}}) = L^2_s(D_{\varepsilon^{r-1}})$ as spaces for any $s \geq 0$, but below we need the estimate $\|B_j\|_{H^s} \leq C\|\hat{B}_j\|_{L^2_s(D_{\varepsilon^{r-1}})}$ with C independent of ε, cf. (1.5).

Lemma 4.5 *Let $s > 1$ and $V \in H^{\lceil s\rceil-1+\delta}_{\mathrm{loc}}(\mathbb{R}^2), \delta > 0$. For $\hat{\mathbf{B}} \in L^2_s(D_{\varepsilon^{r-1}})$ and $\vec{\tilde{\psi}} \in \mathcal{X}^s$, we have*

$$\begin{aligned}\|\vec{\tilde{g}}\|_{\mathcal{X}^s} \leq C\Bigg[&\varepsilon^2\left(\sum_{j=1}^{N}\|\hat{B}_j\|_{L^2_s(D_{\varepsilon^{r-1}})}\right)^3 + \varepsilon^2\left(\sum_{j=1}^{N}\|\hat{B}_j\|_{L^2_s(D_{\varepsilon^{r-1}})}\right)^2\|\vec{\tilde{\psi}}\|_{\mathcal{X}^s} \\ &+\varepsilon\left(\sum_{j=1}^{N}\|\hat{B}_j\|_{L^2_s(D_{\varepsilon^{r-1}})}\right)\|\vec{\tilde{\psi}}\|^2_{\mathcal{X}^s} + \|\vec{\tilde{\psi}}\|^3_{\mathcal{X}^s}\Bigg].\end{aligned} \tag{4.15}$$

Proof. Relying on the norm equivalence in (1.5), we derive most of the estimates in physical variables. Due to (4.8) and (4.9) we have $g = |\phi|^2\phi = |\varepsilon^{-1}\psi^{(0)}_{\mathrm{LS}} + \rho + \psi|^2(\varepsilon^{-1}\psi^{(0)}_{\mathrm{LS}} + \rho + \psi)$. We need to estimate norms of terms of the types

$$\left(\varepsilon^{-1}\psi^{(0)}_{\mathrm{LS}}\right)^3,\ \left(\varepsilon^{-1}\psi^{(0)}_{\mathrm{LS}}\right)^2 f,\ \varepsilon^{-1}\psi^{(0)}_{\mathrm{LS}}f^2,\ \rho^2\psi,\ \rho\psi^2,\ \rho^3, \text{ and } \psi^3. \tag{4.16}$$

First note that $\varepsilon^{-1}\psi^{(0)}_{\mathrm{LS}}(x) = \varepsilon\sum_{j=1}^N B_j(\varepsilon x)u_{n_j}(k^{(j)};x)$. Below we implicitly use $\|B_j(\varepsilon\cdot)u_{n_j}(k^{(j)};\cdot)\|_{H^s} \leq \|u_{n_j}(k^{(j)};\cdot)\|_{C^{\lceil s\rceil}}\|B_j(\varepsilon\cdot)\|_{H^s}$, which holds by interpolation, see, e.g., [38, §4.2]. Next, $\|u_{n_j}(k^{(j)};\cdot)\|_{C^{\lceil s\rceil}} \leq C\|u_{n_j}(k^{(j)};\cdot)\|_{H^{\lceil s\rceil+1+\delta}} \leq C\|V\|_{H^{\lceil s\rceil-1+\delta}_{\mathrm{loc}}}$ for all $\delta > 0$, where the first inequality holds by Sobolev embedding and the second one by the differential equation.

In estimating all except the first term in (4.16) we use the Banach algebra property in Lemma 4.2.

For the first two terms we need to estimate $\|(\varepsilon B_j(\varepsilon\cdot))^n\|_{H^s}$ for $n=2,3$. We have

$$
\begin{aligned}
\|(\varepsilon B_j(\varepsilon\cdot))^n\|_{H^s}^2 &= \int (1+|k|)^{2s}|\mathcal{F}((\varepsilon B_j(\varepsilon\cdot))^n)(k)|^2\,\mathrm{d}k\\
&\le C\left[\int |\mathcal{F}((\varepsilon B_j(\varepsilon\cdot))^n)(k)|^2\,\mathrm{d}k + \varepsilon^{2n-4}\int |k|^{2s}\left|\widehat{B_j^n}\left(\frac{k}{\varepsilon}\right)\right|^2\,\mathrm{d}k\right]\\
&\le C\left[\|(\varepsilon B_j(\varepsilon\cdot))^n\|_{L^2}^2 + \varepsilon^{2n-2+2s}\int |K|^{2s}|\widehat{B_j^n}(K)|^2\,\mathrm{d}K\right]\\
&\le C\left[\varepsilon^{2(n-1)}\|B_j\|_{L^\infty}^{2(n-1)}\|\varepsilon B_j(\varepsilon\cdot)\|_{L^2}^2 + \varepsilon^{2n-2+2s}\|B_j^n\|_{H^s}^2\right]\\
&\le C\left[\varepsilon^{2(n-1)}\|B_j\|_{H^s}^{2(n-1)}\|B_j\|_{L^2}^2 + \varepsilon^{2(n-1)+2s}\|B_j\|_{H^s}^{2n}\right]
\end{aligned}
$$

and hence

$$\|(\varepsilon B_j(\varepsilon\cdot))^n\|_{H^s} \le C\varepsilon^{n-1}\|B_j\|_{H^s}^n \text{ for } n=1,2,3. \tag{4.17}$$

Note that for $n\ge 2$ this is much better than the naive estimate $\|(\varepsilon B_j(\varepsilon\cdot))^n\|_{H^s}\le C\|\varepsilon B_j(\varepsilon\cdot)\|_{H^s}^n \le C\|B\|_{H^s}^n$ based on (4.17) with $n=1$. Next, for the third term in (4.16) we get

$$\|\varepsilon B_j(\varepsilon\cdot)f(\cdot)\|_{H^s}\le C\varepsilon\|B_j\|_{H^s}\|f\|_{H^s}, \tag{4.18}$$

and this together with (4.17) can be used to prove (4.15). To show (4.18), we start with

$$\|\varepsilon B_j(\varepsilon\cdot)f(\cdot)\|_{H^s}\le \|\varepsilon B_j(\varepsilon\cdot)f(\cdot)\|_{L^2} + C\left\||k|^s\left(\frac{1}{\varepsilon}\hat{B}_j\left(\frac{\cdot}{\varepsilon}\right)*\hat{f}(\cdot)\right)\right\|_{L^2}.$$

The first term is estimated as $\|\varepsilon B_j(\varepsilon\cdot)f(\cdot)\|_{L^2}\le\varepsilon\|B_j\|_\infty\|f\|_{L^2}\le\varepsilon\|B_j\|_{H^s}\|f\|_{L^2}$, and for the second we note that $w(k)\le\varepsilon w(\frac{k-l}{\varepsilon})+w(l)$ where $w(k)=|k|^s$. Thus, similarly to the proof of Lemma 4.2, we have, using Young's inequality,

$$
\begin{aligned}
\left\|w(k)\left(\frac{1}{\varepsilon}\hat{B}_j\left(\frac{\cdot}{\varepsilon}\right)*\hat{f}(\cdot)\right)\right\|_{L^2} &\le C\left\|\left|w\left(\frac{\cdot}{\varepsilon}\right)\hat{B}_j\left(\frac{\cdot}{\varepsilon}\right)\right|*|\hat{f}(\cdot)| + \left|\frac{1}{\varepsilon}\hat{B}_j\left(\frac{\cdot}{\varepsilon}\right)\right|*|w(\cdot)\hat{f}(\cdot)|\right\|_{L^2}\\
&\le C\left[\left\|w\left(\frac{\cdot}{\varepsilon}\right)\hat{B}_j\left(\frac{\cdot}{\varepsilon}\right)\right\|_{L^2}\|\hat{f}\|_{L^1} + \left\|\frac{1}{\varepsilon}\hat{B}_j\left(\frac{\cdot}{\varepsilon}\right)\right\|_{L^1}\|w\hat{f}\|_{L^2}\right].
\end{aligned}
$$

Now $\|w\left(\frac{\cdot}{\varepsilon}\right)\hat{B}_j\left(\frac{\cdot}{\varepsilon}\right)\|_{L^2}\le\varepsilon\|\hat{B}_j\|_{L^2_s}$, $\left\|\frac{1}{\varepsilon}\hat{B}_j\left(\frac{\cdot}{\varepsilon}\right)\right\|_{L^1}\le C\varepsilon\|\hat{B}_j\|_{L^2_s}$, and $\|\hat{f}\|_{L^1}\le C\|\hat{f}\|_{L^2_s}$ (see (4.2)) yield (4.18).

The 4th, 5th and 7th term in (4.16) are estimated simply using Lemma 4.2 and (4.11). The 6th term is treated the same way and is bounded by $\varepsilon^3\big(\sum_{j=1}^N\|\hat{B}_j\|_{L^2_s(D_{\varepsilon^{r-1}})}\big)^3$ so that it is of higher order than the first term on the right hand side of (4.15). $\square$

We seek now a small solution of (4.14). First, for ε sufficiently small we have

$$\min_{\substack{k\in\mathrm{supp}(\vec{\tilde{\psi}})\\ n\in\mathbb{N}}}|\omega_n(k)-\omega_*|\ge C\varepsilon^{2r} \tag{4.19}$$

due to $\partial_{k_1}\omega_{n_j}(k^{(j)}) = \partial_{k_2}\omega_{n_j}(k^{(j)}) = 0$ and the definiteness of the Hessian of ω_{n_j} at $k = k^{(j)}$. We rewrite now (4.14) as

$$\tilde{\psi}_n(k) = (\omega_n(k) - \omega_*)^{-1}(-\sigma\tilde{g}_n(k) + \varepsilon^2\Omega\tilde{\psi}_n(k)) =: \tilde{F}_n(\vec{\tilde{\psi}})(k)$$

and solve $\vec{\tilde{\psi}} = \vec{\tilde{F}}(\vec{\tilde{\psi}})$ on a neighborhood of 0, namely on $B_{\varepsilon^\eta} := \{\vec{\tilde{\psi}} \in \mathcal{X}^s : \|\vec{\tilde{\psi}}\|_{\mathcal{X}^s} \le \varepsilon^\eta\}, \eta > 0$, via the Banach fixed point theorem.

Performing similar estimates as in the proof of Lemma (4.5) and using (4.19), we get

$$\|F(\psi) - F(\phi)\|_{H^s} \le C\varepsilon^{-2r}\left[\varepsilon^2 + \varepsilon(\|\psi\|_{H^s} + \|\phi\|_{H^s}) + \|\psi\|_{H^s}^2 + \|\phi\|_{H^s}^2\right]\|\psi - \phi\|_{H^s},$$

where $C = C(\sum_{j=1}^N \|B_j\|_{H^s(\mathbb{R}^2)}, |\Omega|)$. The contraction property thus holds if

$$\varepsilon^{2-2r} + \varepsilon^{1-2r}(\|\psi\|_{H^s} + \|\phi\|_{H^s}) + \varepsilon^{-2r}(\|\psi\|_{H^s}^2 + \|\phi\|_{H^s}^2) < 1 \tag{4.20}$$

for all $\vec{\tilde{\psi}}, \vec{\tilde{\phi}} \in B_{\varepsilon^\eta}$ and if $\vec{\tilde{F}}$ maps B_{ε^η} to B_{ε^η}, i.e., using (4.15), if

$$\varepsilon^{2-2r} + \varepsilon^{2-2r}\|\vec{\tilde{\psi}}\|_{\mathcal{X}^s} + \varepsilon^{1-2r}\|\vec{\tilde{\psi}}\|_{\mathcal{X}^s}^2 + \varepsilon^{-2r}\|\vec{\tilde{\psi}}\|_{\mathcal{X}^s}^3 < C\varepsilon^\eta \tag{4.21}$$

for all $\vec{\tilde{\psi}} \in B_{\varepsilon^\eta}$.

Condition (4.20) is satisfied if $r < 1$ and $\eta > \max(2r-1, r)$ and (4.21) holds if $\max(2r-1, r) < \eta < 2-2r$ and $r < 1$. In combination these yield

$$\eta \in (r, 2-2r), \quad 0 < r < \tfrac{2}{3}. \tag{4.22}$$

Here is the reason for the upper bound in (3.11).

In conclusion we have the existence of $\vec{\tilde{\psi}}$ satisfying

$$\|\vec{\tilde{\psi}}\|_{\mathcal{X}^s} \le C\varepsilon^{2-2r}\sum_{j=1}^N \|\hat{B}_j\|_{L^2_s(D_{\varepsilon^{r-1}})} \quad \text{with} \;\; 0 < r < 2/3. \tag{4.23}$$

We now turn to (4.13). Plugging (4.8) into (4.7), truncating over $k \in D_j$ and mapping $k \in D_j - m$ to $p \in D_{\varepsilon^{r-1}}$ via $p = \varepsilon^{-1}(k + m - k^{(j)})$ yields the so called extended CMEs (eCMEs) in the form

$$\Omega\hat{B}_j - \left(\frac{1}{2}\partial_{k_1}^2\omega_{n_j}(k^{(j)})p_1^2 + \frac{1}{2}\partial_{k_2}^2\omega_{n_j}(k^{(j)})p_2^2 + \partial_{k_1}\partial_{k_2}\omega_{n_j}(k^{(j)})p_1p_2\right)\hat{B}_j - \hat{Q}_j = \varepsilon^{\tilde{r}}\hat{R}_j(p), \tag{4.24}$$

$j = 1, \ldots, N$. Here $\hat{Q}_j$ denotes the nonlinear term $\hat{\mathcal{N}}_j$ in (3.32) with $\hat{A}_i$ replaced by $\hat{B}_i$ and truncated on $p \in D_{\varepsilon^{r-1}}$, i.e.,

$$\hat{Q}_j(p^{(j,m)}) = \frac{\sigma}{\varepsilon^4}\chi_{D_{\varepsilon^{r-1}}}(p^{(j,m)})\langle(\tilde{\psi}_{\rm LS}^{(0)} *_B \bar{\tilde{\psi}}_{\rm LS}^{(0)} *_B \tilde{\psi}_{\rm LS}^{(0)})(\varepsilon p^{(j,m)} + k^{(j)} - m;\cdot), p_{n_j}(k^{(j)};\cdot)e^{im\cdot\cdot}\rangle_{L^2(\mathbb{P}^2)}$$

with $\tilde{\psi}_{\rm LS}^{(0)}$ given in (4.10). In the persistence step below we need to control the remainder $\varepsilon^{\tilde{r}}\hat{R}_j(p)$, where we set $\tilde{r} = \min\{r, 2-2r, 1\}$, and which has the form

$$\varepsilon^{\tilde{r}} \hat{R}_j(p^{(j,m)}) = \varepsilon^{-2} q_j(\varepsilon p^{(j,m)}) \hat{B}_j(p^{(j,m)}) + \varepsilon^{-1} \sigma \chi_{D_{\varepsilon^{r-1}}}(p^{(j,m)}) \tilde{g}_{n_j}(\varepsilon p^{(j,m)} + k^{(j)} - m) - \hat{Q}_j(p^{(j,m)}). \quad (4.25)$$

The first term in (4.25) comes from the Taylor expansion of ω_{n_j}, i.e.,

$$q_j(\varepsilon p) := \omega_{n_j}(k^{(j)} + \varepsilon p) - T_2(\omega_{n_j}(k^{(j)}))(\varepsilon p) \quad (4.26)$$

with $T_2(\omega_{n_j}(k^{(j)})$ the second order Taylor polynomial. Since ω_{n_j} is analytic near $k^{(j)}$, we have that q_j is analytic as well and starts with a cubic polynomial. Given $\hat{B}_j \in L^2_s(D_{\varepsilon^{r-1}})$, in the persistence step we need to bound the remainder in $L^2_{s-2}(D_{\varepsilon^{r-1}})$. For the cubic terms $q_j^{(3)}(\varepsilon p)$ in $q_j(\varepsilon p)$ we get

$$\begin{aligned}
\|\varepsilon^{-2} q_j^{(3)}(\varepsilon \cdot) \hat{B}_j(\cdot)\|^2_{L^2_{s-2}(D_{\varepsilon^{r-1}})} &\le C \int_{D_{\varepsilon^{r-1}}} \varepsilon^2 |p|^6 |\hat{B}_j(p)|^2 (1+|p|^2)^{s-2} \, \mathrm{d}p \\
&\le C \int_{D_{\varepsilon^{r-1}}} \varepsilon^2 |p|^2 |\hat{B}_j(p)|^2 (1+|p|^2)^{s} \, \mathrm{d}p \le C \varepsilon^{2r} \|\hat{B}_j\|^2_{L^2_s(D_{\varepsilon^{r-1}})}, \quad (4.27)
\end{aligned}$$

and clearly higher order contributions from $q_j(\varepsilon p)$ gain additional powers in ε.

The leading order contribution to the remaining two terms in $\varepsilon^{\tilde{r}} \hat{R}_j(p)$ comes from $\vec{\tilde{\psi}}$, which is estimated via (4.23) to be $\mathcal{O}(\varepsilon^{2-2r})$. This is guaranteed to be small since $r < 2/3$. The last contribution to $\varepsilon^{\tilde{r}} \hat{R}_j(p)$ comes from the $\tilde{\psi}$-independent terms in $\chi_{D_{\varepsilon^{r-1}}}(p) \tilde{g}_{n_j}(\varepsilon p + k^{(j)} - m)$, i.e., from

$$\begin{aligned}
\frac{\sigma}{\varepsilon^4} \chi_{D_{\varepsilon^{r-1}}}(p) \Big[&\langle (\tilde{\eta}^{(0)}_{\mathrm{LS}} *_B \tilde{\bar{\eta}}^{(0)}_{\mathrm{LS}} *_B \tilde{\eta}^{(0)}_{\mathrm{LS}})(\varepsilon p + k^{(j)} - m; \cdot), p_{n_j}(\varepsilon p + k^{(j)}; \cdot) e^{\mathrm{i}m \cdot \cdot} \rangle_{L^2(\mathbb{P}^2)} \\
&- \langle (\tilde{\psi}^{(0)}_{\mathrm{LS}} *_B \tilde{\bar{\psi}}^{(0)}_{\mathrm{LS}} *_B \tilde{\psi}^{(0)}_{\mathrm{LS}})(\varepsilon p + k^{(j)} - m; \cdot), p_{n_j}(k^{(j)}; \cdot) e^{\mathrm{i}m \cdot \cdot} \rangle_{L^2(\mathbb{P}^2)} \Big].
\end{aligned}$$

This difference is estimated in L^2_s to leading order via $\|\tilde{\rho}\|_{L^2_s}$, which is $\mathcal{O}(\varepsilon)$ according to (4.11).

These estimates are not strictly needed for our first main result, which follows directly from the above Lyapunov-Schmidt reduction, and which *assumes* the existence of solutions of the extended CMEs (4.24), but the estimates show that we may expect solutions of the extended CMEs (4.24) near solutions of the CME, as will be worked out in §4.3.

Theorem 4.6 *Let* $s>1, V \in H^{\lceil s \rceil - 1 + \delta}_{\mathrm{loc}}(\mathbb{R}^2)$ *for some* $\delta>0$ *and let* $0 < r < \frac{2}{3}$. *There exist* $\varepsilon_0, C_1, C_2 > 0$ *such that for all* $0 < \varepsilon < \varepsilon_0$ *the following holds. Assume that there exists a solution* $(\hat{B}_j)_{j=1}^N \in L^2(D_{\varepsilon^{r-1}})$ *of the extended CMEs (4.24) with* $\|\hat{B}_j\|_{L^2_s(D_{\varepsilon^{r-1}})} \le C_1$. *Then* (1.2) *has a solution* $\phi \in H^s(\mathbb{R}^2)$ *with*

$$\|\phi(\cdot) - \varepsilon \sum_{j=1}^N B_j(\varepsilon \cdot) u_{n_j}(k^{(j)}; \cdot)\|_{H^s(\mathbb{R}^2)} \le C_2(\varepsilon^{2-2r} + \varepsilon), \quad (4.28)$$

where $u_{n_j}(k^{(j)}; x)$ *are the resonant Bloch waves and* $B_j = \mathcal{F}^{-1} \hat{B}_j$.

Proof. By construction, a solution $(\hat{B}_j)_{j=1}^N$ of (4.24) yields via (4.8) a solution $\vec{\tilde{\phi}}(k)$ of (4.6). The estimate (4.23) and the norm equivalence in Lemma 4.1 yield $\|\phi - \varepsilon^{-1} \eta^{(0)}_{\mathrm{LS}}\|_{H^s(\mathbb{R}^2)} \le C \varepsilon^{2-2r}$. And because $\varepsilon \sum_{j=1}^N B_j(\varepsilon x) u_{n_j}(k^{(j)}; x) = \varepsilon^{-1}\psi_{\rm LS}^{(0)}(x)$, we have $\varepsilon^{-1}\eta_{\rm LS}^{(0)}(x) - \varepsilon \sum_{j=1}^N B_j(\varepsilon x) u_{n_j}(k^{(j)}; x) = \rho(x)$, which is bounded in (4.11) and the approximation result (4.28) follows. □

4.3 Justification Step II: Persistence

The remaining step is to make a connection between solitary waves of the CMEs (3.32) and the eCMEs (4.24). To obtain existence of solutions of the eCMEs (4.24), we show a persistence result of special so called reversible non-degenerate CME solitons to the eCME. This is quite similar to [13, §5] but in order to deal with an arbitrary set of extrema $k^{(j)}, j \in \{1, ..., N\}$, the definition of reversibility needs to be generalized.

The eCMEs (4.24) differ from the CMEs (3.32) in three ways: the $\hat{B}_j(p)$ are supported on $D_{\varepsilon^{r-1}}$, the convolutions are truncated on $D_{\varepsilon^{r-1}}$, and the remainder $\varepsilon^{\tilde{r}}\hat{R}_j(p)$ is included. The idea is to handle these differences as perturbations and thus seek a solution $\hat{\mathbf{B}} = (\hat{B}_j)_{j=1,...,N}$ of the eCMEs near a given solution $\hat{\mathbf{A}} = (\hat{A}_j)_{j=1,...,N}$ of the CMEs. Note that the $\hat{A}_j$ are not supported on $D_{\varepsilon^{r-1}}$ and thus first must be truncated.

We start with a formal discussion. We write the CME in abstract form as $\hat{\mathbf{F}}(\hat{\mathbf{A}}) = 0$ and the eCME as

$$\chi_{D_{\varepsilon^{r-1}}} \hat{\mathbf{F}}(\hat{\mathbf{B}}) = \varepsilon^{\tilde{r}} \hat{\mathbf{R}}(\hat{\mathbf{B}}), \tag{4.29}$$

assume a solution $\hat{\mathbf{A}} \in [L_q^2(\mathbb{R}^2)]^N$ for any $q \geq 0$ of the CME, and look for solutions $\hat{\mathbf{B}} \in [L_s^2(D_{\varepsilon^{r-1}})]^N$ of the eCME in the form $\hat{\mathbf{B}} = \hat{\mathbf{A}}_\varepsilon + \hat{\mathbf{b}}$ with $\hat{\mathbf{A}}_\varepsilon = \chi_{D_{\varepsilon^{r-1}}} \hat{\mathbf{A}}$ and $\operatorname{supp} \hat{\mathbf{b}} \subset D_{\varepsilon^{r-1}}$. This yields

$$\begin{aligned} &\hat{\mathbf{J}}_\varepsilon \hat{\mathbf{b}} = \hat{\mathbf{N}}(\hat{\mathbf{b}}), \ p \in D_{\varepsilon^{r-1}} \quad \text{ with} \\ &\hat{\mathbf{J}}_\varepsilon = \chi_{D_{\varepsilon^{r-1}}} D_{\hat{\mathbf{A}}} \hat{\mathbf{F}}(\hat{\mathbf{A}}) \quad \text{and} \quad \hat{\mathbf{N}}(\hat{\mathbf{b}}) = \varepsilon^{\tilde{r}} \hat{\mathbf{R}}(\hat{\mathbf{A}}_\varepsilon + \hat{\mathbf{b}}) - (\chi_{D_{\varepsilon^{r-1}}} \hat{\mathbf{F}}(\hat{\mathbf{A}}_\varepsilon + \hat{\mathbf{b}}) - \hat{\mathbf{J}}_\varepsilon \hat{\mathbf{b}}). \end{aligned} \tag{4.30}$$

Note that in (4.30) we may replace $\hat{\mathbf{N}}(\hat{\mathbf{b}})$ by $\chi_{D_{\varepsilon^{r-1}}} \hat{\mathbf{N}}(\hat{\mathbf{b}})$ to display explicitly that $p \in D_{\varepsilon^{r-1}}$.

We have $\hat{\mathbf{F}}(\hat{\mathbf{A}}_\varepsilon + \hat{\mathbf{b}}) - \hat{\mathbf{J}}_\varepsilon \hat{\mathbf{b}} = \hat{\mathbf{F}}(\hat{\mathbf{A}}_\varepsilon) + (D_{\hat{\mathbf{A}}} \hat{\mathbf{F}}(\hat{\mathbf{A}}_\varepsilon) - \hat{\mathbf{J}}_\varepsilon)\hat{\mathbf{b}} + \hat{\mathbf{G}}(\hat{\mathbf{b}})$ with $\hat{\mathbf{G}}(\hat{\mathbf{b}})$ quadratic in $\hat{\mathbf{b}}$. Moreover,

$$\hat{\mathbf{F}}(\hat{\mathbf{A}}_\varepsilon) = L^{\rm CME}\, \chi_{D_{\varepsilon^{r-1}}} \hat{\mathbf{A}} - \chi_{D_{\varepsilon^{r-1}}} \hat{\mathcal{N}}(\hat{\mathbf{A}}) - (\hat{\mathcal{N}}(\hat{\mathbf{A}}_\varepsilon) - \chi_{D_{\varepsilon^{r-1}}} \hat{\mathcal{N}}(\hat{\mathbf{A}})) = -(\hat{\mathcal{N}}(\hat{\mathbf{A}}_\varepsilon) - \chi_{D_{\varepsilon^{r-1}}} \hat{\mathcal{N}}(\hat{\mathbf{A}})),$$

where $L^{\rm CME}$ denotes the linear part of the operator in (3.32) and $\hat{\mathcal{N}}$ denotes the $N-$dimensional vector of the nonlinear terms in (3.32). $\chi_{D_{\varepsilon^{r-1}}} \hat{\mathcal{N}}(\hat{\mathbf{A}})(p)$ is a sum of convolutions $\hat{A}_{j_1} * \hat{A}_{j_2} * \hat{\bar{A}}_{j_3}$, and hence in $\hat{\mathcal{N}}(\hat{\mathbf{A}}_\varepsilon) - \chi_{D_{\varepsilon^{r-1}}} \hat{\mathcal{N}}(\hat{\mathbf{A}})$ this yields terms of the form

$$\int_{p_1} \int_{|p_2| \geq \varepsilon^{r-1}} \chi_{D_{\varepsilon^{r-1}}}(p - p_1) \chi_{D_{\varepsilon^{r-1}}}(p_1 - p_2) \hat{A}_{j_1}(p - p_1) \hat{A}_{j_2}(p_1 - p_2) \hat{\bar{A}}_{j_3}(p_2)\, {\rm d}p_1\, {\rm d}p_2,$$

which can be bounded by $C\varepsilon^q$ for any $q > 0$ in $L_s^2(D_{\varepsilon^{r-1}})$ due to the fast decay of $\hat{\mathbf{A}}$. Therefore,

$$\|\hat{\mathbf{F}}(\hat{\mathbf{A}}_\varepsilon)\|_{L_s^2} \leq C\varepsilon^q, \quad \text{and} \quad \|(D_{\hat{\mathbf{A}}} \hat{\mathbf{F}}(\hat{\mathbf{A}}_\varepsilon) - \hat{\mathbf{J}}_\varepsilon)\hat{\mathbf{b}}\|_{L_s^2} \leq C\varepsilon^q \|\hat{\mathbf{b}}\|_{L_s^2} \tag{4.31}$$

by a similar estimate.

Thus $\|\hat{\mathbf{N}}(\hat{\mathbf{b}})\|_{L_{s-2}^2} \leq C(\varepsilon^{\tilde{r}} + \varepsilon^q + \varepsilon^{\tilde{r}} \|\hat{\mathbf{b}}\|_{L_s^2} + \varepsilon^q \|\hat{\mathbf{b}}\|_{L_s^2} + \|\hat{\mathbf{b}}\|_{L_s^2}^2)$, and this suggests applying the contraction mapping theorem to (4.30) in the form

$$\hat{\mathbf{b}} = \hat{\mathbf{J}}_\varepsilon^{-1}\hat{\mathbf{N}}(\hat{\mathbf{b}}). \tag{4.32}$$

To discuss $\hat{\mathbf{J}}_\varepsilon^{-1}$, we start with $\mathbf{J} : H^s(\mathbb{R}^2) \to H^{s-2}(\mathbb{R}^2)$. The continuous spectrum $\sigma_c(\mathbf{J})$ of $\mathbf{J}$ equals that of L^{CME}. Thus, if $\omega = \omega_* + \varepsilon^2\Omega$ is in a gap, then Ω and the quadratic forms defined by $\frac{1}{2}\partial_{k_1}^2\omega_{n_j}(k^{(j)})p_1^2 + \frac{1}{2}\partial_{k_2}^2\omega_{n_j}(k^{(j)}p_2^2 + \partial_{k_1}\partial_{k_2}\omega_{n_j}(k^{(j)})p_1p_2$, $j{=}1,\ldots,N$ have opposite signs such that $\sigma_c(\mathbf{J})$ is bounded away from zero. However, the problem is that $\mathbf{J}$ has a nontrivial kernel since Ker $\mathbf{J}$ contains at least $\partial_{y_1}\mathbf{A}, \partial_{y_2}\mathbf{A}$ and i$\mathbf{A}$ which follows from the translational and phase invariances of the CME. For $\hat{\mathbf{J}}_\varepsilon^{-1} : L^2_s(D_{\varepsilon^{r-1}}) \to L^2_{s+2}(D_{\varepsilon^{r-1}})$ (if it exists) this implies that it cannot be bounded independently of ε.

The solution is to consider (4.32) in a subspace $X_{\mathrm{rev}} \subset L^2_s(D_{\varepsilon^{r-1}})$ where $\hat{\mathbf{J}}_\varepsilon^{-1}$ is bounded, and where $\hat{\mathbf{b}} \in X_{\mathrm{rev}}$ implies $\hat{\mathbf{N}}(\hat{\mathbf{b}}) \in X_{\mathrm{rev}}$. This can be achieved by symmetries of the problem (1.2) if we assume that $\mathbf{J}$ on $H^s(\mathbb{R}^2)$ has only $\partial_{y_1}\mathbf{A}, \partial_{y_2}\mathbf{A}$ and i$\mathbf{A}$ in its kernel.

The original problem (1.2) is equivariant under the symmetries

$$\phi(x_1, x_2) = \xi_1\bar{\phi}(-x_1, x_2) = \xi_2\bar{\phi}(x_1, -x_2) \tag{4.33}$$

where $\xi_1, \xi_2 = \pm 1$, i.e., $(\xi_1, \xi_2) = (1, 1)$ or $(\xi_1, \xi_2) = (1, -1)$ or $(\xi_1, \xi_2) = (-1, 1)$ or $(\xi_1, \xi_2) = (-1, -1)$, and similarly

$$\phi(x_1, x_2) = \xi_1\bar{\phi}(x_2, x_1) = \xi_2\bar{\phi}(-x_2, -x_1), \tag{4.34}$$

where again $\xi_1, \xi_2 = \pm 1$.

These symmetries have their counterparts in symmetries of the CMEs. As we show, the reversibility in the following definition firstly guarantees that $\hat{\mathbf{b}} \in X_{\mathrm{rev}}$ implies $\hat{\mathbf{N}}(\hat{\mathbf{b}}) \in X_{\mathrm{rev}}$, secondly provides a leading order approximation $\varepsilon\eta^{(0)}$ satisfying (4.33) or (4.34), and lastly (under a non-degeneracy condition) guarantees the existence of a Gross-Pitaevskii solution ϕ satisfying (4.33) or (4.34).

Definition 4.7 *A solution* $\mathbf{A}$ *of (3.33) is called reversible if it fulfills one of the symmetries in* (4.35) *or* (4.36)*.*

$$A_i(y) = s_1\bar{A}_{i''}(-y_1, y_2) = s_2\bar{A}_{i'}(y_1, -y_2) \text{ for all } i \in \{1, \ldots, N\} \text{ with } s_{1,2} = \pm 1 \text{ independent of } i, \tag{4.35}$$

and where the indices i' *and* i'' *are defined by*

- $k^{(i')} = (-k_1^{(i)}, k_2^{(i)})$ *if* $k_1^{(i)} < \frac{1}{2}$ *and* $i' = i$ *if* $k_1^{(i)} = \frac{1}{2}$,
- $k^{(i'')} = (k_1^{(i)}, -k_2^{(i)})$ *if* $k_2^{(i)} < \frac{1}{2}$ *and* $i'' = i$ *if* $k_2^{(i)} = \frac{1}{2}$.

$$A_i(y) = s_1\bar{A}_{i'}(y_2, y_1) = s_2\bar{A}_{i''}(-y_2, -y_1) \text{ for all } i \in \{1, \ldots, N\} \text{ with } s_{1,2} = \pm 1 \text{ independent of } i, \tag{4.36}$$

and where the indices i' *and* i'' *are defined by*

- $k^{(i')} = (k_2^{(i)}, k_1^{(i)})$,
- $k^{(i'')} = (-k_2^{(i)}, -k_1^{(i)})$ *if* $k_{1,2}^{(i)} < \frac{1}{2}$, $\quad i'' = i$ *if* $k^{(i)} = M$,
  $k^{(i'')} = (-k_2^{(i)}, k_1^{(i)})$ *if* $k_1^{(i)} = \frac{1}{2}, k^{(i)} \neq M$, $\quad k^{(i'')} = (k_2^{(i)}, -k_1^{(i)})$ *if* $k_2^{(i)} = \frac{1}{2}, k^{(i)} \neq M$.

We define the space

$$X_{\rm rev} = \{\hat{\mathbf{f}} \in L^2_s(D_{\varepsilon^{r-1}}) : \mathbf{f} \textit{ satisfies } (4.35) \textit{ or } (4.36)\}.$$

Definition 4.8 *A solution* $\mathbf{A}$ *of* (3.33) *is called non-degenerate if Ker* $\mathbf{J} = \{\partial_{y_1}\mathbf{A}, \partial_{y_2}\mathbf{A}, \mathrm{i}\mathbf{A}\}$.

We explain next that under the assumption $\hat{\mathbf{A}} \in X_{\rm rev}$ we, indeed, have the property

$$\hat{\mathbf{b}} \in X_{\rm rev} \Rightarrow \hat{\mathbf{N}}(\hat{\mathbf{b}}) \in X_{\rm rev}. \tag{4.37}$$

Note that this resembles the question of inheritance of symmetries of the full problem for ϕ by the Lyapunov-Schmidt reduction, see e.g. Prop. 3.3 in [18]. We are, however, using a generalized Lyapunov-Schmidt reduction in which a whole neighborhood of the kernel is isolated. Secondly, we wish to carry symmetries of the scalar problem for ϕ over to symmetries of the vector problem for the envelopes $\hat{\mathbf{B}}$. We inspect, therefore, problem (4.37) by hand. For the sake of brevity we present the analysis only for the first symmetry in (4.35), i.e.,

$$A_i(y) = s_1 \bar{A}_{i''}(-y_1, y_2) \quad \text{for all } i \in \{1, \ldots, N\}. \tag{4.38}$$

This is equivalent to $\hat{A}_i(p) = s_1\hat{\bar{A}}_{i''}(-p_1, p_2)$ for all $i \in \{1, \ldots, N\}$ in Fourier variables. If we denote $S\mathbf{A}(y) = \pm(\bar{A}_{1''}, \bar{A}_{2''}, \ldots, \bar{A}_{N''})(-y_1, y_2)$, condition (4.38) reads $\mathbf{A} = S\mathbf{A}$.

In the term $\chi_{D_{\varepsilon^{r-1}}}\hat{\mathbf{F}}(\hat{\mathbf{A}}_\varepsilon + \hat{\mathbf{b}})$ in $\hat{\mathbf{N}}(\hat{\mathbf{b}})$ the operator $\mathbf{F}$ is the CME operator, which commutes with S. In the argument of $\hat{\mathbf{F}}$ the function $\mathbf{A}_\varepsilon$ satisfies (4.38) because $\mathbf{A}$ does and the symmetric cut-off $\hat{\mathbf{A}}_\varepsilon = \chi_{D_{\varepsilon^{r-1}}}(p)\hat{\mathbf{A}}$ preserves the symmetry. The linear operator $\mathbf{J}_\varepsilon$ commutes with S because it is a symmetric cut-off of the linear operator corresponding to CMEs. It remains to discuss the term $\varepsilon^{\tilde{r}}\hat{\mathbf{R}}(\hat{\mathbf{A}}_\varepsilon + \hat{\mathbf{b}})$ in $\hat{\mathbf{N}}(\hat{\mathbf{b}})$. The first term on the right hand side of (4.25) is the band structure with its quadratic Taylor expansion subtracted away, i.e.

$$\begin{aligned} &\varepsilon^{-2}q_j(\varepsilon p)\hat{B}_j(p) = \\ &\left(\varepsilon^{-2}\omega_{n_j}(k^{(j)} + \varepsilon p - m) - \tfrac{1}{2}\partial^2_{k_1}\omega_{n_j}(k^{(j)})p_1^2 - \tfrac{1}{2}\partial^2_{k_2}\omega_{n_j}(k^{(j)})p_2^2 - \partial_{k_1}\partial_{k_2}\omega_{n_j}(k^{(j)})p_1p_2\right)\hat{B}_j(p), \end{aligned} \tag{4.39}$$

where we write p instead of $p^{(j,m)}$ and keep in mind that Range(p) depends on j and m. To ensure that the right hand side of (4.39) lies in $X_{\rm rev}$, we need to assume evenness of V in both x_1 and x_2. In that case $\omega_n(k) = \omega_n(-k_1, k_2) = \omega_n(k_1, -k_2)$, cf. (2.5). This implies, first of all,

$$\begin{aligned} &\partial^2_{k_1}\omega_{n_j}(k^{(j)}) = \partial^2_{k_1}\omega_{n_j}(k^{(j')}) = \partial^2_{k_1}\omega_{n_j}(k^{(j'')}), \quad \partial^2_{k_2}\omega_{n_j}(k^{(j)}) = \partial^2_{k_2}\omega_{n_j}(k^{(j')}) = \partial^2_{k_2}\omega_{n_j}(k^{(j'')}), \\ &\partial_{k_1}\partial_{k_2}\omega_{n_j}(k^{(j)}) = -\partial_{k_1}\partial_{k_2}\omega_{n_j}(k^{(j')}) = -\partial_{k_1}\partial_{k_2}\omega_{n_j}(k^{(j'')}). \end{aligned}$$

Hence, the quadratic part in (4.39) satisfies (4.38), i.e.

$$\begin{aligned} &\left(-\tfrac{1}{2}\partial^2_{k_1}\omega_{n_j}(k^{(j)})p_1^2 - \tfrac{1}{2}\partial^2_{k_2}\omega_{n_j}(k^{(j)})p_2^2 - \partial_{k_1}\partial_{k_2}\omega_{n_j}(k^{(j)})p_1p_2\right)\hat{b}_j(p) = \\ &\qquad s_1\left(-\tfrac{1}{2}\partial^2_{k_1}\omega_{n_j}(k^{(j'')})p_1^2 - \tfrac{1}{2}\partial^2_{k_2}\omega_{n_j}(k^{(j'')})p_2^2 - \partial_{k_1}\partial_{k_2}\omega_{n_j}(k^{(j'')})p_1p_2\right)\hat{\bar{b}}_{j''}(p_1, -p_2) \end{aligned}$$

since $\bar{\hat{b}}_{j''}(p_1,-p_2)=\hat{\bar{b}}_{j''}(-p_1,p_2)$

For the first term in (4.39) note first that

$$\omega_{n_j}(k^{(j)}+\varepsilon p-m)\hat{b}_j(p)=\omega_{n_j}\left(\begin{pmatrix}k_1^{(j)}\\-k_2^{(j)}\end{pmatrix}+\varepsilon\begin{pmatrix}p_1\\-p_2\end{pmatrix}-\begin{pmatrix}m_1\\-m_2\end{pmatrix}\right)\bar{\hat{b}}_{j''}(p_1,-p_2). \tag{4.40}$$

If $k_2^{(j)}<1/2$, then $(k_1^{(j)},-k_2^{(j)})^T=k^{(j'')}$ and $m_2=0$ so that $\omega_{n_j}(k^{(j)}+\varepsilon p-m)\hat{b}_j(p)=\omega_{n_j}(k^{(j'')}+\varepsilon\begin{pmatrix}p_1\\-p_2\end{pmatrix}-m)\bar{\hat{b}}_{j''}(p_1,-p_2)$, which is (4.38).

If $k_2^{(j)}=1/2$, then $k^{(j)}=k^{(j'')}$ and we have two cases, either $m_2=0$ or $m_2=1$. For $m_2=0$ we get $-k_2^{(j)}+m_2=-1/2=k_2^{(j)}-1$ so that

$$\omega_{n_j}(k^{(j)}+\varepsilon p-m)\hat{b}_j(p)=\omega_{n_j}\left(k^{(j'')}+\varepsilon\begin{pmatrix}p_1\\-p_2\end{pmatrix}-m''\right)\bar{\hat{b}}_{j''}(p_1,-p_2) \tag{4.41}$$

with $m''=(m_1,1)\in M_{j''}=M_j$. Finally, for $m_2=1$ one has $-k_2^{(j)}+m_2=1/2=k_2^{(j)}-0$ so that (4.41) holds with $m''=(m_1,0)\in M_{j''}=M_j$.

The last term on the right hand side of (4.25) commutes with S because it is a symmetric cut-off of the nonlinear part of the CME operator.

The term $\tilde{g}_{n_j}(k)$ is more complicated because it involves $p_{n_j}(k;x)$ in the convolutions instead of the $k-$independent $p_{n_j}(k^{(j)};x)$. After solving (4.14) for the regular part $\tilde{\psi}$ in terms of $\hat{\mathbf{B}}$, we obtain that $\tilde{g}_{n_j}(k)$ consists of the terms

$$\hat{I}^{\alpha\beta\gamma j}=\sum_{m\in M_j}\chi_{D_j}(k+m)\langle(\xi_\alpha *_B \xi_\beta *_B \xi_\gamma^c)(k;\cdot),p_{n_j}(k;\cdot)\rangle_{L^2(\mathbb{P}^2)} \tag{4.42}$$

and of higher order convolutions of the ξ_i's coming from $\tilde{\psi}$, which starts with cubic convolutions. Here $\xi_\alpha=p_{n_\alpha}(k;x)\sum_{m\in M_\alpha}\chi_{D_\alpha}(k+m)\hat{B}_\alpha\left(\frac{k+m-k^{(\alpha)}}{\varepsilon}\right)$ and $\xi_\alpha^c=\overline{p_{n_\alpha}}(k;x)\sum_{m\in M_\alpha}\chi_{-D_\alpha}(k-m)\hat{\bar{B}}_\alpha\left(\frac{k-m+k^{(\alpha)}}{\varepsilon}\right)$. It suffices to show reversibility for the simplest convolution type (4.42). We have

$$\hat{I}^{\alpha\beta\gamma j}=\sum_{m\in M_j}\sum_{n,o,q\in\mathcal{A}_{\alpha,\beta,\gamma,j,m}}\tilde{w}_{noqm}^{\alpha\beta\gamma j}(k)$$

with

$$\begin{aligned}\tilde{w}_{noqm}^{\alpha\beta\gamma j}(k)=&\chi_{D_j}(k+m)\int_{\mathbb{T}^2}\int_{\mathbb{T}^2}\chi_{D_\alpha}(k-r+n)\hat{B}_\alpha\left(\tfrac{k-r+n-k^{(\alpha)}}{\varepsilon}\right)\chi_{D_\beta}(r-s+o)\hat{B}_\beta\left(\tfrac{r-s+o-k^{(\beta)}}{\varepsilon}\right)\times\\&\times\ \chi_{-D_\gamma}(s-q)\hat{\bar{B}}_\gamma\left(\tfrac{s-q+k^{(\gamma)}}{\varepsilon}\right)\langle p_{n_\alpha}(k-r;\cdot)p_{n_\beta}(r-s;\cdot)\overline{p_{n_\gamma}}(s;\cdot),p_{n_j}(k;\cdot)\rangle_{L^2(\mathbb{P}^2)}\,\mathrm{d}s\,\mathrm{d}r.\end{aligned}$$

Employing the same change of variables as in (3.24) and using (3.29), we get

$$\hat{\mu}_{noqm}^{\alpha\beta\gamma j}(p) := \tilde{w}_{noqm}^{\alpha\beta\gamma j}(\varepsilon p + k^{(j)} - m) = \chi_{D_{\varepsilon r-1}}(p) \int\limits_{\Omega_1}\int\limits_{\Omega_2} \chi_{D_{\varepsilon r-1}}(p-\tilde{r})\hat{B}_\alpha(p-\tilde{r})\chi_{D_{\varepsilon r-1}}(\tilde{r}-\tilde{s})\hat{B}_\beta(\tilde{r}-\tilde{s})\times$$
$$\times\ \chi_{D_{\varepsilon r-1}}(\tilde{s})\hat{\bar{B}}_\gamma(\tilde{s})\langle p_{n_\alpha}(k^{(\alpha)} - n + \varepsilon(p-\tilde{r});\cdot)p_{n_\beta}(k^{(\beta)} - o + \varepsilon(\tilde{r}-\tilde{s});\cdot)\times$$
$$\times\ \overline{p_{n_\gamma}}(-k^{(\gamma)} + q + \varepsilon\tilde{s};\cdot), p_{n_j}(\varepsilon p + k^{(j)} - m;\cdot)\rangle_{L^2(\mathbb{P}^2)}\,\mathrm{d}\tilde{s}\,\mathrm{d}\tilde{r}, \tag{4.43}$$

where $\Omega_1 = D_{2\varepsilon r-1} \cap \frac{\mathbb{T}^2 - k^{(\beta)} + k^{(\gamma)} + o - q}{\varepsilon}$ and $\Omega_2 = D_{\varepsilon r-1} \cap \frac{\mathbb{T}^2 + k^{(\gamma)} - q}{\varepsilon}$.

For the symmetry S we need to consider $\hat{\bar{\mu}}_{noqm}^{\alpha\beta\gamma j}(-p_1, p_2) = \bar{\hat{\mu}}_{noqm}^{\alpha\beta\gamma j}(p_1, -p_2)$. Assuming that **b** (and **A**) satisfies (4.38), we get that also **B** does. Thus, using $\hat{\bar{B}}_i\left(\left(\begin{smallmatrix} p_1 \\ -p_2 \end{smallmatrix}\right) - \tilde{r}\right) = \hat{\bar{B}}_i\left(\left(\begin{smallmatrix} -p_1 \\ p_2 \end{smallmatrix}\right) + \tilde{r}\right) = s_1\hat{B}_{i''}\left(p - \left(\begin{smallmatrix} \tilde{r}_1 \\ -\tilde{r}_2 \end{smallmatrix}\right)\right)$ and the change of variables $\tilde{r} \to (\tilde{r}_1, -\tilde{r}_2)$, $\tilde{s} \to (\tilde{s}_1, -\tilde{s}_2)$, we get

$$\bar{\hat{\mu}}_{noqm}^{\alpha\beta\gamma j}(p_1, -p_2) = \chi_{D_{\varepsilon r-1}}(p)s_1^3 \int\limits_{\Omega_1'}\int\limits_{\Omega_2'} \chi_{D_{\varepsilon r-1}}(p-\tilde{r})\hat{B}_{\alpha''}(p-\tilde{r})\chi_{D_{\varepsilon r-1}}(\tilde{r}-\tilde{s})\hat{B}_{\beta''}(\tilde{r}-\tilde{s})\chi_{D_{\varepsilon r-1}}(\tilde{s})\hat{\bar{B}}_{\gamma''}(\tilde{s})\times$$
$$\times \left\langle \overline{p_{n_\alpha}}\left(k^{(\alpha)} - n + \varepsilon\left(\left(\begin{smallmatrix} p_1 \\ -p_2 \end{smallmatrix}\right) - \left(\begin{smallmatrix} \tilde{r}_1 \\ -\tilde{r}_2 \end{smallmatrix}\right)\right);\cdot\right)\overline{p_{n_\beta}}\left(k^{(\beta)} - o + \varepsilon\left(\left(\begin{smallmatrix} \tilde{r}_1 \\ -\tilde{r}_2 \end{smallmatrix}\right) - \left(\begin{smallmatrix} \tilde{s}_1 \\ -\tilde{s}_2 \end{smallmatrix}\right)\right);\cdot\right)\times$$
$$\times p_{n_\gamma}\left(-k^{(\gamma)} + q + \varepsilon\left(\begin{smallmatrix} \tilde{s}_1 \\ -\tilde{s}_2 \end{smallmatrix}\right);\cdot\right), \overline{p_{n_j}}\left(k^{(j)} - m + \varepsilon\left(\begin{smallmatrix} p_1 \\ -p_2 \end{smallmatrix}\right);\cdot\right)\right\rangle_{L^2(\mathbb{P}^2)} \mathrm{d}\tilde{s}\,\mathrm{d}\tilde{r}, \tag{4.44}$$

where $\Omega_1' = D_{2\varepsilon r-1} \cap \varepsilon^{-1}\left(\mathbb{T}^2 - \begin{pmatrix} k_1^{(\beta)} \\ -k_2^{(\beta)} \end{pmatrix} + \begin{pmatrix} k_1^{(\gamma)} \\ -k_2^{(\gamma)} \end{pmatrix} + \left(\begin{smallmatrix} o_1 \\ -o_2 \end{smallmatrix}\right) - \left(\begin{smallmatrix} q_1 \\ -q_2 \end{smallmatrix}\right)\right)$, and $\Omega_2' = D_{\varepsilon r-1} \cap \varepsilon^{-1}\left(\mathbb{T}^2 + \begin{pmatrix} k_1^{(\gamma)} \\ -k_2^{(\gamma)} \end{pmatrix} - \left(\begin{smallmatrix} q_1 \\ -q_2 \end{smallmatrix}\right)\right)$. Finally, we use the Bloch function symmetry $\overline{p_n}((k_1, -k_2); x) = p_n(k; (2\pi - x_1, x_2))$. In our case this means

$$\overline{p_{n_\alpha}}\left(k^{(\alpha)} - n + \varepsilon\left(\left(\begin{smallmatrix} p_1 \\ -p_2 \end{smallmatrix}\right) - \left(\begin{smallmatrix} \tilde{r}_1 \\ -\tilde{r}_2 \end{smallmatrix}\right)\right); x\right) = p_{n_\alpha}\left(\begin{pmatrix} k_1^{(\alpha)} \\ -k_2^{(\alpha)} \end{pmatrix} - \left(\begin{smallmatrix} n_1 \\ -n_2 \end{smallmatrix}\right) + \varepsilon(p - \tilde{r}); \left(\begin{smallmatrix} 2\pi - x_1 \\ x_2 \end{smallmatrix}\right)\right).$$

If $k_2^{(\alpha)} \in (-1/2, 1/2)$, then $n_2 = 0$ and we have for $k^{(\alpha'')} = \begin{pmatrix} k_1^{(\alpha)} \\ -k_2^{(\alpha)} \end{pmatrix}$ and $n_{\alpha''} = n_\alpha$

$$p_{n_\alpha}\left(\begin{pmatrix} k_1^{(\alpha)} \\ -k_2^{(\alpha)} \end{pmatrix} - \left(\begin{smallmatrix} n_1 \\ -n_2 \end{smallmatrix}\right) + \varepsilon(p - \tilde{r}); \left(\begin{smallmatrix} 2\pi - x_1 \\ x_2 \end{smallmatrix}\right)\right) = p_{n_{\alpha''}}\left(k^{(\alpha'')} - n + \varepsilon(p - \tilde{r}); \left(\begin{smallmatrix} 2\pi - x_1 \\ x_2 \end{smallmatrix}\right)\right).$$

If $k_2^{(\alpha)} = 1/2$, then $n_2 \in \{0, 1\}$ and $\begin{pmatrix} k_1^{(\alpha)} \\ -k_2^{(\alpha)} \end{pmatrix} - \left(\begin{smallmatrix} n_1 \\ -n_2 \end{smallmatrix}\right) = k^{(\alpha)} - \tilde{n}$ for some $\tilde{n} \in M_\alpha$ with $\tilde{n} \neq n$. Thus

$$p_{n_\alpha}\left(\begin{pmatrix} k_1^{(\alpha)} \\ -k_2^{(\alpha)} \end{pmatrix} - \left(\begin{smallmatrix} n_1 \\ -n_2 \end{smallmatrix}\right) + \varepsilon(p - \tilde{r}); \left(\begin{smallmatrix} 2\pi - x_1 \\ x_2 \end{smallmatrix}\right)\right) = p_{n_\alpha}\left(k^{(\alpha)} - \tilde{n} + \varepsilon(p - \tilde{r}); \left(\begin{smallmatrix} 2\pi - x_1 \\ x_2 \end{smallmatrix}\right)\right)$$

so that $\alpha'' = \alpha$.

Performing the same analysis for the Bloch functions $p_{n_\beta}, p_{n_\gamma}$ and p_{n_j} in (4.44) and using $s_1^3 = s_1$, we get

$$\hat{\mu}^{\alpha\beta\gamma j}_{noqm}(-p_1,p_2) = s_1 \hat{\mu}^{\alpha''\beta''\gamma'' j''}_{\tilde{n}\tilde{o}\tilde{q}\tilde{m}}(p),$$

where some of the doubly primed and 'tilded' indices may equal the bare ones and where $\tilde{n} \in M_\alpha, \tilde{o} \in M_\beta, \tilde{q} \in M_\gamma$, and $\tilde{m} \in M_j$. After the sum in n, o, q and m, we get

$$\hat{I}^{\alpha\beta\gamma j}(-p_1,p_2) = s_1 \hat{\bar{I}}^{\alpha''\beta''\gamma'' j''}(p).$$

In conclusion $\mathbf{N} = S\mathbf{N}$ and (4.37) holds.

With Ker $\mathbf{J} = \{\partial_{y_1}\mathbf{A}, \partial_{y_2}\mathbf{A}, \mathrm{i}\mathbf{A}\}$ the operator $\hat{\mathbf{J}}_\varepsilon$ has an $\mathcal{O}(1)$ bounded inverse on X_{rev}, i.e., with a bound independent of ε. Moreover, $\|\hat{\mathbf{J}}_\varepsilon^{-1}\hat{\mathbf{R}}\|_{L^2_s(D_{\varepsilon^{r-1}})} \leq C\|\hat{\mathbf{R}}\|_{L^2_{s-2}(D_{\varepsilon^{r-1}})}$, which is why, e.g., estimating $\varepsilon^{-2}q_j^{(3)}(\varepsilon p)\hat{B}_j(p)$ in L^2_{s-2} is sufficient, cf. (4.27). Thus, (4.32) can now be solved producing

Theorem 4.9 *Let $s > 1$ and let V be even in x_1 as well as x_2. There exists an $\varepsilon_0 > 0$ such that for all $0 < \varepsilon < \varepsilon_0$ the following holds. Let $\omega = \omega_* + \varepsilon^2\Omega$ be in a band gap, let $\mathbf{A}$ be a reversible non-degenerate solution of the CME (3.33) with $\hat{\mathbf{A}} \in L^2_q(\mathbb{R}^2)$ for all $q \geq 0$, and let $0 < r < 2/3$. Then there exists a $C > 0$ and a solution $\hat{\mathbf{B}}$ of the eCME such that*

$$\|\hat{\mathbf{B}} - \hat{\mathbf{A}}\|_{L^2_s(D_{\varepsilon^{r-1}})} \leq C\varepsilon^{\tilde{r}} \qquad \tilde{r} = \min\{r, 2-2r\}. \tag{4.45}$$

Corollary 4.10 *The solution ϕ constructed in Theorems 4.6 and 4.9 is a localized solution of (1.2), it is symmetric according to (4.33) or (4.34), and*

$$\|\phi(\cdot) - \varepsilon\sum_{j=1}^N A_j(\varepsilon\cdot)u_{n_j}(k^{(j)};\cdot)\|_{H^s(\mathbb{R}^2)} \leq C\varepsilon^{\tilde{r}}. \tag{4.46}$$

Proof. We first show that the reversibility of $\hat{\mathbf{B}}$ and the symmetry of Bloch functions provides a $\varepsilon^{-1}\eta^{(0)}_{\mathrm{LS}}$ that satisfies (4.33) or (4.34). Let us work out explicitly only the first symmetry in (4.33). Recall that

$$\eta^{(0)}_{\mathrm{LS}}(x) = \int_{\mathbb{T}^2} e^{\mathrm{i}k\cdot x}\sum_{j=1}^N p_{n_j}(k;x)\sum_{m\in M_j}\chi_{D_j}(k+m)\hat{B}_j\left(\tfrac{k+m-k^{(j)}}{\varepsilon}\right)\mathrm{d}k.$$

Using $p_{n_j}(k;(-x_1,x_2)) = \overline{p_{n_{j''}}}((k_1,-k_2);x)$ with $n_{j''} = n_j$, $\hat{B}_j(p) = s_1\hat{\bar{B}}_{j''}(-p_1,p_2) = s_1\bar{\hat{B}}_{j''}(p_1,-p_2)$ together with the definition $k^{(j'')} = (k_1^{(j)}, -k_2^{(j)})$ if $k_2^{(j)} < \frac{1}{2}$ and $k^{(j'')} = k^{(j)}$ if $k_2^{(j)} = \frac{1}{2}$, we get for $k_2^{(j)} < \frac{1}{2}$

$$\eta^{(0)}_{\mathrm{LS}}(-x_1,x_2) = s_1\int_{\mathbb{T}^2}\overline{e^{\mathrm{i}(k_1,-k_2)\cdot x}}\sum_{j=1}^N\overline{p_{n_{j''}}}((k_1,-k_2);x)\sum_{m\in M_j}\chi_{D_j}(k+m)\bar{\hat{B}}_{j''}\left(\frac{\binom{k_1}{-k_2}+\binom{m_1}{-m_2}-k^{(j'')}}{\varepsilon}\right)\mathrm{d}k.$$

Because $m_2 = 0\ \forall m \in M_j$ when $k_2^{(j)} < \frac{1}{2}$ and since $M_j = M_{j''}$, and $\chi_{D_j}(k+m) = \chi_{D_{j''}}\left(\binom{k_1}{-k_2} + m\right)$, we have after the change of variables $k \to (k_1, -k_2)$

$$\eta_{\rm LS}^{(0)}(-x_1,x_2) = s_1 \int_{\mathbb{T}^2} \overline{e^{ik\cdot x}} \sum_{j=1}^{N} \overline{p_{n_{j''}}}(k;x) \sum_{m\in M_{j''}} \chi_{D_{j''}}(k+m)\bar{\hat{B}}_{j''}\left(\tfrac{k+m-k^{(j'')}}{\varepsilon}\right) \mathrm{d}k = s_1 \overline{\eta_{\rm LS}^{(0)}}(x).$$

For $k_2^{(j)} = \frac{1}{2}$ we have

$$\eta_{\rm LS}^{(0)}(-x_1,x_2) = s_1 \int_{\mathbb{T}^2} \overline{e^{i(k_1,-k_2)\cdot x}} \sum_{j=1}^{N} \overline{p_{n_{j''}}}((k_1,-k_2);x) \sum_{m\in M_j} \chi_{D_j}(k+m)\bar{\hat{B}}_{j''}\left(\frac{\binom{k_1}{-k_2}+\binom{m_1}{-m_2}-k^{(j'')}+\binom{0}{1}}{\varepsilon}\right) \mathrm{d}k$$

because $\begin{pmatrix} k_1^{(j)} \\ -k_2^{(j)} \end{pmatrix} = k^{(j'')} - \binom{0}{1}$. Next, using firstly $\binom{m_1}{-m_2} + \binom{0}{1} = \tilde{m}$ with some $\tilde{m} = M_j, \tilde{m} \neq m$, and secondly $\chi_{D_j}(k+m) = \chi_{D_{j''}}\left(\binom{k_1}{-k_2} + \tilde{m}\right)$, we arrive (after the change of variables $k \to (k_1, -k_2)$) at

$$\eta_{\rm LS}^{(0)}(-x_1,x_2) = s_1 \int_{\mathbb{T}^2} \overline{e^{ik\cdot x}} \sum_{j=1}^{N} \overline{p_{n_{j''}}}(k;x) \sum_{\tilde{m}\in M_{j''}} \chi_{D_{j''}}(k+\tilde{m})\bar{\hat{B}}_{j''}\left(\tfrac{k+\tilde{m}-k^{(j'')}}{\varepsilon}\right) \mathrm{d}k = s_1 \overline{\eta_{\rm LS}^{(0)}}(x).$$

Next, ψ in (4.14) inherits the symmetry, so that ϕ is symmetric. The estimate (4.46) follows from the triangle inequality with (4.28) and (4.45). □

Remark 4.11 In (4.45) we use $\tilde{r} = \min\{r, 2-2r\}$ although below (4.24) we defined $\tilde{r} := \min\{r, 2-2r, 1\}$. This is because for $0 < r < 2/3$ we have $\min\{r, 2-2r, 1\} = \min\{r, 2-2r\}$.

The optimal value of $\tilde{r}$ is $\tilde{r} = 2/3$ attained at $r = 2/3$. Based on the formal asymptotic expansion in (3.13) and (3.20), we see that the next order term $\tilde{\psi}^{(1)}$ (just like $\tilde{\psi}^{(0)}$) consists of terms of the type $\hat{F}\left(\frac{k-k^{(j)}}{\varepsilon}\right) q(k^{(j)};x)$, where F is an envelope and q a periodic carrier wave. $\psi^{(1)}$, therefore, consists of terms $\varepsilon^2 F(\varepsilon x) q(k^{(j)};x) e^{i2\pi k^{(j)}\cdot x}$ and $\|\psi^{(1)}\|_{H^s(\mathbb{R}^2)} = \mathcal{O}(\varepsilon)$. As a result, the formal asymptotics predict ε^1 convergence. Thus, while the estimate (4.46) guarantees convergence of the CME approximation, it does not appear to be sharp. But if all third derivatives of ω_{n_j} vanish at $k^{(j)}$, like for separable potentials [13], we have $\tilde{r} = \min\{2r, 2-2r, 1\}$ with $0 < r < 2/3$ and the optimal value is $\tilde{r} = 1$ attained at $r = 1/2$. It is, however, unclear which non-separable potentials result in vanishing third derivatives of the bands at gap edge extrema. ⌟

Remark 4.12 As said in the previous remark, the formal asymptotics predict that the error ψ in (4.8) has the form

$$\psi(x) = \varepsilon^2 F(\varepsilon x) w(x)$$

with $F \in H^q(\mathbb{R}^2)$ for all $q \geq 0$ and $w(x) \in C_b^{\lceil s \rceil}(\mathbb{R}^2)$ with $w(2\pi, x_2) = e^{2\pi i k_1} w(0, x_2)$, $w(x_1, 2\pi) = e^{2\pi i k_2} w(x_1, 0)$. In this case

$$\|\psi\|_{H^s} = \mathcal{O}(\varepsilon^\beta) \text{ implies } \|\psi\|_{L^\infty} = \mathcal{O}(\varepsilon^{\beta+1}).$$

To see this, assume that $\|\psi\|_{L^\infty} \geq C_1 \varepsilon^\alpha$ with $\alpha < 1+\beta$, i.e., $|\psi(x_0)| \geq C_1 \varepsilon^\alpha$ for some $x_0 \in \mathbb{R}^2$. Since ψ is continuous, this implies $|\psi(x)| \geq \frac{C_1}{2} \varepsilon^\alpha$ in some neigborhood of x_0 of diameter $\delta_1 > 0$, say. Moreover, since F is continuous, $|\varepsilon^2 F(\varepsilon x)| \geq \frac{C_1}{2}\varepsilon^\alpha$ for $|x| \leq \delta_2/\varepsilon$. Therefore, for ε sufficiently small, $|\psi(x)| \geq \frac{C_1}{4}\varepsilon^\alpha$ for x in $\mathcal{O}(\delta_1)$ wide neighborhoods of $x_j = x_0 + (2j_1\pi, 2j_2\pi)$, where $j_1, j_2 \in \mathbb{Z}$ and $|j_1|, |j_2| \leq C_2/\varepsilon$, see Fig.7 for a 1D sketch. Since in 2D there are at least $C_3\varepsilon^{-2}$ such neighborhoods, we obtain

$$\|\psi\|_{H^s} \geq \frac{C_3 C_1 \varepsilon^\alpha \delta}{4\varepsilon} > C\varepsilon^{\alpha-1} \tag{4.47}$$

which contradicts $\|\psi\|_{H^s} \leq C\varepsilon^\beta$ as $\varepsilon \to 0$ due to $\alpha - 1 < \beta$.

To make this argument rigorous for the error $\psi = \phi - \varepsilon \sum_{j=1}^N A_j(\varepsilon\cdot)u_{n_j}(k^{(j)};\cdot)$, we could split off the next term in the formal asymptotic expansion and show that the remainder is of higher order, i.e., $\psi = \varepsilon^2\psi^{(1)} + \varepsilon^3\psi_*$. Then we can estimate $\|\varepsilon^3\psi_*\|_{H^s} = \mathcal{O}(\varepsilon^{3-2r})$ using the analysis from §4.2, but here we refrain from these tedious calculations. ⌋

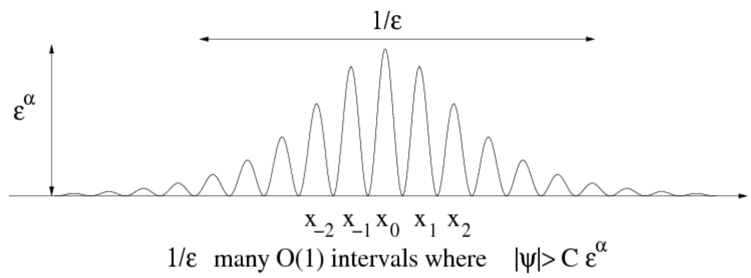

Figure 7: Sketch of the argument for $\|\psi\|_{H^s} = \mathcal{O}(\varepsilon^\beta) \Rightarrow \|\psi\|_{L^\infty} = \mathcal{O}(\varepsilon^{\beta+1})$.

5 Numerical Results on Reversible Gap Solitons

We numerically compute some representative cases of gap solitons and their asymptotic envelope approximations $\varepsilon\phi^{(0)}(x) = \varepsilon\sum_{j=1}^N A_j(\varepsilon x)u_{n_j}(k^{(j)};x)$. We do not attempt to provide an exhaustive study of possible GS solutions but rather select only several cases to corroborate our analysis. Namely, we select GSs bifurcating from the edges s_2 and s_5. The latter case is of particular interest as it features a situation whose occurrence is impossible for separable potentials $V(x)$. To our knowledge this case has not been studied before and the presented GSs are novel. We also check the reversibility and non-degeneracy conditions which are sufficient for persistence, see §4.3. In addition, we compute the convergence rate in ε, i.e., in the square root of the distance to the gap edge, of the error $\|\phi_{\rm GS}^{\rm num} - \varepsilon\phi^{(0)}\|_{H^2}$.

A 4th order centered finite difference discretization is used for (1.2). The computational domain is a square $x \in [-D_{\rm GS}/2, D_{\rm GS}/2]^2$ selected large enough so that the asymptotic approximation $\varepsilon\phi^{(0)}(x)$ of the GS is well-decayed at the boundary and zero Dirichlet boundary conditions are then used. Equation (1.2) is then solved via Newton's iteration using $\varepsilon\phi^{(0)}$ as the initial guess. The computational domain is in practice reduced to its quarter using the corresponding reversibility symmetry.

5.1 Gap Solitons near $\omega = s_2$

Near the edge ω=s_2 we limit our attention to real, even GSs and to symmetric vortices of charge 1. As the coupled mode system near $\omega = s_2$ is a scalar nonlinear Schrödinger equation, see §3.2.2, one can search for solutions of the form $A(y) = R(r)e^{\mathrm{i}m\theta}$, where $r = \frac{1}{\sqrt{\alpha}}\sqrt{y_1^2 + y_2^2}, \theta = \arg(y_1 + \mathrm{i}y_2)$, and $m \in \mathbb{N}$. We choose $R > 0$ and $m = 0$ corresponding to the so called Townes soliton, and $m = 1$ corresponding to a vortex of charge 1. The function $R(r)$ satisfies the ODE

$$R'' + \frac{1}{r}R' + \Omega R - \frac{m^2}{r^2}R - \sigma\gamma R^3 = 0, \tag{5.1}$$

where $R(0) > 0$, $R'(0) = 0$ for $m = 0$ and $R(0) = 0$, $R'(0) > 0$ for $m = 1$. For $m \neq 0$ the initial-value problem for the ODE (5.1) is ill-posed but can be turned into a well-posed one via the transformation $Q = r^{-m}R(r)$ leading to

$$Q'' + \frac{2m+1}{r}Q' + \Omega Q - \sigma\gamma r^{2m}Q^3 = 0 \tag{5.2}$$

with $Q(0) > 0$, such that $R(r) \sim r^{|m|}$ as $r \to 0$. We solve equation (5.2) numerically via a shooting method searching for $Q(r)$ vanishing as $r \to \infty$.

For $m = 0$ we have the reversibility $A(-y_1, y_2) = A(y_1, -y_2) = A(y)$, which is the same as (4.35) with $s_1 = s_2 = 1$ since A is real. The non-degeneracy condition on $\mathbf{J}$ in Theorem 4.9 is known to be satisfied by the positive ground state A [24, 9] and conditions of this theorem are, therefore, satisfied.

Figure 8 shows the profiles of the envelope A, of the asymptotic approximation $\varepsilon\phi^{(0)}(x)$ and of the GS $\phi(x)$ computed via Newton's iteration on (1.2). A GS deep inside the gap (s_2, s_3) obtained via a homotopy continuation in ω from the $\phi(x)$ in Fig. 8 is plotted in Fig. 9(a), while (b) shows the ε-convergence of the approximation error. Here the $\varepsilon^{1.46}$ convergence rate is better than the estimate proved in Corollary 4.10 and even better than the rate ε^1 predicted by formal asymptotics in Rem. 4.11.

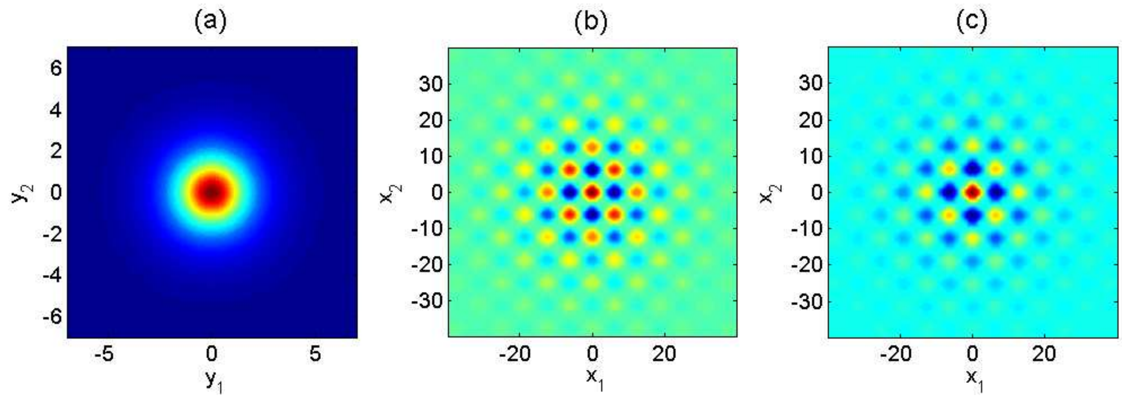

Figure 8: Profiles of the even real GS at $\omega = s_2 + \varepsilon^2\Omega, \varepsilon = 0.1, \Omega = 1$. (a) $A(y)$; (b) the corresponding leading-order GS approximation $\varepsilon A(y)u_1(M; x)$; (c) the numerically computed GS at $\omega = s_2 + \varepsilon^2\Omega$.

For $m = 1$ the solution is complex and we have the reversibility $A(-y_1, y_2) = -A(y_1, -y_2) = -\bar{A}(y)$, which is (4.35) with $s_1 = -s_2 = -1$. Figure 10 shows the modulus and phase of the envelope A, of the asymptotic approximation $\varepsilon\phi^{(0)}(x) = \varepsilon A(\varepsilon x)u_1(M; x)$ and of the computed GS. The non-degeneracy

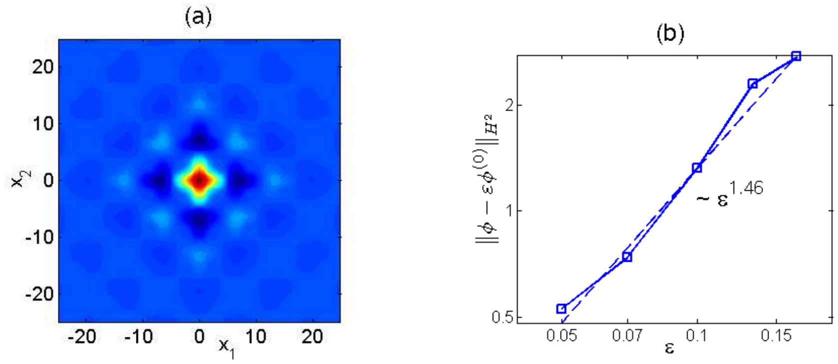

Figure 9: (a) Profile of a GS corresponding to the even real family that bifurcates from $\omega = s_2$ in Fig. 8. The plotted GS is deep inside the gap (s_2, s_3) at $\omega \approx 1.78$ (corresponding to $\varepsilon \approx 0.28$). (b) ε-convergence of the error $\|\phi - \varepsilon\phi^{(0)}\|_{H^2(\mathbb{R}^2)}$.

of the envelope is illustrated in Fig. 11(a), which plots the 4 smallest eigenvalues (in modulus) of the Jacobian operator $\mathbf{J}$ of the CMEs evaluated at the vortex A: 3 eigenvalues converge to zero as the computational domain size grows while the fourth one stays bounded away from zero. Figure 11(b) presents the ε-convergence of the approximation error $\|\phi - \varepsilon\phi^{(0)}\|_{H^2(\mathbb{R}^2)}$. The resulting convergence is very close to ε^1, which is the prediction based on formal asymptotics.

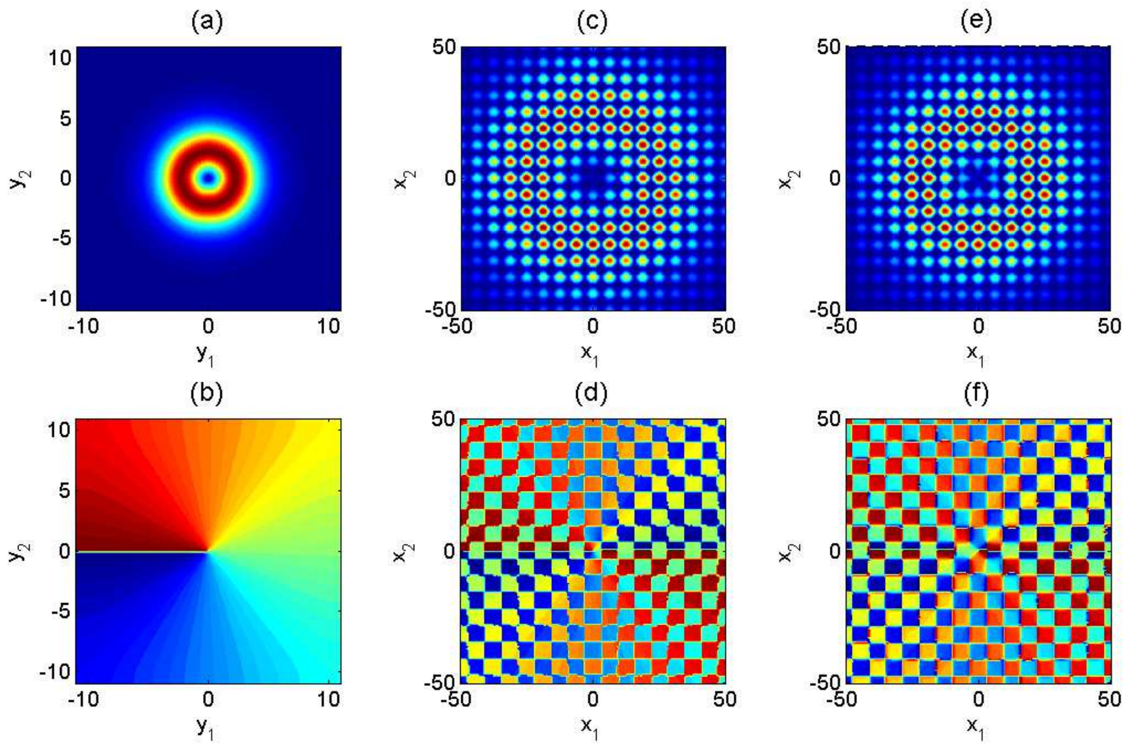

Figure 10: Profiles of the vortex GS at $\omega = s_2 + \varepsilon^2\Omega, \varepsilon = 0.09, \Omega = 1$. (a) and (b) modulus and phase of $A(y)$ resp.; (c) and (d) modulus and phase of the corresponding leading-order GS approximation $\varepsilon A(y)u_1(M;x)$ resp.; (e) and (f) modulus and phase of the numerically computed GS at $\omega = s_2 + \varepsilon^2\Omega$ resp.

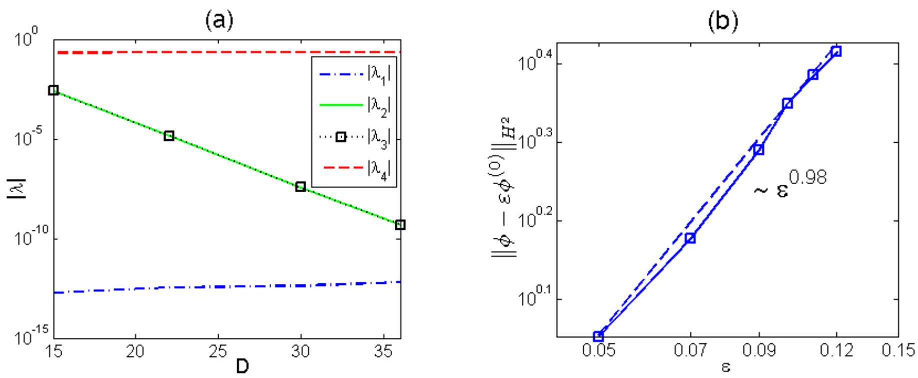

Figure 11: (a) The four smallest eigenvalues of the Jacobian $\mathbf{J}$ in Theorem 4.9 at the solution $\mathbf{A}$ in Fig. 10 (a-d) for a range of sizes of the computational domain. (b) ε-convergence of the error $\|\phi - \varepsilon\phi^{(0)}\|_{H^2(\mathbb{R}^2)}$.

5.2 Gap Solitons near $\omega = s_5$

We limit our attention here to gap solitons with real positive envelopes satisfying the symmetries $A_1 = A_3, A_2 = A_4$ and $A_1(-y_1, y_2) = A_1(y_1, -y_2) = A_1(-y_2, y_1) = A_2(y_1, y_2)$, which is (4.36) with $s_1 = s_2 = 1$ for each A_j. Such solutions of the CME system (3.39) can be found by first setting $\alpha_2 = 0$ and computing radially symmetric positive solutions $A_1 = A_2 = A_3 = A_4 = R(r)$, where $r = \frac{1}{\sqrt{\alpha_1}}\sqrt{y_1^2 + y_2^2}$, via a shooting method and then performing a homotopy continuation in α_2 on the system of the first two equations in (3.39) employing the symmetry $A_1 = A_3, A_2 = A_4$ up to the original value $\alpha_2 = 0.096394$.

We normalize the Bloch functions $v_1(x) := u_6((k_c, k_c); x), v_2(x) := u_6((-k_c, k_c); x)$, $v_3(x) := u_6((-k_c, -k_c); x)$ and $v_4(x) := u_6((k_c, -k_c); x)$ so that

$$v_2(-x_1, x_2) = v_1(x_1, x_2),\ v_3(x_1, -x_2) = v_2(x_1, x_2) \text{ and } v_4(-x_1, x_2) = v_3(x_1, x_2) \quad \text{(see §2)}. \tag{5.3}$$

This normalization implies that $\varepsilon\phi^{(0)}(x)$ is real and even in both variables. These symmetries are used to reduce the computational domain to one quadrant and restrict to the real arithmetic.

Figure 12 shows the envelope $A_1(y)$, the GS approximation $\varepsilon\phi^{(0)}$ and the computed GS ϕ. The envelope $A_1(y)$ in Fig. 12 is not radially symmetric due to the mixed derivative $\partial_{y_1}\partial_{y_2}$ in (3.39), but looks radially symmetric because the coefficient α_2 is relatively small ($\alpha_2 \approx 0.0964$). Profiles of $A_2, \ldots, A_4$ are not plotted as they can be obtained from A_1 via the above mentioned symmetries.

A closer look at the structure of ϕ near the origin, an illustration of the non-degeneracy of $\mathbf{A}$, and the ε-convergence of the approximation error are provided in Fig. 13. The obtained rate is about $\varepsilon^{0.94}$, which is once again close to the rate ε^1 predicted by the formal asymptotics.

6 Conclusions

We have derived systems of Coupled Mode Equations (CME) which approximate stationary gap solitons (GSs) of the 2D periodic Nonlinear Schrödinger Equation/Gross Pitaevskii equation near a band edge.

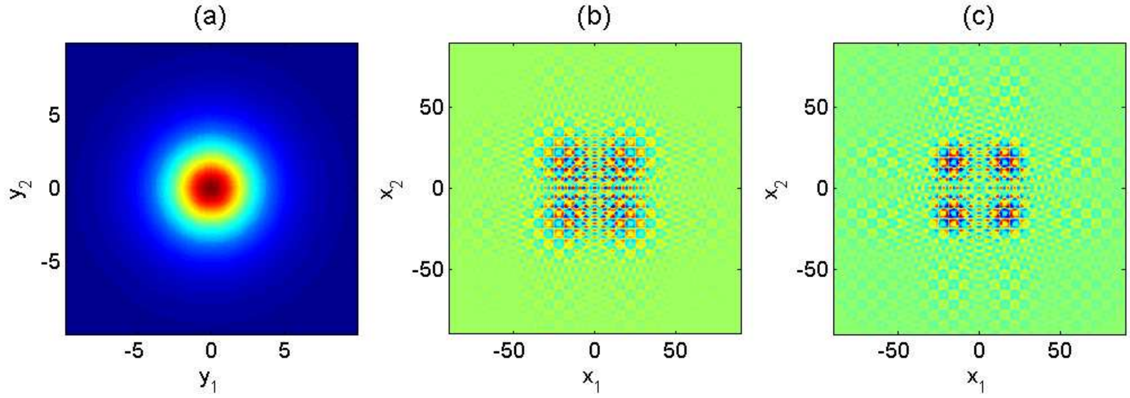

Figure 12: Profiles of the even, real GS at $\omega = s_5 + \varepsilon^2\Omega, \varepsilon = 0.1, \Omega = -1$. (a) $A_1(y)$; (b) the corresponding leading-order GS approximation $\varepsilon \sum_{j=1}^{4} A_j(y)v_j(x)$; (c) the numerically computed GS at $\omega = s_5 + \varepsilon^2\Omega$.

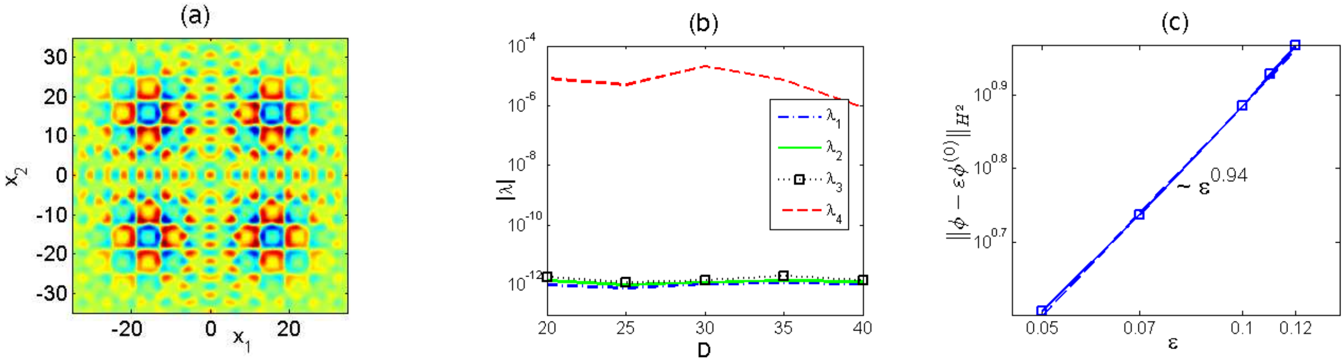

Figure 13: (a) Detail of the profile in Fig. 12 (c). (b) the non-degeneracy of **A**. (c) ε-convergence of the error $\|\phi - \varepsilon\phi^{(0)}\|_{H^2(\mathbb{R}^2)}$.

In contrast to [13] we do not assume separability of the periodic potential V. While in the case of a separable $V(x)$ [13] the derivation is possible in physical variables, here in general it has to be performed in Bloch variables. We have rigorously proved via the Lyapunov-Schmidt reduction that reversible non-degenerate solitons of the CME yield GSs of the Gross-Pitaevskii equation. We have also provided an $H^s(\mathbb{R}^2)$, $s > 1$ estimate on the approximation error showing that it is $\mathcal{O}(\varepsilon^{2/3})$ for GSs with the spectral parameter $\mathcal{O}(\varepsilon^2)$ close to the band edge. Our analysis requires some smoothness of V, namely $V \in H_{\text{loc}}^{\lceil s\rceil-1+\delta}(\mathbb{R}^2), \delta > 0$, and, in the persistence step, evenness of V. The analysis has been corroborated by numerical examples including one which features novel GSs bifurcating from a band edge Bloch wave located outside the set of vertices of the first Brillouin zone, which is impossible in the case of separable potentials.

Acknowledgement. The work of T. Dohnal is supported by the Humboldt Research Fellowship. The authors wish to thank Dmitry Pelinovsky and Guido Schneider for stimulating discussions.

References

[1] F. Abdullaev, A. Abdumalikov, and R. Galimzyanov, "Gap solitons in Bose-Einstein condensates in linear and nonlinear optical lattices," Physics Letters A **367** 149-155 (2007)

[2] A.B. Aceves, B. Costantini and C. De Angelis, "Two-dimensional gap solitons in a nonlinear periodic slab waveguide," J. Opt. Soc. Am. B **12**, 1475–1479 (1995)

[3] A.B. Aceves, G. Fibich and B. Ilan, "Gap-soliton bullets in waveguide gratings," Physica D **189**, 277–286 (2004)

[4] R.A. Adams and J.J.F. Fournier, Sobolev Spaces, Academic Press, Amsterdam, 2003.

[5] D. Agueev and D. Pelinovsky, "Modeling of wave resonances in low-contrast photonic crystals," SIAM J. Appl. Math. **65**, 1101–1129 (2005)

[6] N. Aközbek and S. John, "Optical solitary waves in two and three dimensional photonic bandgap structures," Phys. Rev. E **57**, 2287–2319 (1998)

[7] B.B. Baizakov, B.A. Malomed, and M. Salerno, "Multidimensional solitons in periodic potentials," Europhys. Lett., **63** 642-648 (2003)

[8] K. Busch, G. Schneider, L. Tkeshelashvili, and H. Uecker, "Justification of the nonlinear Schrödinger equation in spatially periodic media," Z. Angew. Math. Phys. **57**, 905–939 (2006)

[9] S. M. Chang, S. Gustafson, K. Nakanishi and T. P. Tsai, "Spectra of Linearized Operators for NLS Solitary Waves," SIAM J. Math. Anal., **39** 10701111 (2007).

[10] J. Denzler, "Nonpersistence of breather families for the perturbed sine Gordon equation," Comm. Math. Phys. **158**:2, 397–430 (1993)

[11] C.M. de Sterke, D.G. Salinas and J.E. Sipe, "Coupled-mode theory for light propagation through deep nonlinear gratings," Phys. Rev. E **54**, 1969–1989 (1996)

[12] T. Dohnal and A.B. Aceves, "Optical soliton bullets in (2+1)D nonlinear Bragg resonant periodic geometries," Stud. Appl. Math. **115**, 209–232 (2005)

[13] T. Dohnal, D. Pelinovsky and G. Schneider, "Coupled-mode equations and gap solitons in a two-dimensional nonlinear elliptic problem with a separable periodic potential," J. Nonlin. Sci. **19**, 95–131 (2009)

[14] M.S. Eastham, The Spectral Theory of Periodic Differential Equations, Scottish Academic Press, Edinburgh, 1973

[15] N.K. Efremidis, J. Hudock, D.N. Christodoulides, J.W. Fleischer, O. Cohen, and M. Segev, "Two-Dimensional Optical Lattice Solitons," Phys. Rev. Lett. **91** 213906 (2003).

[16] R. Fischer, D. Träger, D.N. Neshev, A.A. Sukhorukov, W. Krolikowski, C. Denz, and Y.S. Kivshar "Reduced-symmetry two-dimensional solitons in photonic lattices," Phys. Rev. Lett. **96**, 23905 (2006).

[17] J.W. Fleischer, M. Segev, N.K. Efremidis, and D.N. Christodoulides, "Observation of two-dimensional discrete solitons in optically induced nonlinear photonic lattices," Nature **422**, 147–150 (2003).

[18] M. Golubitsky and D.G. Schaeffer, Singularities and Groups in Bifurcation Theory, v. 1, Springer-Verlag, New-York, 1985.

[19] R.H. Goodman, M.I. Weinstein, and P.J. Holmes, "Nonlinear propagation of light in one-dimensional periodic structures," J. Nonlinear. Science **11**, 123–168 (2001)

[20] E.P. Gross, "Structure of a quantized vortex in boson systems," Il Nuovo Cimiento **20** 454 (1961) ["Hydrodynamics of a Superfluid Concentrate," J. Math. Phys **4**, 195–207 (1963)]

[21] J.M. Harrison, P. Kuchment, A. Sobolev, and B. Winn, "On occurrence of spectral edges for periodic operators inside the Brillouin zone," J. Phys. A: Math. Theor. **40**, 7597–7618 (2007)

[22] L. Hörmander, The Analysis of linear partial differential operators III, Springer-Verlag, Berlin, 1985.

[23] Y.V. Kartashov, A.A. Egorov, V.A. Vysloukh, and L. Torner, "Surface vortex solitons," Opt. Express **14** 4049–4057 (2006).

[24] M.K. Kwong, "Uniqueness of of positive solutions of $\Delta u - u + u^p = 0$ in $\mathbb{R}^n$," Arch. Rational Mech. Anal., **105** 243–266 (1989).

[25] S.F. Mingaleev and Y.S. Kivshar, "Self-trapping and stable localized modes in nonlinear photonic crystals," Phys. Rev. Lett. **86**, 5474-5477 (2001).

[26] E.A. Ostrovskaya and Yu.S. Kivshar, "Matter-wave gap solitons in atomic band-gap structures," Phys. Rev. Lett. **90**, 160407 (2003).

[27] A. Pankov, "Periodic nonlinear Schrödinger equation with application to photonic crystals," Milan J. Math. **73**, 259–287 (2005)

[28] D. Pelinovsky and G. Schneider, "Justification of the coupled-mode approximation for a nonlinear elliptic problem with a periodic potential," Appl. Anal. **86**, 1017–1036 (2007)

[29] D.E. Pelinovsky, A.A. Sukhorukov, and Y. Kivshar, "Bifurcations and stability of gap solitons in periodic structures," Phys. Rev. E **70**, 036618 (2004)

[30] L.P. Pitaevskii, "Vortex lines in an imperfect Bose gas," Zh. Eksp. Teor. Fiz. **40**, 646-651 (1961) [Sov. Phys. JETP **13**, 451-454 (1961)]

[31] M. Reed and B. Simon, Methods of Modern Mathematical Physics. IV. Analysis of Operators, Academic Press, New York, 1978

[32] G. Schneider and H. Uecker, "Nonlinear coupled mode dynamics in hyperbolic and parabolic periodically structured spatially extended systems," Asymp. Anal. **28**, 163–180 (2001)

[33] Z. Shi and J. Yang, "Solitary waves bifurcated from Bloch-band edges in two-dimensional periodic media," Phys. Rev. E **75**, 056602 (2007)

[34] J.E. Sipe, C.M. De Sterke, and B.J. Eggleton, "Rigorous derivation of coupled mode equations for short, high-intensity grating-coupled, co-propagating pulses," Journal of Modern Optics, **49** 1437–1452 (2002).

[35] I. Stakgold, Green's Functions and Boundary Value Problems, J. Wiley & Sons, New York, 1979.

[36] C.A. Stuart, "Bifurcations into spectral gaps," Bull. Belg. Math. Soc. - Simon Stevin Suppl., 1995, 59pp.

[37] A.A. Sukhorukov, and Y. Kivshar, "Nonlinear guided waves and spatial solitons in a periodic layered medium," J. Opt. Soc. Am. B **19**, 772–781 (2002).

[38] M. E. Taylor, Partial Differential Equations I, Springer, New York, (1996)

[39] C.H. Wilcox, "Theory of Bloch waves," J. Analyse Math. **33** 146–167 (1978).