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

March 3, 2009

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. (2008), online] 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 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

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

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

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

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

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 [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 $V \in H_{\text{loc}}^m(\mathbb{R}^2), m > 1$. 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. 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 physical
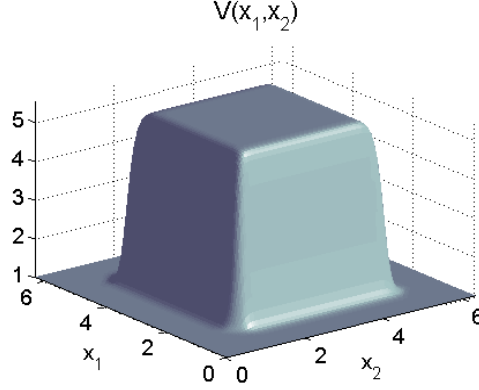
Figure 1: The periodic potential V in (1.3) over the Wigner-Seitz cell.

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

$$\hat{\phi}(k) := (\mathcal{F}\phi)(k) := \frac{1}{2\pi} \int_{\mathbb{R}^2} \phi(x)e^{-ik\cdot x}dx, \qquad \phi(x) = (\mathcal{F}^{-1}\hat{\phi})(k) := \frac{1}{2\pi} \int_{\mathbb{R}^2} \hat{\phi}(k)e^{ik\cdot x}dk, \qquad (1.4)$$

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

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

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 $s \geq 0$. Thus, henceforth we use $H^s(\mathbb{R}^2) := \mathcal{F}^{-1}L_s^2(\mathbb{R}^2)$ as definition for $0 \leq s \in \mathbb{R}$. This also gives a very simple proof of the Sobolev embedding theorem $\|\phi\|_{C^k} \leq C\|\phi\|_{H^s}$ for $k < s - 1$, see Lemma 4.2 below.

Main result. Let $s > 1$ 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. 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 $q \geq 0$. Then for $\omega = \omega_* + \varepsilon^2\Omega$ with ε^2 sufficiently small there exists a GS ϕ_{GS}

3

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

$$\|\phi_{GS} - \varepsilon\phi^{(0)}\|_{H^s(\mathbb{R}^2)} \le C\varepsilon^{4/5}, \tag{1.7}$$

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 $\|\phi_{GS} - \varepsilon\phi^{(0)}\|_{L^\infty(\mathbb{R}^2)} \le C\varepsilon^{1+\beta}$ 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 $\varepsilon^{4/5}$ 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). ⌋

Remark 1.2 Time-dependent CMEs have been justified in 1D for infinitesimal [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. ⌋

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

$$\begin{aligned}
&Lu_n(k; x) = \omega_n(k)u_n(k; x), \ x \in \mathbb{P}^2 := [0, 2\pi)^2, \\
&u_n(k; (2\pi, x_2)) = e^{i2\pi k_1}u_n(k; (0, x_2)), \ u_n(k; (x_1, 2\pi)) = e^{i2\pi k_2}u_n(k; (x_1, 0)).
\end{aligned} \tag{2.1}$$

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

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

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

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, via symmetries the $\omega_n(k)$ can be recovered from their values in the irreducible Brillouin zone B_0, see Fig. 2. From (2.2) we also note that for $x \in \mathbb{P}^2$ and $n \in \mathbb{Z}^2$ we have

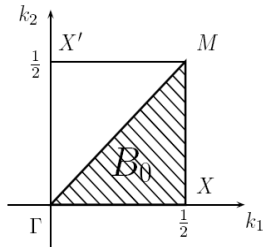$$u_n(k; (x_1 + 2n_1\pi, x_2 + 2n_2\pi)) = e^{2\pi i n \cdot k} u_n(k; x). \tag{2.4}$$

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

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 $k \in B_0$.

We choose the contrast η of the periodic structure (1.3) 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
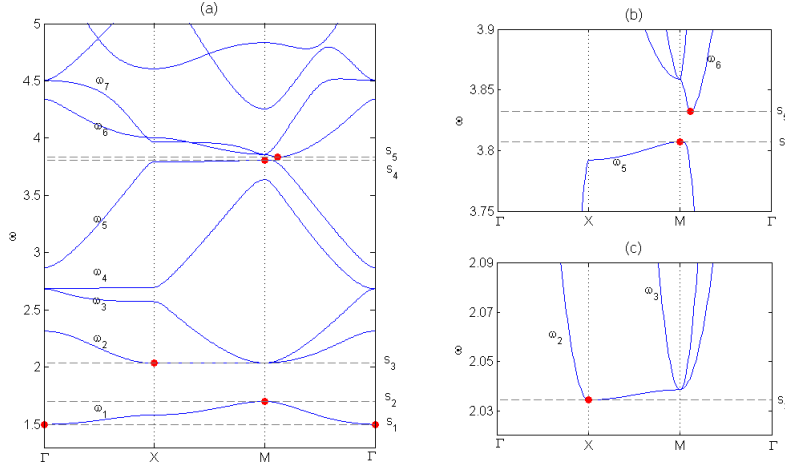
Figure 3: (a) Band structure of L with $\eta = 5.35$ along ∂B_0. Red dots label band extrema at gap edges $s_1, \ldots, s_5$. (b) Detail in the second finite gap. (c) Detail near the edge s_3 showing that ω_2 is not flat for k between X and M.
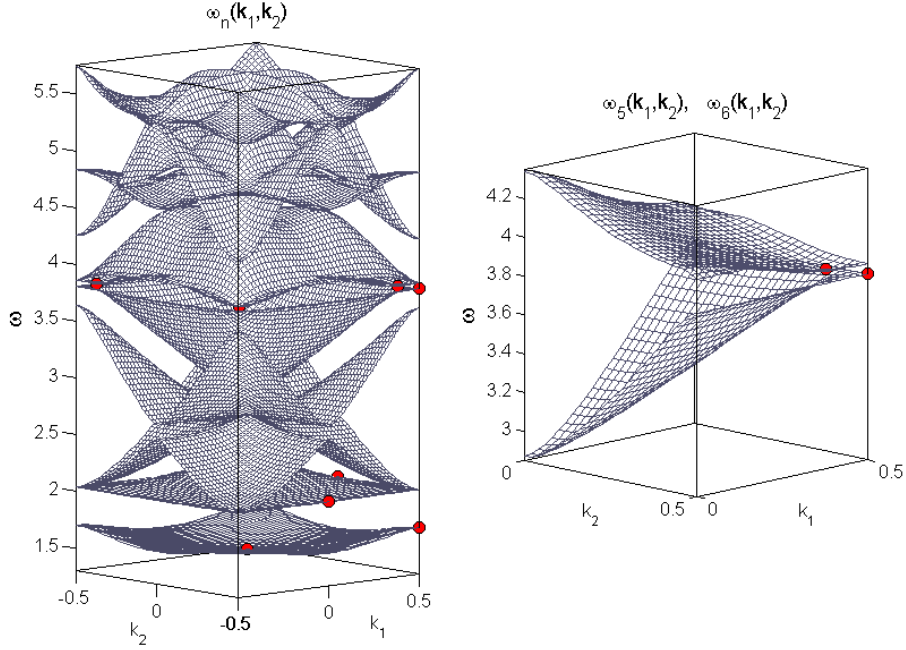

Figure 4: Band structure of L with $\eta = 5.35$. On the right a detail of ω_5 and ω_6 near the second finite gap.

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

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

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

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 $(k_1, -k_2) \doteq (k_1, k_2)$, $a_2 = 0$ does not hold in general 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

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

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 $\mathrm{e}^{-i2\pi k_1}$ and $\mathrm{e}^{-i2\pi k_2}$. Thus

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

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 $x_1 = x_2$.
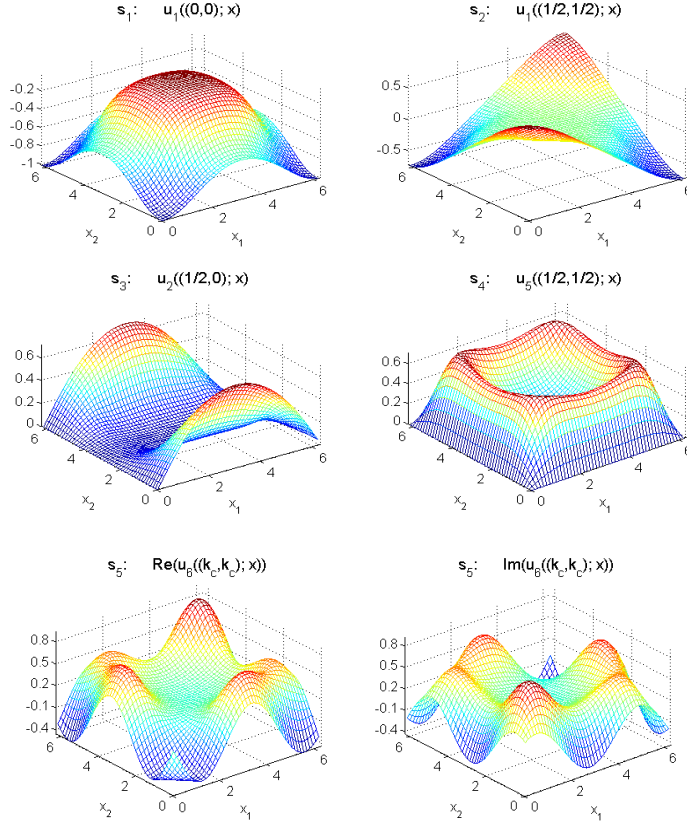
7

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

Fig. 5 shows that all the Bloch functions at $s_1, s_2, \ldots, s_4$ are either even or odd in each variable. This actually follows from (2.5) and the fact that these gap edges occur at $k \in S = \{\Gamma, X, X', M\}$. As each coordinate of any $k \in S$ 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

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

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.*

8

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 $\left\{u_{n_j}(k^{(j)}; x)\right\}_{j=1}^{N}$ 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 $\omega_{n_j}(k), j = 1, \ldots N$ at the corresponding points $k^{(j)} \in B$.

Assumption A.2 At least one of the values $\partial_{k_1}^2\omega_{n_j}(k^{(j)}), \partial_{k_2}^2\omega_{n_j}(k^{(j)})$ or $\partial_{k_1}\partial_{k_2}\omega_{n_j}(k^{(j)})$ is nonzero for each $j \in \{1, \ldots, N\}$.

Remark 3.1 Assumptions A.1 and A.2 are satisfied by the potential (1.3) with $\eta = 5.35$ at all the gap edges $s_1, \ldots, s_5$. ⌋

Remark 3.2 The nonzero second derivatives in A.2 ensure that the resulting CMEs are of second order. The assumption that the extrema are maxima or minima implies that $(\partial_{k_1}\partial_{k_2}\omega_{n_j}(k^{(j)}))^2 \leq \partial_{k_1}^2\omega_{n_j}(k^{(j)})\partial_{k_2}^2\omega_{n_j}(k^{(j)})$. 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. ⌋

Remark 3.3 The Bloch waves $u_{n_j}(k^{(j)}; \cdot), j=1, \ldots, N$ defined by the extrema are called "resonant". ⌋

Remark 3.4 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. $u_{n_1}, \ldots, u_{n_N}$ 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 are 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

$$\phi(x) = \varepsilon \sum_{j=1}^{N} A_j(y) u_{n_j}(k^{(j)}; x) + \varepsilon^2 \phi^{(1)} + \varepsilon^3 \phi^{(2)} + \mathcal{O}(\varepsilon^4) \tag{3.2}$$

$$\omega = \omega_* + \varepsilon^2 \Omega, \qquad y = \varepsilon x, \qquad 0 < \varepsilon \ll 1.$$

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 2π−antiperiodic in x_1 and even and 2π−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 $\phi^{(1)} = \partial_{y_1} A_1 v_1^{(x_1)} + \partial_{y_1} A_2 v_2^{(x_1)} + \partial_{y_2} A_1 v_1^{(x_2)} + \partial_{y_2} A_2 v_2^{(x_2)}$. (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

$$\begin{aligned}
[L - s_3]\phi^{(2)} = \ & \Omega(A_1 v_1 + A_2 v_2) + \Delta_{y_1,y_2} A_1 v_1 + \Delta_{y_1,y_2} A_2 v_2 \\
& +2 \left[\partial_{y_1}^2 A_1 \partial_{x_1} v_1^{(x_1)} + \partial_{y_1}^2 A_2 \partial_{x_1} v_2^{(x_1)} + \partial_{y_2}^2 A_1 \partial_{x_2} v_1^{(x_2)} + \partial_{y_2}^2 A_2 \partial_{x_2} v_2^{(x_2)} \right. \\
& \left. + \partial_{y_1} \partial_{y_2} A_1 \partial_{x_1} v_1^{(x_2)} + \partial_{y_1} \partial_{y_2} A_2 \partial_{x_1} v_2^{(x_2)} + \partial_{y_1} \partial_{y_2} A_1 \partial_{x_2} v_1^{(x_1)} + \partial_{y_1} \partial_{y_2} A_2 \partial_{x_2} v_2^{(x_1)} \right] \\
& -\sigma \left[\sum_{j=1}^2 |A_j|^2 A_j v_j^3 + 2|A_1|^2 A_2 v_1^2 v_2 + 2|A_2|^2 A_1 v_2^2 v_1 + A_1^2 \bar{A}_2 v_1^2 v_2 + A_2^2 \bar{A}_1 v_2^2 v_1 \right],
\end{aligned}$$

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 $\text{Ker}(L^* - s_3)$. 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:

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

$$
\alpha_1 = 1 + 2 \int_{-2\pi}^{2\pi} \int_{-2\pi}^{2\pi} v_1 \partial_{x_1} v_1^{(x_1)} dx, \qquad \alpha_2 = 1 + 2 \int_{-2\pi}^{2\pi} \int_{-2\pi}^{2\pi} v_1 \partial_{x_2} v_1^{(x_2)} dx,
$$

$$
\gamma_1 = \int_{-2\pi}^{2\pi} \int_{-2\pi}^{2\pi} v_1^4 dx \qquad \text{and} \qquad \gamma_2 = \int_{-2\pi}^{2\pi} \int_{-2\pi}^{2\pi} v_1^2 v_2^2 dx.
$$

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

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

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

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

$$
\tilde{\phi}((k_1 + 1, k_2); x) = e^{-ix_1} \tilde{\phi}(k; x), \quad \tilde{\phi}((k_1, k_2 + 1); x) = e^{-ix_2} \tilde{\phi}(k; x).
\tag{3.7}
$$

11

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

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

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

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

$$\left[\tilde{L} - \omega\right]\tilde{\phi} + \sigma\,\tilde{\phi} *_B \bar{\tilde{\phi}} *_B \tilde{\phi} = 0, \tag{3.9}$$

on $(k;x) \in \mathbb{T}^2 \times \mathbb{P}^2$, where we recall from (2.3) that $\tilde{L}(k;x) = (i\partial_{x_1}-k_1)^2 + (i\partial_{x_2}-k_2)^2 + V(x)$. Corresponding to (3.2), the asymptotic ansatz we use in the $\tilde{\phi}$ variables reads

$$\tilde{\phi}(k;x) = \frac{1}{\varepsilon}\tilde{\psi}^{(0)}(k;x) + \tilde{\psi}^{(1)}(k;x) + \varepsilon\tilde{\psi}^{(2)}(k;x) + \mathcal{O}(\varepsilon^2),$$

$$\tilde{\psi}^{(0)}(k;x) = \sum_{j=1}^{N} \chi_{D_j}(k)\hat{A}_j\left(\frac{k-k^{(j)}}{\varepsilon}\right)p_{n_j}(k^{(j)};x), \qquad \omega = \omega_* + \Omega\varepsilon^2, \qquad 0 < \varepsilon \ll 1, \tag{3.10}$$

where p_{n_j} is defined in (2.2), $\chi_{D_j}(k)$ is the characteristic function of the set D_j, and D_j is a disk with radius ε^r

$$\frac{1}{3} < r < 1 \tag{3.11}$$

centered at $k^{(j)}$, and mapped periodically on $\mathbb{T}^2$ if $\{k \in \mathbb{R}^2 : |k-k^{(j)}| < \varepsilon^r\}$ overlaps with the exterior of $\mathbb{T}^2$. We denote this as

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

where $\doteq$ means equal modulo 1 in each component, see Fig. 6 for an example. The reason for (3.11) will be explained in §4.2.
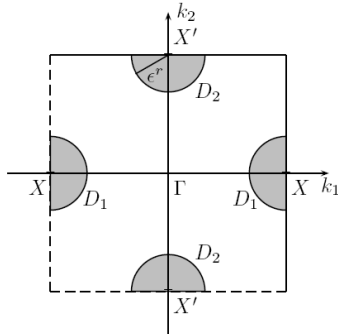

Figure 6: Decomposition of the k-space for the example (1.3) with $\eta = 5.35$ at $\omega_* = s_3$.

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

idea of the ansatz (3.10) is as follows. A direct calculation shows that

$$\left(\mathcal{T}(\varepsilon A_j(\varepsilon x)u_{n_j}(k^{(j)};x))\right)(k;x) = \frac{1}{\varepsilon}\tilde{A}_j\left(\frac{k-k^{(j)}}{\varepsilon};x\right)p_{n_j}(k^{(j)};x).$$

However, it turns out to be more convenient to approximate the function $\tilde{A}_j\left(\frac{k-k^{(j)}}{\varepsilon};x\right)$, which satisfies the periodicities

$$\tilde{A}_j\left(\frac{(k_1+1,k_2)-k^{(j)}}{\varepsilon};x\right) = \tilde{A}_j\left(\frac{(k_1,k_2)-k^{(j)}}{\varepsilon};x\right) = \tilde{A}_j\left(\frac{(k_1,k_2+1)-k^{(j)}}{\varepsilon};x\right),$$

by the function $\chi_{D_j}(k)\hat{A}_j\left(\frac{k-k^{(j)}}{\varepsilon}\right)$, where $\hat{A}$ is the Fourier transform of A. We have

$$\mathcal{T}^{-1}\left[\frac{1}{\varepsilon}\chi_{D_j}(k)\hat{A}_j\left(\frac{k-k^{(j)}}{\varepsilon}\right)\right](x) = \varepsilon A(\varepsilon x)e^{ik^{(j)}\cdot x} - \varepsilon h(x;\varepsilon), \tag{3.13}$$

where $h(x;\varepsilon) := \int_{|p|\geq\varepsilon^{r-1}}\hat{A}_j(p)e^{i\varepsilon p\cdot x}\,dp$. For $\hat{A}_j \in L_s^2(\mathbb{R}^2)$, $s > 1$ we have, e.g.,

$$|h(x)| = \left|\int_{|p|\geq\varepsilon^{r-1}}\hat{A}_j(p)e^{i\varepsilon p\cdot x}\,dp\right| \leq \|\hat{A}\|_{L_s^2}\left(\int_{|p|\geq\varepsilon^{r-1}}\frac{1}{1+|p|^{2s}}\,dp\right)^{1/2} \leq C\|\hat{A}\|_{L_s^2}\varepsilon^{(1-r)(s-1)},$$

and similar estimates for other relevant norms of $h(\cdot;\varepsilon)$. At this point we merely need $0 < r < 1$. In this sense $\varepsilon^{-1}\psi^{(0)}(x)$ is an approximation of $\varepsilon\phi^{(0)}(x)$ in (1.6) (see also (3.2)).

Letting $p^{(j)} := \frac{k-k^{(j)}}{\varepsilon}$ we have

$$\tilde{L}(k;x) = \tilde{L}(k^{(j)}+\varepsilon p^{(j)};x) = \tilde{L}(k^{(j)};x) - 2\varepsilon\left[(i\partial_{x_1}-k_1^{(j)})p_1^{(j)} + (i\partial_{x_2}-k_2^{(j)})p_2^{(j)}\right] + \varepsilon^2\left[p_1^{(j)^2} + p_2^{(j)^2}\right]. \tag{3.14}$$

Now substituting (3.10) in (3.9) and using (3.14), we again obtain a hierarchy of equations:

$$\mathcal{O}(\varepsilon^{-1}): \qquad \hat{A}_j(p^{(j)})\left[\tilde{L}(k^{(j)};x) - \omega_*\right]p_{n_j}(k^{(j)};x) = 0 \ \text{ for } k \in D_j, j \in \{1,\ldots,N\},$$

which holds by definition of $\omega_* = \omega_{n_j}(k^{(j)})$. There is no contribution to $\mathcal{O}(\varepsilon^{-1})$ for $k \in \mathbb{T}^2 \setminus \cup_j D_j$.

$$\mathcal{O}(1): \qquad \left[\tilde{L}(k^{(j)};x) - \omega_*\right]\tilde{\psi}^{(1)}(k;x) = 2\hat{A}_j(p^{(j)})\left[p_1^{(j)}(i\partial_{x_1}-k_1^{(j)}) + p_2^{(j)}(i\partial_{x_2}-k_2^{(j)})\right]p_{n_j}(k^{(j)};x)$$

for $k \in D_j, j \in \{1,\ldots,N\}$. 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 $k_m, m \in \{1,2\}$ and evaluating at $k = k^{(j)}$, we obtain

$$\left[\tilde{L}(k^{(j)};x) - \omega_*\right]\partial_{k_m}p_{n_j}(k^{(j)};x) = 2(i\partial_{x_m}-k_m^{(j)})p_{n_j}(k^{(j)};x). \tag{3.15}$$

Therefore, $\tilde{\psi}^{(1)}(k;x) = \sum_{m=1}^2 p_m^{(j)}\hat{A}_j(p^{(j)})\partial_{k_m}p_{n_j}(k^{(j)};x)$ for $k \in D_j$, $j \in \{1,\ldots,N\}$.

13

For $k \in \mathbb{T}^2 \setminus \cup_j D_j$ we get $\left[\tilde{L}(k^{(j)}; x) - \omega_*\right] \tilde{\psi}^{(1)}(k; x) = 0$, which has no nontrivial solution (on $k \in \mathbb{T}^2 \setminus \cup_j D_j$). Thus, $\tilde{\psi}^{(1)}(k; x) \equiv 0$ on $k \in \mathbb{T}^2 \setminus \cup_j D_j$ and we may write

$$\tilde{\psi}^{(1)}(k; x) = \sum_{j=1}^{N} \chi_{D_j}(k) \sum_{m=1}^{2} p_m^{(j)} \hat{A}_j(p^{(j)}) \partial_{k_m} p_{n_j}(k^{(j)}; x). \tag{3.16}$$

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

$$\begin{aligned}
\left[\tilde{L}(k^{(j)}; x) - \omega_*\right] \tilde{\psi}^{(2)}(k; x) =& \Omega \hat{A}_j(p^{(j)}) p_{n_j}(k^{(j)}; x) + 2 \left[p_1^{(j)}(\mathrm{i}\partial_{x_1} - k_1^{(j)}) + p_2^{(j)}(\mathrm{i}\partial_{x_2} - k_2^{(j)})\right] \tilde{\psi}^{(1)}(k; x) \\
& - \left(p_1^{(j)^2} + p_2^{(j)^2}\right) \hat{A}_j(p^{(j)}) p_{n_j}(k^{(j)}; x) \\
& - \sigma \chi_{D_j}(k) \frac{1}{\varepsilon^4} (\tilde{\psi}^{(0)} *_B \tilde{\psi}^{(0)} *_B \tilde{\bar{\psi}}^{(0)})(k; x)
\end{aligned}$$

for $k \in D_j, j \in \{1, \ldots, N\}$. Substituting $\tilde{\psi}^{(1)}$ from (3.16), we get

$$\begin{aligned}
\left[\tilde{L}(k^{(j)}; x) - \omega_*\right] \tilde{\psi}^{(2)}(k; x) =& \Omega \hat{A}_j(p) p_{n_j}(k^{(j)}; x) \\
& - \left[p_{n_j}(k^{(j)}; x) - 2(\mathrm{i}\partial_{x_1} - k_1^{(j)}) \partial_{k_1} p_{n_j}(k^{(j)}; x)\right] p_1^2 \hat{A}_j(p) \\
& - \left[p_{n_j}(k^{(j)}; x) - 2(\mathrm{i}\partial_{x_2} - k_2^{(j)}) \partial_{k_2} p_{n_j}(k^{(j)}; x)\right] p_2^2 \hat{A}_j(p) \\
& + 2 \left[(\mathrm{i}\partial_{x_1} - k_1^{(j)}) \partial_{k_2} p_{n_j}(k^{(j)}; x) + (\mathrm{i}\partial_{x_2} - k_2^{(j)}) \partial_{k_1} p_{n_j}(k^{(j)}; x)\right] p_1 p_2 \hat{A}_j(p) \\
& - \sigma \chi_{D_j}(k) \frac{1}{\varepsilon^4} (\tilde{\psi}^{(0)} *_B \tilde{\psi}^{(0)} *_B \tilde{\bar{\psi}}^{(0)})(k; x),
\end{aligned} \tag{3.17}$$

where we dropped the superscript (j) of p since the meaning of p as argument of $\hat{A}_j$ is clear. We rewrite the nonlinear term as

$$\begin{aligned}
G := \frac{\sigma}{\varepsilon^4} \chi_{D_j}(k) (\tilde{\psi}^{(0)} *_B \tilde{\psi}^{(0)} *_B \tilde{\bar{\psi}}^{(0)}) =& \frac{\sigma}{\varepsilon^4} \chi_{D_j}(k) \left[\sum_{l=1}^{N} \xi_l *_B \xi_l *_B \xi_l^c + 2 \sum_{\substack{l,m=1 \\ l \neq m}}^{N} \xi_l *_B \xi_m *_B \xi_l^c \right. \\
& \left. + \sum_{\substack{l,m=1 \\ l \neq m}}^{N} \xi_l *_B \xi_l *_B \xi_m^c + \sum_{\substack{l,m=1 \\ l \neq m, l \neq n, m \neq n}}^{N} \xi_l *_B \xi_m *_B \xi_n^c \right],
\end{aligned} \tag{3.18}$$

where $\xi_l = \xi_l(k; x) := \chi_{D_l}(k) \hat{A}_l \left(\frac{k - k^{(l)}}{\varepsilon}\right) p_{n_l}(k^{(l)}; x)$ and $\xi_l^c = \xi_l^c(k; x) := \chi_{-D_l}(k) \hat{\bar{A}}_l \left(\frac{k + k^{(l)}}{\varepsilon}\right) \overline{p_{n_l}}(k^{(l)}; x)$. The last sum or the three last sums in (3.18) are absent if $N = 2$ or $N = 1$ respectively.

A transformation in the integral leads to

$$
\xi_l *_B \xi_m *_B \xi_n^c = \varepsilon^4 p_l(k^{(l)}; x) p_m(k^{(m)}; x) \overline{p_n}(k^{(n)}; x) \int_{D_{2\varepsilon^{r-1}}} \int_{D_{\varepsilon^{r-1}}} \chi_{D_{\varepsilon^{r-1}}} \left(\frac{k - (k^{(l)} + k^{(m)} - k^{(n)})}{\varepsilon} - \eta \right)
$$

$$
\hat{A}_l \left(\frac{k - (k^{(l)} + k^{(m)} - k^{(n)})}{\varepsilon} - \eta \right) \chi_{D_{\varepsilon^{r-1}}} (\eta - \tau) \hat{A}_m(\eta - \tau) \chi_{D_{\varepsilon^{r-1}}} (\tau) \hat{\bar{A}}_n(\tau) d\tau d\eta
$$

(3.19)

for any $l, m, n \in \{1, \ldots, N\}$, where $D_{\varepsilon^{r-1}} = \{p \in \mathbb{R}^2 : |p| < \varepsilon^{r-1}\}$. This formula also shows that $G = \mathcal{O}(1)$ as required for a consistent asymptotic expansion. For $\varepsilon \to 0$ the integrals in (3.19) approximate the Fourier transform of $A_l A_m \bar{A}_n$, namely

$$
(\xi_l *_B \xi_m *_B \xi_n^c)(k) = \varepsilon^4 p_l(k^{(l)}; x) p_m(k^{(m)}; x) \overline{p_n}(k^{(n)}; x) \widehat{A_l A_m \bar{A}_n} \left(\frac{k - (k^{(l)} + k^{(m)} - k^{(n)})}{\varepsilon} \right) + \tilde{h}(k, x; \varepsilon)
$$

where, similarly to (3.13), for $\hat{A}_l, \hat{A}_m, \hat{A}_n \in L_s^2(\mathbb{R}^2)$ we have $\|\tilde{h}(\cdot, \cdot; \varepsilon)\|_{L^2(\mathbb{T}^2, H^s(\mathbb{P}^2, \mathbb{C}))} = \mathcal{O}(\varepsilon^{1-r})$.

It is straightforward to check that the function $\chi_{D_j}(k)$ in G annihilates all terms in the sum except such terms $\xi_l *_B \xi_m *_B \xi_n^c$ for which $k^{(l)} + k^{(m)} - k^{(n)} \doteq k^{(j)}$, where we recall that $\doteq$ means equal modulo 1 in each component. Obvious examples (for $N \geq 2$) of not annihilated terms are $\xi_j *_B \xi_j *_B \xi_j^c$ and $\xi_l *_B \xi_j *_B \xi_l^c \quad \forall j, l \in \{1, \ldots, N\}$ appearing in the first two sums in G.

We return now to equation (3.17) for $\tilde{\psi}^{(2)}$. Its solvability condition is $L^2(\mathbb{P}^2)$-orthogonality to $\text{Ker}(\tilde{L}(k^{(j)}; x) - \omega_*) = \text{span}\{p_{n_j}(k^{(j)}; x), j = 1, \ldots, N\}$. This yields the CMEs in Fourier variables

$$
\Omega \hat{A}_j - \left(\frac{1}{2} \partial_{k_1}^2 \omega_{n_j}(k^{(j)}) p_1^2 + \frac{1}{2} \partial_{k_2}^2 \omega_{n_j}(k^{(j)}) p_2^2 + \partial_{k_1} \partial_{k_2} \omega_{n_j}(k^{(j)}) p_1 p_2 \right) \hat{A}_j - \hat{\mathcal{N}}_j = 0,
$$

(3.20)

$j \in \{1, \ldots, N\}$, where $\hat{\mathcal{N}}_j = \langle G(k; \cdot), p_{n_j}(k^{(j)}; \cdot) \rangle_{L^2(\mathbb{P}^2)}$. Inverse Fourier transform yields

$$
\Omega A_j + \left(\frac{1}{2} \partial_{k_1}^2 \omega_{n_j}(k^{(j)}) \partial_{y_1}^2 + \frac{1}{2} \partial_{k_2}^2 \omega_{n_j}(k^{(j)}) \partial_{y_2}^2 + \partial_{k_1} \partial_{k_2} \omega_{n_j}(k^{(j)}) \partial_{y_1} \partial_{y_2} \right) A_j - \mathcal{N}_j = 0,
$$

(3.21)

where the structure and coefficients in $\mathcal{N}_j$ for our example (1.3) will be explained in §3.2.2. The appearance of second derivatives of the bands ω_{n_j} in (3.20) is due to the following

Lemma 3.5 *For any $l, m \in \{1, 2\}$*

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

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 the extremal point $k^{(j)}$ coincides with one of the vertices of the first irreducible Brillouin zone or of its reflection and $\omega_{n_j}(k^{(j)})$ has geometric multiplicity one.

15

Lemma 3.6 *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 $k^{(j)} \in S = \{\Gamma, X, X', M\}$ 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

$$\langle(\mathrm{i}\partial_{x_m} - k_m^{(j)})\partial_{k_l}p_{n_j}(k^{(j)}; \cdot), p_{n_j}(k^{(j)}; \cdot)\rangle_{L^2(\mathbb{P}^2)} = \langle\partial_{k_l}p_{n_j}(k^{(j)}; \cdot), (\mathrm{i}\partial_{x_m} - k_m^{(j)})p_{n_j}(k^{(j)}; \cdot)\rangle_{L^2(\mathbb{P}^2)}. \quad (3.22)$$

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

$$[L - \omega_*]u = 2\partial_{x_l}u_{n_j}(k^{(j)}; x), \qquad u(2\pi, x_2) = e^{\mathrm{i}2\pi k_1^{(j)}}u(0, x_2), \quad u(x_1, 2\pi) = e^{\mathrm{i}2\pi k_2^{(j)}}u(x_1, 0), \quad (3.23)$$

analogously to (3.3). The inner product in (3.22) 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 $k^{(j)} \in S$, u_{n_j} is even or odd in x_l (Lemma 2.2). From (3.23) 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 $[-\pi, \pi]^2$. $\qquad\square$

3.2.2 CMEs for the Example (1.3)

We now calculate the explicit form of the CMEs (3.21) 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 $k^{(j)} \notin S$ 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 3.5, the functions $\partial_{k_l}p_{n_j}(k^{(j)}; x)$ have to be computed. They are solutions of the singular system (3.15) 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 $k^{(1)} = \Gamma$. Thus

$$\left[\Omega + \alpha(\partial_{y_1}^2 + \partial_{y_2}^2)\right]A - \sigma\gamma|A|^2A = 0, \quad (3.24)$$

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 $k^{(1)} = M$. Thus we again have (3.24), now with $\alpha = \frac{1}{2}\partial_{k_1}^2\omega_1(M) = \frac{1}{2}\partial_{k_2}^2\omega_1(M)$ and $\gamma = \langle p_1(M; \cdot)^2, p_1(M; \cdot)^2\rangle_{L^2(\mathbb{P}^2)} = \|p_1(M; \cdot)\|_{L^4(\mathbb{P}^2)}^4$. 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 $k^{(2)} = X'$. Because $2k^{(1)} - k^{(2)} = (1, -1/2) \doteq k^{(2)}$, the term $\xi_1 *_B \xi_1 *_B \xi_2^c$ appears in the equation for $k \in D_2$. Similarly, as $2k^{(2)} - k^{(1)} \doteq$

$k^{(1)}$, the term $\xi_2 *_B \xi_2 *_B \xi_1^c$ appears in the equation for $k \in D_1$. The resulting CMEs are

$$
\begin{aligned}
\left[\Omega + \alpha_1 \partial_{y_1}^2 + \alpha_2 \partial_{y_2}^2\right] A_1 - \sigma \left[\gamma_1 |A_1|^2 A_1 + 2\gamma_2 |A_2|^2 A_1 + \gamma_3 A_2^2 \bar{A}_1\right] &= 0, \\
\left[\Omega + \alpha_2 \partial_{y_1}^2 + \alpha_1 \partial_{y_2}^2\right] A_2 - \sigma \left[\gamma_1 |A_2|^2 A_2 + 2\gamma_2 |A_1|^2 A_2 + \gamma_3 A_1^2 \bar{A}_2\right] &= 0,
\end{aligned}
\tag{3.25}
$$

where

$$
\begin{aligned}
\alpha_1 &= \tfrac{1}{2}\partial_{k_1}^2 \omega_2(X) = \tfrac{1}{2}\partial_{k_2}^2 \omega_2(X'), \quad \alpha_2 = \tfrac{1}{2}\partial_{k_2}^2 \omega_2(X) = \tfrac{1}{2}\partial_{k_1}^2 \omega_2(X'), \\
\gamma_1 &= \langle p_2(X;\cdot)^2, p_2(X;\cdot)^2 \rangle_{L^2(\mathbb{P}^2)} = \langle p_2(X';\cdot)^2, p_2(X';\cdot)^2 \rangle_{L^2(\mathbb{P}^2)} = \|p_2(X;\cdot)\|_{L^4(\mathbb{P}^2)}^4 = \|p_2(X';\cdot)\|_{L^4(\mathbb{P}^2)}^4, \\
\gamma_2 &= \langle |p_2(X';\cdot)|^2, |p_2(X;\cdot)|^2 \rangle_{L^2(\mathbb{P}^2)}, \quad \text{and} \\
\gamma_3 &= \langle p_2(X';\cdot)^2, p_2(X;\cdot)^2 \rangle_{L^2(\mathbb{P}^2)} = \langle p_2(X;\cdot)^2, p_2(X';\cdot)^2 \rangle_{L^2(\mathbb{P}^2)}.
\end{aligned}
$$

The identities in α_1, α_2 and γ_1 hold due to (2.6). The equality in γ_3 yields $\gamma_3 \in \mathbb{R}$. It follows from the symmetries (2.7), (2.6) and from (3.7). Indeed, using these properties, we arrive at

$$
\gamma_3 = \int_{\mathbb{P}^2} p_2(X;(x_2,x_1))^2 p_2(X;(x_1,x_2))^2 e^{2ix_1} dx_1 dx_2,
$$

$$
\bar{\gamma}_3 = \int_{\mathbb{P}^2} p_2(X;(x_1,x_2))^2 p_2(X;(x_2,x_1))^2 e^{2ix_2} dx_1 dx_2.
$$

The equality $\bar{\gamma}_3 = \gamma_3$ follows by interchanging integration in one of the integrals. Numerically,

$$
\alpha_1 \approx 2.599391, \alpha_2 \approx 0.040561, \gamma_1 \approx 0.090082, \gamma_2 \approx 0.003032, \text{ and } \gamma_3 \approx 0.000154.
$$

CMEs near $\omega_* = s_4$: Here $N = 1, n_1 = 5$ and $k^{(1)} = M$. The governing equation is (3.24) with $\alpha = \tfrac{1}{2}\partial_{k_1}^2 \omega_5(M) = \tfrac{1}{2}\partial_{k_2}^2 \omega_5(M) \approx -0.300655$ 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 $k^{(j)} \notin S$ here.

We start with the last two sums of G (see (3.18)). 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 $k^{(j)}, j \in \{1, \dots, 4\}$ for any choice of $l, m \in \{1, \dots, 4\}, l \neq m$. 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 $k^{(j)} \notin S$ is that Lemma 3.6 does not apply and mixed derivatives of A_j

may appear. The system of CMEs thus becomes

$$
\begin{aligned}
0 = {}& \left[\Omega + \alpha_1(\partial_{y_1}^2 + \partial_{y_2}^2) + \alpha_2 \partial_{y_1}\partial_{y_2}\right] A_1 \\
& - \sigma \left[\gamma_1 |A_1|^2 A_1 + 2\big(\gamma_2 \left(|A_2|^2 + |A_4|^2\right) A_1 + \tilde{\gamma}_1 |A_3|^2 A_1 + \tilde{\gamma}_2 A_2 A_4 \bar{A}_3\big)\right], \\
0 = {}& \left[\Omega + \alpha_1(\partial_{y_1}^2 + \partial_{y_2}^2) - \alpha_2 \partial_{y_1}\partial_{y_2}\right] A_2 \\
& - \sigma \left[\gamma_1 |A_2|^2 A_2 + 2\big(\gamma_2 \left(|A_1|^2 + |A_3|^2\right) A_2 + \tilde{\gamma}_1 |A_4|^2 A_2 + \bar{\tilde{\gamma}}_2 A_1 A_3 \bar{A}_4\big)\right], \\
0 = {}& \left[\Omega + \alpha_1(\partial_{y_1}^2 + \partial_{y_2}^2) + \alpha_2 \partial_{y_1}\partial_{y_2}\right] A_3 \\
& - \sigma \left[\gamma_1 |A_3|^2 A_3 + 2\big(\gamma_2 \left(|A_2|^2 + |A_4|^2\right) A_3 + \tilde{\gamma}_1 |A_1|^2 A_3 + \tilde{\gamma}_2 A_2 A_4 \bar{A}_1\big)\right], \\
0 = {}& \left[\Omega + \alpha_1(\partial_{y_1}^2 + \partial_{y_2}^2) - \alpha_2 \partial_{y_1}\partial_{y_2}\right] A_4 \\
& - \sigma \left[\gamma_1 |A_4|^2 A_4 + 2\big(\gamma_2 \left(|A_1|^2 + |A_3|^2\right) A_4 + \tilde{\gamma}_1 |A_2|^2 A_4 + \bar{\tilde{\gamma}}_2 A_1 A_3 \bar{A}_2\big)\right],
\end{aligned}
\tag{3.26}
$$

where

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

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

$$
\begin{aligned}
\left[\Omega + \alpha_1(\partial_{y_1}^2 + \partial_{y_2}^2) + \alpha_2 \partial_{y_1}\partial_{y_2}\right] A_1 - \sigma \left[\gamma_1 \left(|A_1|^2 + 2|A_3|^2\right) A_1 + 2\gamma_2 \left((|A_2|^2 + |A_4|^2)A_1 + A_2 A_4 \bar{A}_3\right)\right] = 0, \\
\left[\Omega + \alpha_1(\partial_{y_1}^2 + \partial_{y_2}^2) - \alpha_2 \partial_{y_1}\partial_{y_2}\right] A_2 - \sigma \left[\gamma_1 \left(|A_2|^2 + 2|A_4|^2\right) A_2 + 2\gamma_2 \left((|A_1|^2 + |A_3|^2)A_2 + A_1 A_3 \bar{A}_4\right)\right] = 0, \\
\left[\Omega + \alpha_1(\partial_{y_1}^2 + \partial_{y_2}^2) + \alpha_2 \partial_{y_1}\partial_{y_2}\right] A_3 - \sigma \left[\gamma_1 \left(|A_3|^2 + 2|A_1|^2\right) A_3 + 2\gamma_2 \left((|A_2|^2 + |A_4|^2)A_3 + A_2 A_4 \bar{A}_1\right)\right] = 0, \\
\left[\Omega + \alpha_1(\partial_{y_1}^2 + \partial_{y_2}^2) - \alpha_2 \partial_{y_1}\partial_{y_2}\right] A_4 - \sigma \left[\gamma_1 \left(|A_4|^2 + 2|A_2|^2\right) A_4 + 2\gamma_2 \left((|A_1|^2 + |A_3|^2)A_4 + A_1 A_3 \bar{A}_2\right)\right] = 0.
\end{aligned}
\tag{3.27}
$$

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

18

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 $\mathcal{F}^{-1}L_s^1(\mathbb{R}^2)$ 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, $L_s^1(\mathbb{R}^2)$ in Fourier space gives a direct pointwise estimate in physical space via $\|\phi\|_{C^s} \leq C\|\hat{\phi}\|_{L_s^1}$. On the other hand, working in Hilbert spaces H^s is conceptually simpler since it allows to go 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

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

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_s^2, $s \geq 0$, where

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

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

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

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

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 $\phi, \psi \in C^0$, $\|\phi\|_{C^0} \leq C\|\phi\|_{H^s}$, and $\|\phi\psi\|_{H^s} \leq C\|\phi\|_{H^s}\|\psi\|_{H^s}$.*

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

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

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

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

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

$$\|\phi\psi\|_{H^s} \leq C\|\hat{\phi} * \hat{\psi}\|_{L^2_s} \leq C\left(\||w\hat{\phi}| * |\hat{\psi}|\|_{L^2} + \||\hat{\phi}| * |w\hat{\psi}|\|_{L^2}\right)$$

$$\leq C\left(\|w\hat{\phi}\|_{L^2}\|\hat{\psi}\|_{L^1} + \|w\hat{\phi}\|_{L^2}\|\hat{\psi}\|_{L^1}\right) \leq 2C\|\hat{\phi}\|_{L^2_s}\|\hat{\psi}\|_{L^2_s} \leq C\|\phi\|_{H^s}\|\psi\|_{H^s}. \tag{4.5}$$

$\square$

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 $\|\phi\|_{C^k} \leq C\|\phi\|_{H^s}$, $k \in \mathbb{N}$, $k < s - d/2$. 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 $\vec{\tilde{\phi}\tilde{\psi}} \in \mathcal{X}^s$, $\phi \in C_b^k(\mathbb{R}^2)$ for $k < s - 1$, and $\phi(x) \to 0$ as $|x| \to \infty$.*

To justify Taylor expansions of the spectral bands and the Bloch function we use the following lemma, which is proved in [39].

Lemma 4.4 *Let $V(x) \in L^2(\mathbb{P}^2)$, then for each $n \in \mathbb{N}$ the band $\omega_n(k)$ and the Bloch function $u_n(k; x)$ are analytic in k on $k \in \mathbb{T}^2 \setminus Z_n$, where Z_n is the set, where the algebraic multiplicity of $\omega_n(k)$ as an eigenvalue of (2.1) is higher than one.*

4.2 Justification Step I: The extended Coupled Mode Equations

To justify the general stationary CMEs (3.20) as an asymptotic model for stationary gap solitons near an edge $\omega = \omega_*$, we again consider (3.9), 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 $\tilde{g}_n(k) = \langle(\tilde{\phi} *_B \bar{\tilde{\phi}} *_B \tilde{\phi})(k;\cdot), p_n(k;\cdot)\rangle_{L^2(\mathbb{P}^2)}$.

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

$$\vec{\tilde{\phi}}(k) = \varepsilon^{-1}\sum_{j=1}^{N}\hat{B}_j\left(\frac{k - k^{(j)}}{\varepsilon}\right)e_{n_j} + \vec{\tilde{\psi}}(k), \tag{4.8}$$

where e_{n_j} is the unit vector in the n_j direction in $\mathbb{R}^\mathbb{N}$ and $\operatorname{supp}\hat{B}_j \subset D_{\varepsilon^{r-1}} = \{p \in \mathbb{R}^2 : |p| \leq \varepsilon^{r-1}\}$. The part $\widetilde{\psi_{n_j}}(k)$ is set to zero on the critical set $k \in K_c := \cup\{D_l : l \in \{1,\ldots,N\}, n_l = n_j\}$, where $D_j = \{k \in \mathbb{R}^2 : |k - k^{(j)}| < \varepsilon^r\}$ mapped periodically on $\mathbb{T}^2$, $1/3 < r < 1/2$, cf.(3.11) and Fig. 6. We obtain

$$\frac{1}{\varepsilon}(\omega_{n_j}(k) - \omega_* - \varepsilon^2\Omega)\hat{B}_j\left(\frac{k - k^{(j)}}{\varepsilon}\right) = -\sigma\chi_{D_j}(k)\tilde{g}_{n_j}(k), \quad j = 1,\ldots,N, \tag{4.9}$$

$$(\omega_n(k) - \omega_* - \varepsilon^2\Omega)\widetilde{\psi}_n(k) = -\sigma(1 - \sum_{j=1}^{N}\chi_{D_j}(k)\delta_{n,n_j})\tilde{g}_n(k), \quad n \in \mathbb{N}, \tag{4.10}$$

and the goal is to solve (4.10) 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 (4.9). It turns out that the right norm for $\hat{B}_j$ is $\|\hat{B}_j\|_{L_s^2(D_{\varepsilon^{r-1}})}$, where we recall

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

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

Lemma 4.5 *For $\hat{\mathbf{B}} \in L_s^2(D_{\varepsilon^{r-1}})$ and $\vec{\tilde{\psi}} \in \mathcal{X}^s$, $s > 1$, we have*

$$\begin{aligned}
\|\vec{\tilde{g}}\|_{\mathcal{X}^s} \leq C\bigg(&\varepsilon^2\big(\sum_{j=1}^{N}\|\hat{B}_j\|_{L_s^2(D_{\varepsilon^{r-1}})}\big)^3 + \varepsilon^2\big(\sum_{j=1}^{N}\|\hat{B}_j\|_{L_s^2(D_{\varepsilon^{r-1}})}\big)^2\|\vec{\tilde{\psi}}\|_{\mathcal{X}^s} \\
&+ \varepsilon\big(\sum_{j=1}^{N}\|\hat{B}_j\|_{L_s^2(D_{\varepsilon^{r-1}})}\big)\|\vec{\tilde{\psi}}\|_{\mathcal{X}^s}^2 + \|\vec{\tilde{\psi}}\|_{\mathcal{X}^s}^3\bigg).
\end{aligned} \tag{4.11}$$

Proof. Heuristically, (4.11) can best be seen in physical space. With

$$\phi(x) = \mathcal{T}^{-1}\mathcal{D}^{-1}\vec{\tilde{\phi}}(k) = \varepsilon\sum_{j=1}^{N}B_j(\varepsilon x)u_{n_j}(k^{(j)}; x) + \psi = \varepsilon\phi^{(0)} + \psi$$

21

and $g = |\phi|^2\phi = |\varepsilon\phi^{(0)} + \psi|^2(\varepsilon\phi^{(0)} + \psi)$ we have, e.g., for the L^2–norm

$$\|g\|_{L^2} \leq C \left(\varepsilon^2 \big(\sum_{j=1}^N \|B_j u_{n_j}\|_{L^\infty}\big)^2 \|\varepsilon\phi^0\|_{L^2} + \varepsilon^2 \big(\sum_{j=1}^N \|B_j u_{n_j}\|_{L^\infty}\big)^2 \|\psi\|_{L^2} \right.$$
$$\left. + \varepsilon \big(\sum_{j=1}^N \|B_j u_{n_j}\|_{L^\infty}\big) \|\psi\|_{L^\infty} \|\psi\|_{L^2} + \|\psi\|_{L^\infty}^2 \|\psi\|_{L^2} \right).$$

For $s > 1$ we have $\|B_j(\varepsilon\cdot)u_{n_j}\|_{L^\infty} \leq C\|B_j(\cdot)\|_{L^\infty} \leq C\|B_j(\cdot)\|_{H^s} \leq C\|\hat{B}_j\|_{L^2_s(D_{\varepsilon^r-1})}$ and $\|\psi\|_{L^\infty} \leq C\|\psi\|_{H^s} \leq \|\vec{\tilde{\psi}}\|_{\mathcal{X}^s}$ using Lemma 4.1. Next, we invoke smoothness of u_{n_j}, namely for $V \in H^m_{\mathrm{loc}}$ we have $u_{n_j} \in H^{m+2}_{\mathrm{loc}}$, i.e., by Sobolev embedding $u_{n_j} \in C^r_b, r < m + 1$. Using $\partial_x B_j(\varepsilon x) = \varepsilon\partial_y B_j(y), y = \varepsilon x$, and $\|\partial_x^\alpha u_{n_j}\|_\infty \leq C$ for $|\alpha| \leq s$ under the assumption $V \in H^{s-1+\delta}_{\mathrm{loc}}(\mathbb{R}^2), \delta > 0$, we obtain (4.11) for $s \in \mathbb{N}$. Here ∂_x^α is an $|\alpha|$-th order x−derivative.

For the general case $1 < s \in \mathbb{R}$ we proceed as follows. First, by interpolation, see, e.g., [38, §4.2], $\|\phi u\|_{H^s} \leq C\|u\|_{C^{\lceil s\rceil}}\|\phi\|_{H^s}$. Thus, for instance, as in (4.5),

$$\|\varepsilon B_j(\varepsilon x)u_{n_j}\psi\|_{H^s} \leq C\|\varepsilon B_j(\varepsilon x)\psi\|_{H^s} \leq C\left\|\frac{1}{\varepsilon}\hat{B}\left(\frac{\cdot - k^{(j)}}{\varepsilon}\right) * \hat{\psi}\right\|_{L^2_s}$$
$$\leq C\left(\left\|\frac{1}{\varepsilon}\hat{B}\left(\frac{\cdot - k^{(j)}}{\varepsilon}\right)\right\|_{L^1}\|\hat{\psi}\|_{L^2_s} + \left\|\frac{1}{\varepsilon}\hat{B}\left(\frac{\cdot - k^{(j)}}{\varepsilon}\right)\right\|_{L^2_s}\|\hat{\psi}\|_{L^1}\right)$$
$$\leq C\varepsilon\left(\|\hat{B}\|_{L^1}\|\hat{\psi}\|_{L^2_s} + \|\hat{B}\|_{L^2_s}\|\hat{\psi}\|_{L^1}\right) \leq C\varepsilon\|\hat{B}\|_{L^2_s}\|\hat{\psi}\|_{L^2_s} \leq C\varepsilon\|\hat{B}\|_{L^2_s(D_{\varepsilon^r-1})}\|\vec{\tilde{\psi}}\|_{\mathcal{X}^s}.$$

We used $\|u_{n_j}\|_{C^{\lceil s\rceil}} \leq C\|V\|_{H^{\lceil s\rceil-1+\delta}_{\mathrm{loc}}}$, which gives the condition on the smoothness of V, $\left\|\frac{1}{\varepsilon}\hat{B}\left(\frac{\cdot - k^{(j)}}{\varepsilon}\right)\right\|_{L^1} = \varepsilon\|\hat{B}\|_{L^1}$ and, with $w(k) = (1 + |k|)^s$,

$$\int\left|\hat{B}\left(\frac{k - k^{(j)}}{\varepsilon}\right)\right|^2 w^2(k)\,\mathrm{d}k \leq C\int\left|\hat{B}\left(\frac{k - k^{(j)}}{\varepsilon}\right)\right|^2 w^2\left(\frac{k - k^{(j)}}{\varepsilon}\right)\,\mathrm{d}k = C\varepsilon^2\|\hat{B}\|_{L^2_s},$$

due to $w^2\left(\frac{k-k^{(j)}}{\varepsilon}\right) > w^2(k)$ for $\varepsilon < 1$ and large enough $|k|$. $\qquad\square$

Next, for ε sufficiently small we have

$$\min_{\substack{k\in\mathrm{supp}(\vec{\tilde{\psi}})\\ n\in\mathbb{N}}}|\omega_n(k) - \omega_*| \geq C\varepsilon^{2r} \tag{4.12}$$

due to $\partial_{k_1}\omega_{n_j}(k^{(j)}) = \partial_{k_2}\omega_{n_j}(k^{(j)}) = 0$. Thus we may invert $\omega_n(k) - \omega_*$ on supp $\vec{\tilde{\psi}}$ and from (4.11) and (4.1) we find that for $\|\hat{B}_j\|_{L^2_s(D_{\varepsilon^r-1})} \leq C, \ j = 1, \ldots, N$ we may solve (4.10) for $\vec{\tilde{\psi}}$ with

$$\|\vec{\tilde{\psi}}\|_{\mathcal{X}^s} \leq C\varepsilon^{2-2r}\sum_{j=1}^N\|\hat{B}_j\|_{L^2_s(D_{\varepsilon^r-1})}. \tag{4.13}$$

22

We now turn to (4.9). The first sum in the ansatz (4.8) corresponds to $\tilde{\psi}_0$ in the ansatz (3.10) used in the formal derivation of the CME. Therefore, plugging (4.8) into (4.7), truncating over $k \in D_j$ and mapping $k \in D_j$ to $p \in D_{\varepsilon^{r-1}}$ via $p = \varepsilon^{-1}(k - k^{(j)})$ yields the so called extended CMEs (eCMEs) in the form

$$\Omega \hat{B}_j - \left(\frac{1}{2} \partial_{k_1}^2 \omega_{n_j}(k^{(j)}) p_1^2 + \frac{1}{2} \partial_{k_2}^2 \omega_{n_j}(k^{(j)}) p_2^2 + \partial_{k_1} \partial_{k_2} \omega_{n_j}(k^{(j)}) p_1 p_2 \right) \hat{B}_j - \hat{Q}_j = \varepsilon^{\tilde{r}} \hat{R}_j(p), \qquad (4.14)$$

$j = 1, \ldots, N$. Here $\hat{Q}_j$ denote the nonlinear terms $\hat{\mathcal{N}}_j$ in (3.20) truncated on $p \in D_{\varepsilon^{r-1}}$, $\varepsilon^{\tilde{r}} \hat{R}_j(p)$ denotes the remainder, and we set $\tilde{r} = \min(3r - 1, 2 - 2r)$. The contribution of $\vec{\tilde{\psi}}$ to $\varepsilon^{\tilde{r}} \hat{R}_j(p)$ is estimated via (4.13) to be $\mathcal{O}(\varepsilon^{2-2r})$, which requires $r < 1$ to obtain small $\varepsilon^{\tilde{r}} \hat{R}_j(p)$. The next contributions come from third order derivatives of the bands ω_{n_j} and are of the form $\varepsilon^2 \partial_{k_{m_1}} \partial_{k_{m_2}} \partial_{k_{m_3}} \omega_{n_j}(k^{(j)}) p_{m_1} p_{m_2} p_{m_3} \hat{B}_j$, $m_{1,2,3} \in \{1, 2\}$, and therefore bounded by

$$C\varepsilon^2 \| |p|^3 \hat{B}_j \|_{L_s^2(D_{\varepsilon^{r-1}})} \leq C\varepsilon^{3r-1} \| \hat{B}_j \|_{L_s^2(D_{\varepsilon^{r-1}})},$$

enforcing $r > 1/3$.

From this we obtain our first main result:

Theorem 4.6 *Let $s > 1, V \in H_{\mathrm{loc}}^{\lceil s \rceil - 1 + \delta}(\mathbb{R}^2)$ for some $\delta > 0$ and let $\frac{1}{3} < r < 1$. 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 (4.14) with $\| \hat{B}_j \|_{L_s^2(D_{\varepsilon^{r-1}})} \leq C_1$. Then (1.2) has a solution $\phi \in H^s(\mathbb{R}^2)$ with*

$$\| \phi(\cdot) - \varepsilon \sum_{j=1}^N B_j(\varepsilon \cdot) u_{n_j}(k^{(j)}; \cdot) \|_{H^s(\mathbb{R}^2)} \leq C_2 \varepsilon^{2-2r}, \qquad (4.15)$$

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 (4.14) yields via (4.8) a solution $\vec{\tilde{\phi}}(k)$ of (4.6). The norm equivalence in Lemma 4.1 and (4.13) then yield (4.15). $\qquad\qquad \square$

4.3 Justification Step II: Persistence

The remaining step is to make a connection between solitary waves of the CMEs (3.20) and the eCMEs (4.14). To obtain existence of solutions of the eCMEs (4.14) we show a persistence result of special so called reversible CME solitons to the eCME. This is quite similar to [13, §5] but in order to deal with the case $k^{(j)} \notin S$ our notion of reversibility is somewhat more restrictive. Therefore we again repeat the main steps.

The eCMEs (4.14) differ from the CMEs (3.20) 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 $\hat{\mathbf{B}} = (\hat{B}_j)_{j=1,\ldots,N}$ of the eCMEs near a given solution $\hat{\mathbf{A}} = (\hat{A}_j)_{j=1,\ldots,N}$ 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

$$\chi_{D_{\varepsilon^{r-1}}}\hat{\mathbf{F}}(\hat{\mathbf{B}}) = \varepsilon^{\tilde{r}}\hat{\mathbf{R}}(\hat{\mathbf{B}}), \tag{4.16}$$

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

$$\begin{aligned}
&\hat{\mathbf{J}}\hat{\mathbf{b}} = \hat{\mathbf{N}}(\hat{\mathbf{b}}), \ p \in D_{\varepsilon^{r-1}} \quad \text{with} \\
&\hat{\mathbf{J}} = \chi_{D_{\varepsilon^{r-1}}}D_{\hat{\mathbf{A}}}\hat{\mathbf{F}}(\hat{\mathbf{A}}) \quad \text{and} \quad \hat{\mathbf{N}}(\hat{\mathbf{b}}) = \varepsilon^{\tilde{r}}\hat{\mathbf{R}}(\hat{\mathbf{A}}_\varepsilon + \hat{\mathbf{b}}) - (\chi_{D_{\varepsilon^{r-1}}}\hat{\mathbf{F}}(\hat{\mathbf{A}}_\varepsilon + \hat{\mathbf{b}}) - \hat{\mathbf{J}}\hat{\mathbf{b}}).
\end{aligned} \tag{4.17}$$

Note that in (4.17) 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 $\hat{\mathbf{F}}(\hat{\mathbf{A}}_\varepsilon + \hat{\mathbf{b}}) - \hat{\mathbf{J}}\hat{\mathbf{b}} = \hat{\mathbf{F}}(\hat{\mathbf{A}}_\varepsilon) + (D_{\hat{\mathbf{A}}}\hat{\mathbf{F}}(\hat{\mathbf{A}}_\varepsilon) - \hat{\mathbf{J}})\hat{\mathbf{b}} + \hat{\mathbf{G}}(\hat{\mathbf{b}})$ with $\hat{\mathbf{G}}(\hat{\mathbf{b}})$ quadratic in $\hat{\mathbf{b}}$. Moreover,

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

where L^{CME} denotes the linear part of the operator in (3.20) and $\hat{\mathcal{N}}$ denotes the $N-$dimensional vector of the nonlinear terms in (3.20). $\chi_{D_{\varepsilon^{r-1}}}\hat{\mathcal{N}}(\hat{\mathbf{A}})(p)$ is a sum of convolutions $\hat{A}_{j_1} * \hat{A}_{j_2} * \bar{\hat{A}}_{j_3}$, 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

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

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,

$$\|\hat{\mathbf{F}}(\hat{\mathbf{A}}_\varepsilon)\|_{L_s^2} \leq C\varepsilon^q, \quad \text{and} \quad \|(D_{\hat{\mathbf{A}}}\hat{\mathbf{F}}(\hat{\mathbf{A}}_\varepsilon) - \hat{\mathbf{J}})\hat{\mathbf{b}}\|_{L_s^2} \leq C\varepsilon^q\|\hat{\mathbf{b}}\|_{L_s^2} \tag{4.18}$$

by a similar estimate.

Thus $\|\hat{\mathbf{N}}(\hat{\mathbf{b}})\|_{L_s^2} \leq C(\varepsilon^{\tilde{r}} + \varepsilon^q + \varepsilon^q\|\hat{\mathbf{b}}\|_{L_s^2} + \|\hat{\mathbf{b}}\|_{L_s^2}^2)$, and this suggests to apply the contraction mapping theorem to (4.17) in the form

$$\hat{\mathbf{b}} = \hat{\mathbf{J}}^{-1}\hat{\mathbf{N}}(\hat{\mathbf{b}}). \tag{4.19}$$

To discuss $\hat{\mathbf{J}}^{-1}$ 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 $\operatorname{Ker}\mathbf{J}$ 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 $\hat{\mathbf{J}}^{-1} : L_s^2(D_{\varepsilon^{r-1}}) \to L_{s+2}^2(D_{\varepsilon^{r-1}})$ (if it exists) this implies that it cannot be bounded independently of ε.

The solution is to consider (4.19) in a subspace $X_{\mathrm{rev}} \subset L^2(D_{\varepsilon^{r-1}})$ where $\hat{\mathbf{J}}^{-1}$ 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.

24

The original problem (1.2) is equivariant under the symmetries

$$\phi(x_1, x_2) = s_1 \bar{\phi}(-x_1, x_2) = s_2 \bar{\phi}(x_1, -x_2) \tag{4.20}$$

where $s_1, s_2 = \pm 1$, i.e., $(s_1, s_2) = (1, 1)$ or $(s_1, s_2) = (1, -1)$ or $(s_1, s_2) = (-1, 1)$ or $(s_1, s_2) = (-1, -1)$, and similarly

$$\phi(x_1, x_2) = s_1 \bar{\phi}(x_2, x_1) = s_2 \bar{\phi}(-x_2, -x_1), \tag{4.21}$$

where again $s_1, s_2 = \pm 1$.

Definition 4.7 *A solution* $\mathbf{A}$ *of (3.21) is called reversible if it fulfills one of the symmetries (4.20),(4.21), and if the corresponding lowest order approximation* $\phi^{(0)}$ *fulfills the same symmetry.* $\mathbf{A}$ *is called non-degenerate if* $Ker\, \mathbf{J} = \{\partial_{y_1}\mathbf{A}, \partial_{y_2}\mathbf{A}, i\mathbf{A}\}$.

Remark 4.8 If all $k^{(j)} \in S = \{\Gamma, X, M, X'\}$, then the second condition in the definition of reversibility follows from the first by the fact that the Bloch functions are real and either even or odd in both variables. We expect that the second condition is typically satisfied even if some $k^{(j)} \notin S$ but we refrain from a proof. Instead, in §5 we verify this for our specific examples. ⌋

A given symmetry, say S, is then inherited by the Lyapunov Schmidt reduction, see Prop. 3.3 in [18], i.e.,

$$\chi_{D_{\varepsilon^{r-1}}}\hat{\mathbf{F}}(\hat{S}\hat{\mathbf{B}}) = \hat{S}\chi_{D_{\varepsilon^{r-1}}}\mathbf{F}(\hat{\mathbf{B}}) = \hat{S}\varepsilon^{\tilde{r}}\hat{\mathbf{R}}(\hat{\mathbf{B}}) = \varepsilon^{\tilde{r}}\hat{\mathbf{R}}(\hat{S}\hat{\mathbf{B}}), \tag{4.22}$$

and therefore

$$\hat{S}\hat{\mathbf{N}}(\hat{\mathbf{b}}) = \hat{\mathbf{N}}(\hat{S}\hat{\mathbf{b}}). \tag{4.23}$$

Finally, let Ker $\mathbf{J} = \{\partial_{y_1}\mathbf{A}, \partial_{y_2}\mathbf{A}, i\mathbf{A}\}$. If $S\mathbf{A} = \mathbf{A}$ then $\partial_{y_1}\mathbf{A}, \partial_{y_2}\mathbf{A}$ and $i\mathbf{A}$ cannot fulfill the same symmetry and $\hat{\mathbf{J}}$ has an $\mathcal{O}(1)$ bounded inverse on $X_{\mathrm{rev}} = \{\hat{\mathbf{A}} \in L_s^2(D_{\varepsilon^{r-1}}) : \hat{S}\hat{\mathbf{A}} = \hat{\mathbf{A}}\}$, i.e., with bound independent of ε. Using (4.23) we may thus seek $\hat{\mathbf{b}}$ in X_{rev} and obtain the following

Theorem 4.9 *Let* $s > 1$. *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 (3.21) with* $\hat{\mathbf{A}} \in L_q^2(\mathbb{R}^2)$ *for all* $q \geq 0$, *and let* $1/3 < r < 1$. *Then there exists a* $C > 0$ *and a solution* $\hat{\mathbf{B}}$ *of the eCME such that*

$$\|\hat{\mathbf{B}} - \hat{\mathbf{A}}\|_{L_s^2(D_{\varepsilon^{r-1}})} \leq C\varepsilon^{\tilde{r}} \qquad \tilde{r} = \min(3r - 1, 2 - 2r). \tag{4.24}$$

Corollary 4.10 *The solution* ϕ *constructed in Theorems 4.6 and 4.9 is a reversible localized solution of (1.2), and*

$$\left\|\phi - \varepsilon \sum_{j=1}^{N} A_j(\varepsilon\cdot)u_{n_j}(k^{(j)}; \cdot)\right\|_{H^s(\mathbb{R}^2)} \leq C\varepsilon^{\tilde{r}}. \tag{4.25}$$

Proof. ϕ is reversible as ψ in (4.10) inherits the reversibility symmetry. The estimate (4.25) follows from the triangle inequality. ▢

Remark 4.11 The optimal value of $\tilde{r}$ is $\tilde{r} = 4/5$ attained at $r = 3/5$. Based on the formal asymptotic expansion in (3.10) and (3.16), 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 (4.25) 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 $\tilde{r} = \min(4r-1, 2-2r)$ with $1/4 < r < 1$ 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 $\mathcal{O}(\delta_1)$ 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

$$\|\psi\|_{H^s} \geq \frac{C_3 C_1 \varepsilon^\alpha \delta}{4\varepsilon} > C \varepsilon^{\alpha-1} \tag{4.26}$$

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. ⌋
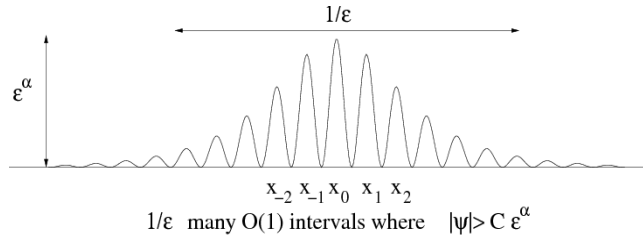

Figure 7: Sketch of the argument for $\|\psi\|_{H^s} = \mathcal{O}(\varepsilon^\beta) \Rightarrow \|\psi\|_{L^\infty} = \mathcal{O}(\varepsilon^{\beta+1})$.

26

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 impossible to occur 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 of the error $\|\phi_{\mathrm{GS}}^{\mathrm{num}} - \varepsilon\phi^{(0)}\|_{H^2}$.

A 4th order centered finite difference discretization is used for (1.2). The computational domain is a square $x \in [-D_{\mathrm{GS}}/2, D_{\mathrm{GS}}/2]^2$ 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 $\omega = s_2$ 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 $A(y) = R(r)e^{im\theta}$, where $r = \frac{1}{\sqrt{\alpha}}\sqrt{y_1^2 + y_2^2}, \theta = \arg(y_1 + iy_2)$, 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 (4.20) with $s_1 = s_2 = 1$ since A is real. Because $u_1(M; x)$ is also even in both x_1 and x_2 and real, the same reversibility property holds for $\varepsilon\phi^{(0)}(x) = \varepsilon A(\varepsilon x)u_1(M; x)$. 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.
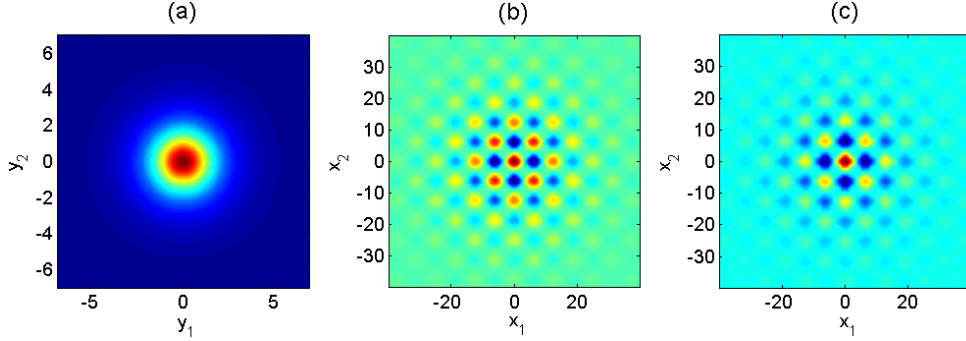
4.11.

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 $\omega = s_2 + \varepsilon^2\Omega$.
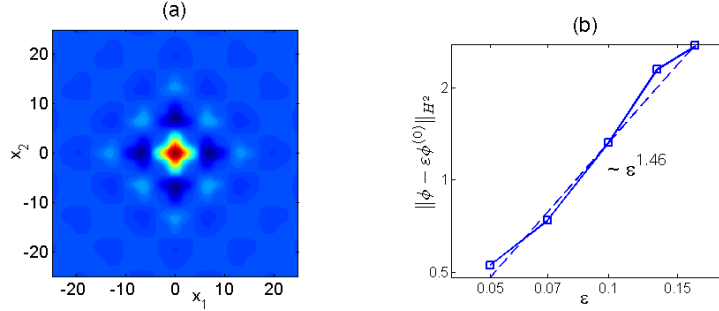

Figure 9: (a) Profile of a GS corresponding to the even real family that bifurcates from $\omega = s_2$ in Fig. 8. The plotted GS is deep inside the gap (s_2, s_3) at $\omega \approx 1.78$ (corresponding to $\varepsilon \approx 0.28$). (b) ε-convergence of the error $\|\phi - \varepsilon\phi^{(0)}\|_{H^2(\mathbb{R}^2)}$.

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 (4.20) with $s_1 = -s_2 = -1$, and due to the even symmetries of the real Bloch functions $u_1(M;x)$ we have the same for $\phi^{(0)}$. 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 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.

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 (4.21) with
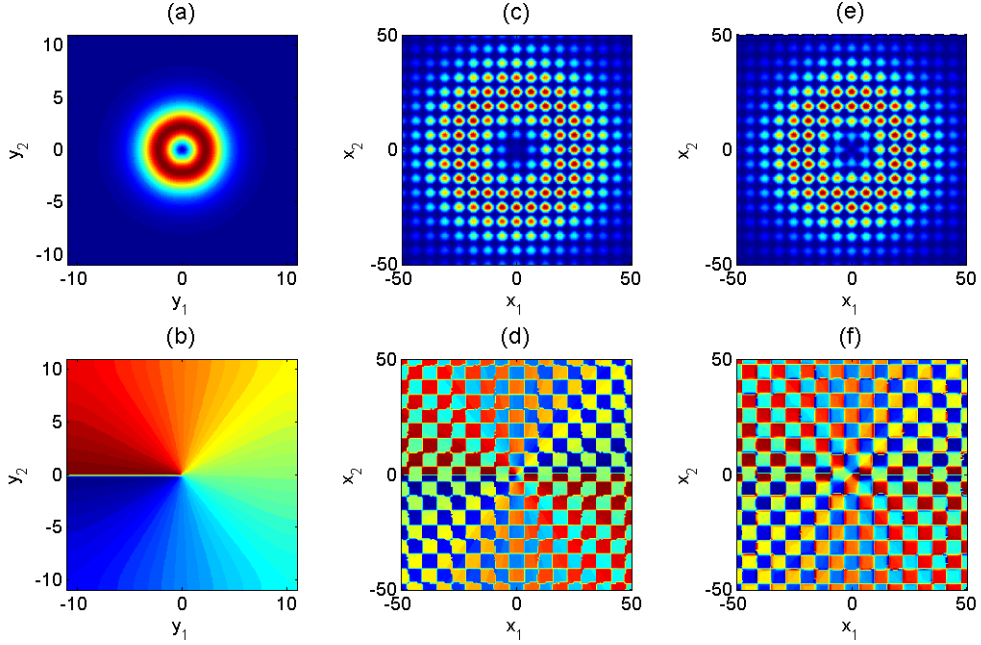
28

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.
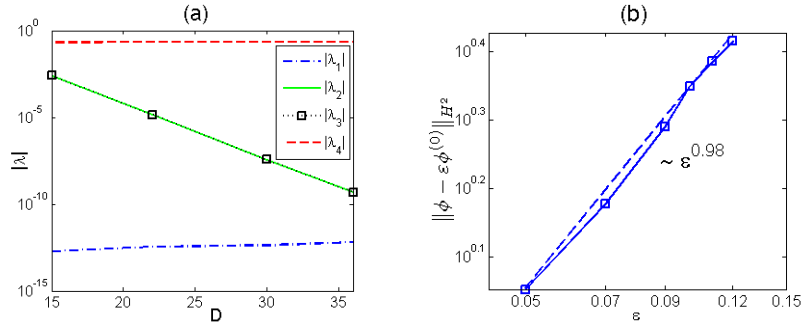

Figure 11: (a) The four smallest eigenvalues of the Jacobian $\mathbf{J}$ in Theorem 4.9 at the solution $\mathbf{A}$ in Fig. 10 (a-d) for a range of sizes of the computational domain. (b) ε-convergence of the error $\|\phi - \varepsilon\phi^{(0)}\|_{H^2(\mathbb{R}^2)}$.

$s_1 = s_2 = 1$ for each A_j. Such solutions of the CME system (3.27) 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 (3.27) 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

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

Using (5.3) and the fact

$$v_{1,3}(x_2, x_1) = v_{1,3}(x_1, x_2) \quad \text{(see §2),} \quad (5.4)$$

we obtain

$$v_2(x_2, x_1) = v_4(x_1, x_2). \quad (5.5)$$

Due to (5.4) and (5.5) the asymptotic approximation $\varepsilon\phi^{(0)}(x) = \varepsilon \sum_{j=1}^4 A_j(\varepsilon x)v_j(x)$ also satisfies (4.21) with $s_1 = s_2 = 1$. For instance,

$$\phi^{(0)}(x_2, x_1) = A_1(y_2, y_1)(v_1(x_2, x_1) + v_3(x_2, x_1)) + A_2(y_2, y_1)(v_2(x_2, x_1) + v_4(x_2, x_1))$$
$$= A_1(y)(v_1(x) + v_3(x)) + A_2(y)(v_4(x) + v_2(x)) = \phi^{(0)}(x_1, x_2).$$

The solution $\mathbf{A}$ is thus reversible according to Def. 4.7. Moreover, the normalization (5.3) also implies that $\varepsilon\phi^{(0)}(x)$ is real and even in both variables, i.e., it satisfies (4.20) with $s_1 = s_2 = 1$. 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 (3.27), but only 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.

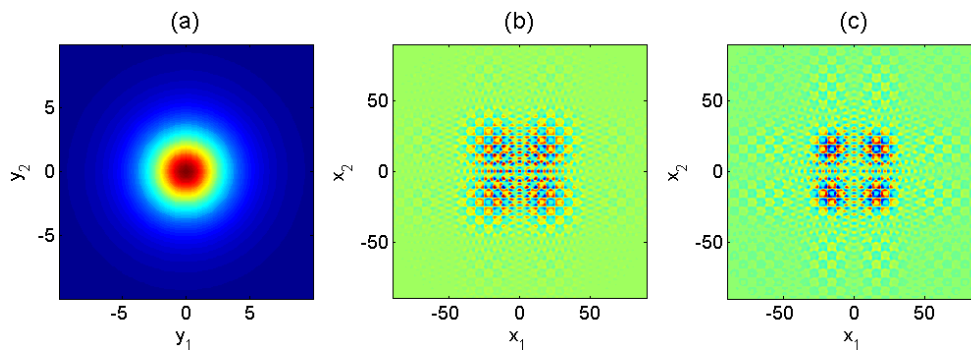Figure 12: Profiles of the even, real GS at $\omega = s_5 + \varepsilon^2\Omega, \varepsilon = 0.1, \Omega = -1$. (a) $A_1(y)$; (b) the corresponding leading-order GS approximation $\varepsilon\sum_{j=1}^4 A_j(y)v_j(x)$; (c) the numerically computed GS at $\omega = s_5 + \varepsilon^2\Omega$.

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 $\varepsilon^{0.94}$,

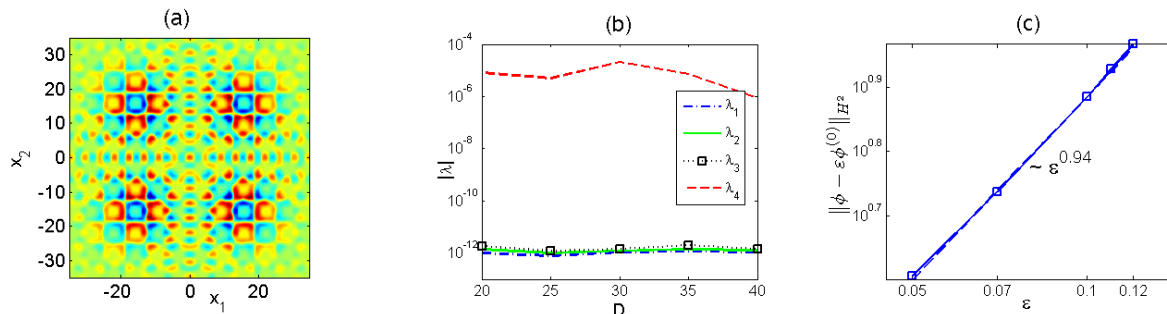which is once again close to the rate ε^1 predicted by the formal asymptotics.

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

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. In contrast to [13] we make no assumption on the form of the periodic potential $V(x)$. 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 $\mathcal{O}(\varepsilon^{4/5})$ for GSs with the spectral parameter $\mathcal{O}(\varepsilon^2)$ close to the band edge. 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.

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. (2008) online.

[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 (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, 1995, suppl., 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).