LONDON
MATHEMATICAL
SOCIETY
EST. 1865

IOP Publishing

PAPER • OPEN ACCESS

Justification of the asymptotic coupled mode approximation of out-of-plane gap solitons in Maxwell equations

View the article online for updates and enhancements.

Justification of the asymptotic coupled mode approximation of out-of-plane gap solitons in Maxwell equations

Tomáš Dohnal[*] and **Giulio Romani**

Institut für Mathematik, Martin-Luther-Universität Halle-Wittenberg, 06099 Halle (Saale), Germany

E-mail: tomas.dohnal@mathematik.uni-halle.de and giulio.romani@mathematik.uni-halle.de

CrossMark

Abstract

In periodic media gap solitons with frequencies inside a spectral gap but close to a spectral band can be formally approximated by a slowly varying envelope ansatz. The ansatz is based on the linear Bloch waves at the edge of the band and on effective coupled mode equations (CMEs) for the envelopes. We provide a rigorous justification of such CME asymptotics in two-dimensional photonic crystals described by the Kerr nonlinear Maxwell system. We use a Lyapunov–Schmidt reduction procedure and a nested fixed point argument in the Bloch variables. The theorem provides an error estimate in $H^2(\mathbb{R}^2)$ between the exact solution and the envelope approximation. The results justify the formal and numerical CME-approximation in Dohnal and Dörfler, [2013 Multiscale Model. Simul. 11 162–191].

(Some figures may appear in colour only in the online journal)

1. Introduction

Maxwell's equations in Kerr nonlinear dielectric materials without free charges are described by

$$\mu_0 \partial_t H = -\nabla \times E, \qquad \epsilon_0 \partial_t D = \nabla \times H, \qquad \nabla \cdot D = \nabla \cdot H = 0, \qquad (1.1)$$

where $E = (E_1, E_2, E_3)$ and $H = (H_1, H_2, H_3)$ are the electric and the magnetic field respectively, $D = (D_1(E), D_2(E), D_3(E))$ is the electric displacement field and ϵ_0 and μ_0 are the permittivity and the permeability of the free space, respectively. We assume the constitutive relations

$$D(x, t) = \epsilon_0(\epsilon(x_1, x_2)E(x, t) + \mathcal{F}(E)(x, t)), \quad x \in \mathbb{R}^3, \ t \in \mathbb{R},$$

where

$$\mathcal{F}_d(E)(x, t) := \sum_{a,b,c=1}^{3} \chi_{a,b,c,d}^{(3)}(x_1, x_2)(E_a E_b E_c)(x, t), \quad d \in \{1, 2, 3\}. \qquad (1.2)$$

We model a two dimensional photonic crystal and hence assume that the dielectric function (relative permittivity) $\epsilon : \mathbb{R}^2 \to \mathbb{R}$ and the cubic electric susceptibility $\chi^{(3)} : \mathbb{R}^2 \to \mathbb{R}^{3 \times 3 \times 3 \times 3}$ are periodic and ϵ is positive. The periodicity is specified by two linearly independent lattice vectors $a^{(1)}, a^{(2)} \in \mathbb{R}^2$ defining the *Bravais lattice* $\Lambda := \operatorname{span}_{\mathbb{Z}}\{a^{(1)}, a^{(2)}\}$ of the crystal. Then the required periodicity reads

$$\epsilon(x) = \epsilon(x + R), \qquad \chi^{(3)}(x) = \chi^{(3)}(x + R) \quad \text{for any } x \in \mathbb{R}^2 \text{ and } R \in \Lambda. \qquad (1.3)$$

Since $\epsilon = \epsilon(x_1, x_2)$ and $\chi^{(3)} = \chi^{(3)}(x_1, x_2)$, the material is homogeneous in the x_3-direction. In the following $\mathcal{Q} \subset \mathbb{R}^2$ denotes the Wigner–Seitz periodicity cell, defined as the set of points in $\mathbb{R}^2$ which are closer to 0 than to any other lattice point in Λ. More precisely,

$$\mathcal{Q} := \{x \in \mathbb{R}^2 \mid |x| < |x - Z|, \quad Z \in \Lambda \backslash \{0\}\} \cup S,$$

where the connected subset $S \subset \partial \mathcal{Q}$ is chosen so that $\bigcup_{Z \in \Lambda}(\mathcal{Q} + Z) = \mathbb{R}^2$ and $\mathcal{Q} \cap (\mathcal{Q} + Z) = \emptyset$ for all $Z \in \Lambda$. We often use the term Λ-periodic to mean the periodicity as in (1.3).

We consider monochromatic waves propagating in the homogeneous x_3-direction, i.e. out of the plane of periodicity of the 2D crystal, and use the ansatz

$$(E, H)(x, t) = e^{i(\kappa x_3 - \omega t)}(u, h)(x_1, x_2; \omega) + \text{c.c.} \qquad (1.4)$$

where $\kappa \in \mathbb{R}$ and c.c. denotes the complex conjugate. We look for profiles u, h localized in both x_1 and x_2 and with ω in a frequency gap. The resulting solutions are called *out-of-plane gap solitons*. Inserting such a monochromatic ansatz into the nonlinearity (1.2) and neglecting the higher harmonics[1], one obtains for $d \in \{1, 2, 3\}$

$$\mathcal{F}_d(E)(x, t) = \sum_{a,b,c=1}^{3} \chi_{a,b,c,d}^{(3)}(x_1, x_2)(\overline{u}_a u_b u_c + u_a \overline{u}_b u_c + u_a u_b \overline{u}_c)(x_1, x_2) e^{i(\kappa x_3 - \omega t)} + \text{c.c.}.$$

---

[1] Neglecting higher harmonics is a common approach in theoretical studies of weakly nonlinear optical waves [41]. Alternatively, one can use a time averaged model for the nonlinear part of the displacement field, see [43–45], where no higher harmonics appear.

We define

$$F_d(u) := \sum_{a,b,c=1}^{3} \chi^{(3)}_{a,b,c,d}(\overline{u}_a u_b u_c + u_a \overline{u}_b u_c + u_a u_b \overline{u}_c)$$

$$= \sum_{a,b,c=1}^{3} \underline{\chi}^{(3)}_{a,b,c,d} u_a u_b \overline{u}_c, \quad d \in \{1,2,3\},$$

where

$$\underline{\chi}^{(3)}_{a,b,c,d} := \chi^{(3)}_{c,b,a,d} + \chi^{(3)}_{a,c,b,d} + \chi^{(3)}_{a,b,c,d}.$$

We rescale the frequency by defining $\tilde{\omega} = \frac{\omega}{c}$, where $c = (\mu_0 \epsilon_0)^{-1/2}$, but drop the tilde again for better readability. Then, with the ansatz in (1.4) Maxwell's equation (1.1) become

$$\begin{cases} \nabla' \times u - i\mu_0 c \omega h = 0, \\ \nabla' \times h + i\epsilon_0 c \omega \epsilon(\cdot) u = i\epsilon_0 c \omega F(u), \end{cases} \tag{1.5}$$

where $\nabla' := (\partial_1, \partial_2, i\kappa)^T$ is the restriction of the standard ∇ applied to our 2D-ansatz (1.4). Notice that, indeed, the divergence equations in (1.1) are automatically satisfied by our ansatz. Equivalently, we may write (1.5) as a second-order equation for the electric field u,

$$(L^{(E)} - \omega^2 \epsilon(\cdot))u := \nabla' \times \nabla' \times u - \omega^2 \epsilon(\cdot) u = \omega^2 F(u), \tag{1.6}$$

and then, having determined a solution u, the magnetic field can be recovered by

$$h = -\frac{i}{\mu_0 c \omega} \left(\nabla' \times u \right).$$

For any ω in a spectral gap of the linear problem $(L^{(E)} - \omega^2 \epsilon(\cdot))u = 0$ equation (1.6) is expected to have localized solutions u with $u(x_1, x_2) \to 0$ as $|(x_1, x_2)| \to \infty$, called *gap solitons*. This has been proved variationally for other problems, e.g. the periodic Gross–Pitaevskii equation, see [33], or equation (1.6) with other (not periodic) coefficients ϵ and $\chi^{(3)}$, see e.g. [4, 31]. From the physics point of view, gap solitons are phenomenologically interesting as they achieve a balance between the periodicity induced dispersion and the focusing or defocusing of the non-linearity. In addition, they exist for frequencies in spectral gaps, i.e. where no linear propagation is possible. Examples of physics references for gap solitons in two dimensions are [20] or [42, section 16.6].

In [43] an approximation of gap solitons of (1.6) with periodic coefficients and for ω in an asymptotic vicinity of a gap edge was formally obtained using a slowly varying envelope approximation. In particular, envelopes of such gap solitons satisfy a system of nonlinear equations with constant coefficients, so-called *couple mode equations* (CMEs), posed in a slow variable. The advantage is that such a system can be numerically solved with less effort than the original Maxwell system (1.6), which is posed in the 'fast' variable x. Then, the solution of (1.6) for ω near a band edge would be asymptotically approximated by the sum of linear Bloch waves at the edge modulated by the corresponding envelopes, see ansatz (1.8). The aim of this paper is to give a rigorous justification of this approximation.

Let us now describe the approximation in more detail. First, recall that the (first) Brillouin zone (here denoted by $\mathbb{B} \subset \mathbb{R}^2$) is the Wigner–Seitz periodicity cell for the reciprocal lattice $\Lambda^* := \text{span}_{\mathbb{Z}}\{b^{(1)}, b^{(2)}\} \subset \mathbb{R}^2$. The vectors $b^{(1)}, b^{(2)}$ satisfy $a^{(i)} \cdot b^{(j)} = 2\pi \delta_{ij}$ for $i, j \in \{1, 2\}$, with δ_{ij} being the Kronecker-delta.

Let now ω_* be a boundary point of the (real) spectrum of the pencil $L^{(E)} - \omega^2 \epsilon$, i.e. such that there is a choice of $\Omega \in \{-1, +1\}$ for which $\omega_* + \tau\Omega$ lies outside and $\omega_* - \tau\Omega$ inside the spectrum for all $\tau > 0$ small enough. We are interested in studying solutions of equation (1.6) when ω lies in a band gap and is asymptotically close to ω_*. Hence, in our main result (theorem 1.1), we choose a small parameter $0 < \varepsilon \ll 1$ and set

$$\omega = \omega_* + \Omega\varepsilon^2, \tag{1.7}$$

where $\Omega \in \{-1, +1\}$ is chosen such that ω lies outside the spectrum. We aim to study the existence of a solution of (1.6) close to the *slowly varying envelope* ansatz

$$u_{\text{ans}}(x) = \varepsilon \sum_{j=1}^{N} A_j(\varepsilon x) u_{n_*}(x, k^{(j)}). \tag{1.8}$$

Here $(A_j)_{j=1}^{N}$ are localized envelopes to be determined below and the function $u_{n_*}(x, k)$, for $n_* \in \mathbb{N}$, is a Bloch wave of $L^{(E)} - \omega_*^2\epsilon$. Writing

$$\nabla'_k := \nabla' + ik \quad \text{for } k \in \mathbb{R}^2,$$

a *Bloch wave* $u_n(\cdot, k): \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{C}^3$ is defined as

$$u_n(x, k) := p_n(x, k)e^{ik \cdot x}, \quad n \in \mathbb{N},$$

where $p_n(\cdot, k)$ is a solution of the periodic eigenvalue problem

$$\begin{aligned}\nabla'_k \times \nabla'_k \times p_n(x, k) &= \omega_n(k)^2 \epsilon(x) p_n(x, k), \quad \text{for all } x \in \mathbb{R}^2, \\ p_n(x + R, k) &= p_n(x, k) \quad \text{for all } x \in \mathbb{R}^2 \text{ and all } R \in \Lambda,\end{aligned} \tag{1.9}$$

which is to be solved for the eigenpair $(\omega_n(k), p_n(\cdot, k))$. Let us choose the normalization

$$\int_{\mathcal{Q}} \epsilon(x)|p_n(x, k)|^2 \, dx = 1, \ n \in \mathbb{N}, \ k \in \mathbb{B},$$

see (2.11) below. Note that for a geometrically simple eigenvalue the eigenfunction $p_n(\cdot, k)$ is unique up to a phase factor $e^{i\alpha}, \alpha \in \mathbb{R}$. We call $p_n(\cdot, k)$ a *Bloch eigenfunction*. Clearly, Bloch waves are quasiperiodic

$$u_n(x + R, k) = u_n(x, k)e^{ik \cdot R} \quad \text{for all } x \in \mathbb{R}^2 \text{ and all } R \in \Lambda.$$

Note that, strictly speaking, a 'Bloch wave' is the time dependent function $u_n(x, k)e^{i\omega_n(k)t}$ but, for the purpose of this paper, we use this name for the factor $u_n(x, k)$.

We assume that the band structure $k \mapsto \{\omega_n(k) : n \in \mathbb{N}\}$ attains the value ω_* at finitely many points, denoted $k^{(1)}, \ldots, k^{(N)} \in \mathbb{B}$, see (1.8). In other words, there exist indices $(n_j)_{j=1}^{N}$ for which $\omega_{n_j}(k^{(j)}) = \omega_*$ holds, with $k \mapsto \omega_{n_j}(k)$ defined by (1.9). Notice that since at each $k \in \mathbb{B}$ the eigenvalues $\omega_n(k)$ are ordered by magnitude, we necessarily have $\omega_{n_*}(k^{(j)}) = \omega_*$ for some $n_* \in \mathbb{N}$ and all $j \in \{1, \ldots, N\}$. Due to the additional assumption of geometric simpleness at $k \in \{k^{(1)}, \ldots, k^{(N)}\}$ at the level ω_*, there is only one Bloch eigenfunction (up to a complex

phase factor) at $k \in \{k^{(1)}, \ldots, k^{(N)}\}$ at the level ω_*. We denote this eigenfunction by $p_{n_*}(\cdot, k^{(j)})$. For a precise formulation of the corresponding assumptions see assumption (A3) in section 3. Also note that the spectrum of the pencil $L^{(E)} - \omega^2 \epsilon$ equals the union of the ranges of the functions ω_n over all n, see sections 2.2 and 2.3.

In figure 1 we plot the material structure, the Brillouin zone, and the band structure for an example adopted from [43]. Only the band structure along the boundary of an 'irreducible' Brillouin zone $\mathbb{B}_0$ is plotted, which is standard practice in the physics literature. It was checked in [43] that the level sets of the first five band edges do not include any points from the interior of $\mathbb{B}_0$. Five spectral edges $s_1, \ldots, s_5$ are labeled. Note that, for instance, the edge $\omega_* = s_3$ has $N = 1$ as the level set includes only the point $k^{(1)} = \Gamma = (0, 0)$. At $\omega_* = s_5$ we have $N = 6$ because the minimal point along the line Γ–M is repeated 6 times in the full Brillouin zone $\mathbb{B}$ due to a discrete rotational symmetry of the lattice.

As shown in [43], in order for the residual $L^{(E)} u_{\text{ans}} - \omega^2 \epsilon u_{\text{ans}} - \omega^2 F(u_{\text{ans}})$ to be small, the functions $\left(A_j\right)_{j=1}^{N}$ in (1.8) have to satisfy the second-order CMEs

$$\Omega A_j + \frac{1}{2} \left(\partial_{k_1}^2 \omega_{n_*}(k^{(j)}) \partial_{y_1}^2 + \partial_{k_2}^2 \omega_{n_*}(k^{(j)}) \partial_{y_2}^2 + 2 \partial_{k_1} \partial_{k_2} \omega_{n_*}(k^{(j)}) \partial_{y_1} \partial_{y_2} \right) A_j + \mathcal{N}_j = 0 \tag{1.10}$$

in $\mathbb{R}^2$, where $y := \varepsilon x$ is the slow variable and the nonlinear term $\mathcal{N}_j$ is given by

$$\mathcal{N}_j = \sum_{(\alpha, \beta, \gamma) \in \sigma_j} I_{\alpha, \beta, \gamma}^j A_\alpha A_\beta \overline{A_\gamma}. \tag{1.11}$$

where

$$\sigma_j := \{ (\alpha, \beta, \gamma) \in \{1, \ldots, N\}^3 : k^{(\alpha)} + k^{(\beta)} - k^{(\gamma)} - k^{(j)} \in \Lambda^* \}. \tag{1.12}$$

The coefficients $I_{\alpha, \beta, \gamma}^j$ are determined by the Bloch wave u_{n_*} at the points $k^{(j)}$, in detail

$$I_{\alpha, \beta, \gamma}^j := \frac{\omega_*}{2} \sum_{a, b, c, d = 1}^{3} \left\langle \chi_{a, b, c, d}^{(3)} u_{n_*, a}(\cdot, k^{(\alpha)}) u_{n_*, b}(\cdot, k^{(\beta)}) \overline{u_{n_*, c}}(\cdot, k^{(\gamma)}), u_{n_*, d}(\cdot, k^{(j)}) \right\rangle. \tag{1.13}$$

The formal derivation of the CMEs (1.10) as an effective model for the envelopes A_j can be summarized as follows. First, ansatz (1.8) is inserted into (1.6) and for each j the terms proportional to $e^{ik^{(j)} \cdot x}$ times a Λ-periodic function are collected. Then, setting the $L^2(\mathcal{Q})$-inner product of the leading order part of these terms with $u_{n_*}(\cdot, k^{(j)})$ to zero, produces the jth equation in (1.10). In the inner product the variable $y := \varepsilon x$ is considered independent of x.

For several examples with the coefficients $\nabla^2 \omega_{n_*}(k^{(j)})$ and $I_{\alpha, \beta, \gamma}^j$ obtained from actual Bloch waves of the corresponding Maxwell problem, localized solutions were found numerically in [43]. Figures 2(a) and (b) show an example solution of CMEs (1.10) corresponding to $\omega_* = s_5$ in figure 1. We also plot the total intensity $|u_{\text{ans}}|^2$ of the corresponding formal approximation (1.8) at $\omega = \omega_* - \varepsilon^2$ with $\varepsilon = 0.1$ in figure 2(c). All plots in figure 2 are adopted from [43].

The main result of this paper is that for $\varepsilon > 0$ small enough the existence of suitable solutions of CMEs (1.10) implies the existence of gap solitons of (1.6) with ω given by (1.7). These gap solitons are approximated by the ansatz (1.8). The following theorem uses assumptions (A1)–(A7) on the band structure and on the functions ϵ and $\chi^{(3)}$, see section 3, the
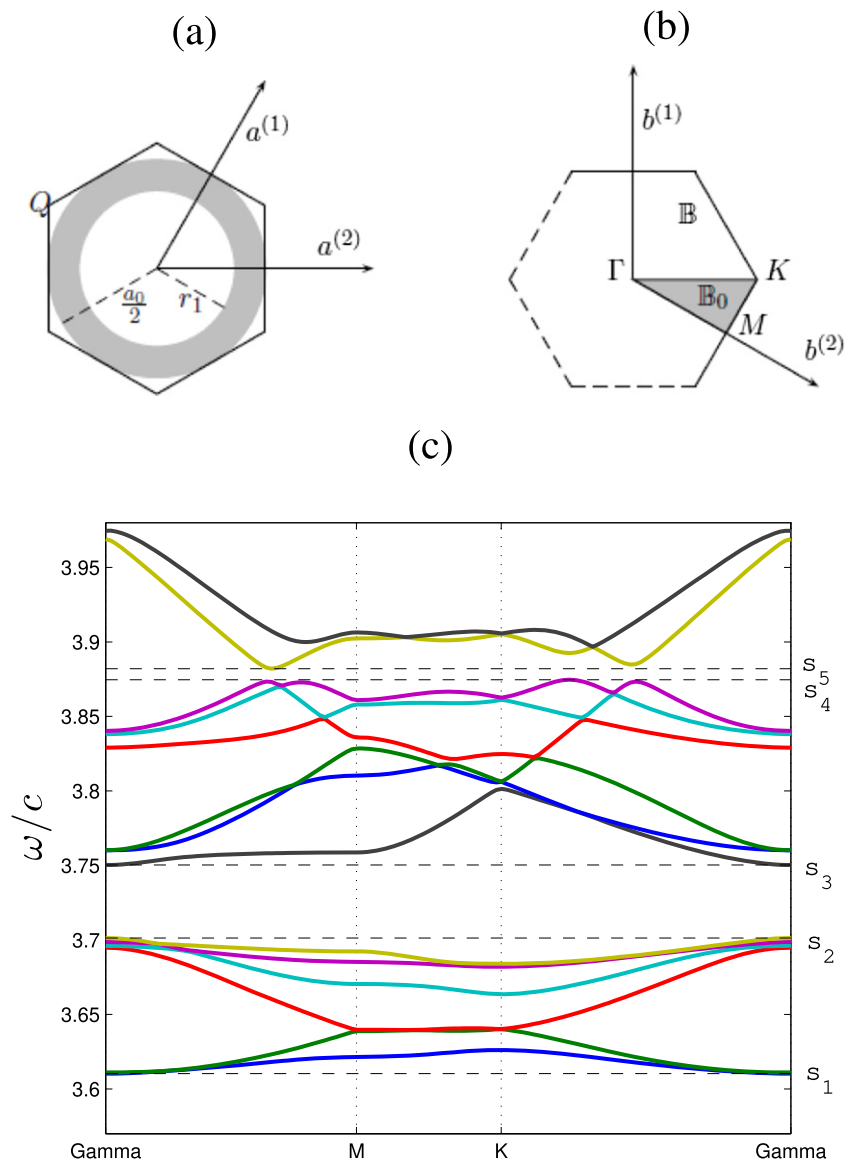
Figure 1. (a) Hexagonal periodicity cell $\mathcal{Q}$ with a cylindrical material structure and the lattice vectors $a^{(1)}, a^{(2)}$. (b) The corresponding Brillouin zone $\mathbb{B}$ (not to scale) with the reciprocal lattice vectors $b^{(1)}, b^{(2)}$ and a shaded irreducible Brillouin zone $\mathbb{B}_0$. (c) Band structure along $\partial \mathbb{B}_0$ for $r_1 = \frac{1.31}{4.9} a_0$, $\epsilon(x) = 2.1025$ for $r_1 \leqslant |x| \leqslant a_0/2$, and $\epsilon(x) = 0$ otherwise. The diameter a_0 of the cylinder can be chosen arbitrary. Reproduced with permission from [43]. Copyright © 2013 Society for Industrial and Applied Mathematics. All rights reserved.

non-degeneracy property, see definition 3.1 as well as the $\mathcal{PT}$ (parity-time reversal) symmetry defined below.

Definition 1.1. A function $A \in L^2(\mathbb{R}^n)$ is called $\mathcal{PT}$-symmetric if $A = \overline{A(-\cdot)}$.

Figure 2. (a) and (b) Components A_2 and A_3 of CMEs for the case $\omega_* = s_5$ in figure 1, where $N = 6$. A solution with the symmetry $A_1 = A_4 = 0$, $A_5 = A_2$, and $A_6 = A_3$ was chosen. (c) The approximation of the square modulus of a gap soliton at $\omega = \omega_* - \varepsilon^2$ with $\varepsilon = 0.1$ as given by $|u_{\mathrm{ans},1}|^2 + |u_{\mathrm{ans},2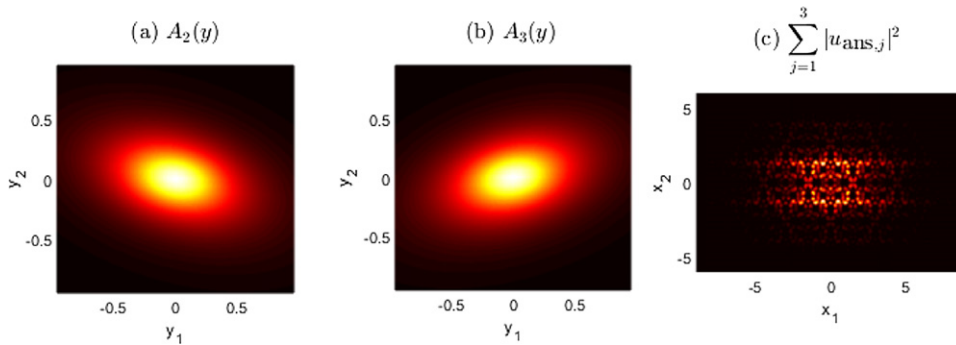}|^2 + |u_{\mathrm{ans},3}|^2$, see (1.8). Reproduced with permission from [43]. Copyright © 2013 Society for Industrial and Applied Mathematics. All rights reserved.

Theorem 1.1. *Let $\kappa \in \mathbb{R} \backslash \{0\}$. Suppose $A = (A_j)_{j=1}^N \in H^{s_A}(\mathbb{R}^2, \mathbb{C}^N)$ is a $\mathcal{PT}$-symmetric non-degenerate solution of the CMEs (1.10) with $s_A > 1$. Then, under assumptions (A1)–(A7), see section 3, there are constants $c > 0$, $\Omega \in \{-1, 1\}$, and $\varepsilon_0 > 0$ such that for each $\varepsilon \in (0, \varepsilon_0)$ there exists a $\mathcal{PT}$-symmetric solution $u \in H^2(\mathbb{R}^2, \mathbb{C})$ of the reduced Maxwell equation (1.6) with ω as in (1.7), which satisfies*

$$\|u - u_{ans}\|_{H^2(\mathbb{R}^2)} \leqslant c \, \varepsilon,$$

where u_{ans} is defined by (1.8).

Before immersing ourselves in the details of the proof, let us make some important remarks.

Remark 1. The assumption $\kappa \neq 0$ is needed in the Helmholtz decomposition in lemma A.1.

Remark 2. Note that $\|u_{\mathrm{ans}}\|_{L^2(\mathbb{R}^2)} = O(1)$ $(\varepsilon \to 0)$ since $\|A_j(\varepsilon \cdot)\|_{L^2(\mathbb{R}^2)} = \varepsilon^{-1} \|A_j\|_{L^2(\mathbb{R}^2)}$. Analogously one has $\|u_{\mathrm{ans}}\|_{H^2(\mathbb{R}^2)} = O(1)$. The next correction term in the asymptotics of the solution u is expected to have the form $\varepsilon^2 \sum_{j=1}^N A_j^{(2)}(\varepsilon x) r_j(x)$ with suitable (smooth) functions $A_j^{(2)}$ and r_j. It is the correction compensating for the residual at the formal order $O(\varepsilon^3)$ after satisfying the coupled mode equations (CMEs), see section 4.7. Hence, the expected correction term is $O(\varepsilon)$ in $H^2(\mathbb{R}^2)$, which implies the error estimate of theorem 1.1 is expected to be optimal.

Remark 3. Theorem 1.1 can also be considered as a result on the bifurcation of gap solitons from the zero solution at $\omega = \omega_*$. A sufficient condition is the existence of a $\mathcal{PT}$-symmetric and non-degenerate solution $A \in H^{s_A}(\mathbb{R}^2, \mathbb{C}^N)$ of the effective CME equations.

Remark 4. The CMEs in (1.10) are a system of coupled nonlinear Schrödinger equations and have the same structure as those for stationary gap solitons of the 2D scalar Gross–Pitaevskii equation with a periodic potential, see [10, 13]. In [43] several localized solutions of CMEs (1.10) with coefficients determined by Bloch waves of the Maxwell system were found numerically. The current paper does not discuss the existence of localized nontrivial solutions to CMEs. Existence results based on bifurcation theory and variational analysis can be found, e.g., in [28–30].

Remark 5. The $\mathcal{PT}$-symmetry has been extensively studied by the physics community in the recent years, mainly with emphasis on localized solutions, as it serves as a model for a balance between gain and loss in the structure. It has been shown to have a lot of applications, e.g. in Bose–Einstein condensates [25], non-Hermitian systems [5], quantum mechanics, optics [37], or surface plasmons polaritons [3, 32]. For a survey on the topic we refer to [19]. Mathematically, the restriction of a fixed point argument to a $\mathcal{PT}$-symmetric (or more generally anti-linearly symmetric) subspace has been used to obtain real nonlinear eigenvalues, see, e.g., [9, 11, 12, 36]. In particular, in our justification result, such symmetry, assumed on the functions ϵ and $\chi^{(3)}$ and then reflected in the band structure, is exploited to remove shift and space invariances in perturbed CMEs. This enables us to invert the linearized operator when working in the symmetric subspace.

Remark 6. The proof of theorem 1.1 is based on a generalized Lyapunov–Schmidt decomposition in Bloch variables and on fixed point arguments. The CMEs can be seen as the effective bifurcation system of the Lyapunov–Schmidt decomposition. This approach has been used, e.g., for wave packets of the Gross–Pitaevskii equation with periodic coefficients in [9, 10, 13, 14].

Remark 7. The term 'CMEs' in the context of asymptotics of wave packets is often used also for a different system, namely for time dependent first order envelope equations. These are derived when the wave packet is built using Bloch waves with nonzero group velocities, see [16, 21, 23, 39].

The rest of the paper is organized as follows. In section 2, after introducing the suitable functional setting, we investigate the linear problem $L^{(E)}u - \omega^2 \epsilon u = 0$, its spectrum and the Bloch waves. Then, using the Bloch transform, we formulate (1.6) in the Bloch variables. In addition, important regularity estimates on the Bloch eigenfunctions are also established here. Next, precise formulations of our assumptions are given in section 3. The proof of theorem 1.1 is provided in section 4 and split into several subsections according to our Lyapunov–Schmidt decomposition of the solution. We trim the solution u by rest terms, which are proved to be small enough in sections 4.3–4.6, and we finally show that the leading order part is ε-close to our ansatz in section 4.7. The appendix collects some auxiliary lemmas which are used in our analysis.

2. Function spaces, spectrum, Bloch transformation and linear estimates

In this section we firstly investigate the eigenvalue problem

$$\nabla' \times \nabla' \times u = \omega^2 \epsilon(x) u, \quad x \in \mathbb{R}^2 \tag{2.1}$$

and the corresponding Bloch eigenvalue problem on the periodicity cell

$$\nabla'_k \times \nabla'_k \times p = \omega_j^2(k) \epsilon(x) p, \quad x \in \mathcal{Q}, \tag{2.2}$$

where p is Λ-periodic, $k \in \mathbb{R}^2$ and (with $\kappa \in \mathbb{R}$ being a fixed parameter)

$$\nabla'_k := (\partial_1 + ik_1, \partial_2 + ik_2, i\kappa)^{\mathrm{T}}.$$

Secondly, we prove estimates on a linear inhomogeneous problem on the periodicity cell. This problem is obtained by applying the Bloch transformation to an inhomogeneous version of

(2.1), which plays a central role in a Banach fixed point iteration for the nonlinear equation in section 4.

The Bloch transformation and its properties are reviewed subsequently.

2.1. Function spaces

We start by defining some function spaces which we use below. Because of the presence of the curl operator in the Maxwell system (1.1) we will make use of H(curl) spaces with the curl defined using the above gradient ∇'. Let us first define

$$L^2_\#(\mathcal{Q}, \mathbb{C}^3) := \{v \in L^2_{\text{loc}}(\mathbb{R}^2, \mathbb{C}^3) : v \text{ is } \Lambda - \text{periodic}\},$$

$$H^s_\#(\mathcal{Q}, \mathbb{C}^3) := \{v \in H^s_{\text{loc}}(\mathbb{R}^2, \mathbb{C}^3) : v \text{ is } \Lambda - \text{periodic}\}, \quad s > 0.$$

The notation $L^2_\#(\mathcal{Q}, \mathbb{C}^3), H^s_\#(\mathcal{Q}, \mathbb{C}^3)$ is chosen to make clear that the elements need to be defined on the periodicity cell $\mathcal{Q}$ and periodically extendable in an L^2_{loc} resp. H^s_{loc} fashion onto $\mathbb{R}^2$. Note that for a vector field $u : \mathcal{Q} \to \mathbb{C}^3$ we define

$$\|u\|^2_{H^s(\mathcal{Q})} := \sum_{j=1}^3 \sum_{|\alpha| \leqslant s} \|D^\alpha u_j\|^2_{L^2(\mathcal{Q})}$$

with $\alpha \in \mathbb{N}_0^3$ being the standard multi-index and $s \in \mathbb{N}_0$.

Next we define

$$H_\#(\text{curl}, \mathcal{Q}) := \{v \in L^2_\#(\mathcal{Q}, \mathbb{C}^3) \,|\, \nabla' \times v \in L^2_\#(\mathcal{Q}, \mathbb{C}^3)\}$$

and

$$H_\#(\text{curl}^2, \mathcal{Q}) := \{v \in H_\#(\text{curl}, \mathcal{Q}) \,|\, \nabla' \times \nabla' \times v \in L^2_\#(\mathcal{Q}, \mathbb{C}^3)\}.$$

We will sometimes use the short notation $H_\#(\text{curl})$ or $H_\#(\text{curl}^2)$.

Note that in the majority of our calculations the gradient ∇' is replaced by ∇'_k. However, this makes no difference in the definition of the function spaces. Indeed, because

$$H_\#(\text{curl}, \mathcal{Q}) = \{v \in L^2_\#(\mathcal{Q}, \mathbb{C}^3) \,|\, \nabla'_k \times v \in L^2_\#(\mathcal{Q}, \mathbb{C}^3)\}$$

and

$$H_\#(\text{curl}^2, \mathcal{Q}) = \{v \in H_\#(\text{curl}, \mathcal{Q}) \,|\, \nabla'_k \times \nabla'_k \times v \in L^2_\#(\mathcal{Q}, \mathbb{C}^3)\}$$

for any $k \in \mathbb{R}^2$, we do not need to define new function spaces for problems involving the gradient ∇'_k.

For later use, we note the identity

$$\nabla'_k \cdot \nabla'_k \times u = 0 \quad \text{for all } u \in H_\#(\text{curl}, \mathcal{Q}),$$

which can be easily checked.

2.2. Spectral problem for the H-field

We build our linear theory on the results of [8] for the spectral problem for the H-field

$$\nabla' \times \left(\frac{1}{\epsilon} \nabla' \times v \right) = \omega^2 v.$$

It follows from the Bloch theory (see [18, 27]) that the spectrum of

$$L^{(H)} : H(\text{curl}_\epsilon^2) \to L^2(\mathbb{R}^2, \mathbb{C}^3), \qquad L^{(H)} := \nabla' \times \left(\frac{1}{\epsilon} \nabla' \times \cdot \right),$$

where

$$H(\text{curl}_\epsilon^2) := \left\{ v \in L^2(\mathbb{R}^2, \mathbb{C}^3) \,|\, \nabla' \times v \in L^2(\mathbb{R}^2, \mathbb{C}^3), \nabla' \times \left(\frac{1}{\epsilon} \nabla' \times v \right) \in L^2(\mathbb{R}^2, \mathbb{C}^3) \right\},$$

is obtained as the union (over all $k \in \mathbb{B}$) of the spectra of

$$L_k^{(H)} : H_\#(\text{curl}_\epsilon^2) \to L_\#^2(\mathcal{Q}, \mathbb{C}^3), \qquad L_k^{(H)} := \nabla'_k \times \left(\frac{1}{\epsilon} \nabla'_k \times \cdot \right),$$

where

$$H_\#(\text{curl}_\epsilon^2) := \left\{ v \in L_\#^2(\mathcal{Q}, \mathbb{C}^3) \,|\, \nabla'_k \times v \in L_\#^2(\mathcal{Q}, \mathbb{C}^3), \nabla'_k \times \left(\frac{1}{\epsilon} \nabla'_k \times v \right) \in L_\#^2(\mathcal{Q}, \mathbb{C}^3) \right\}.$$

We emphasize that $L_k^{(H)}$ acts on periodic functions on the periodicity cell $\mathcal{Q}$.

Let $k \in \mathbb{B}$ be fixed. With the form domain of $L_k^{(H)}$ being

$$V_k := \{ v \in L_\#^2(\mathcal{Q}, \mathbb{C}^3) \,|\, \nabla'_k \times v \in L_\#^2(\mathcal{Q}, \mathbb{C}^3), \ \nabla'_k \cdot v = 0 \}, \tag{2.3}$$

the authors of [8] prove that the spectrum is discrete and satisfies

$$\sigma(L_k^{(H)}) = \{ \omega_1^2(k), \omega_2^2(k), \dots \} \subset [0, \infty),$$

where

$$\omega_1^2(k) \leqslant \omega_2^2(k) \leqslant \dots$$

The corresponding eigenfunctions $(q_j(\cdot, k))_{j \in \mathbb{N}} \subset V_k$ satisfy

$$a_k(q_j(\cdot, k), \varphi) = \omega_j^2(k) \langle q_j(\cdot, k), \varphi \rangle \quad \forall \, \varphi \in V_k, \tag{2.4}$$

where

$$a_k(\psi, \varphi) := \int_{\mathcal{Q}} \frac{1}{\epsilon} \nabla'_k \times \psi \cdot \overline{\nabla'_k \times \varphi} \, dx \qquad \text{and}$$

$$\langle \psi, \varphi \rangle := (\psi, \varphi)_{L^2(\mathcal{Q})} := \int_{\mathcal{Q}} \psi \cdot \overline{\varphi} \, dx.$$

Moreover, they can be chosen $L^2(\mathcal{Q})$-orthonormal, i.e.

$$\langle q_i(\cdot, k), q_j(\cdot, k) \rangle = \delta_{ij} \quad \forall \, i, j \in \mathbb{N}.$$

It follows that for λ in the resolvent set, i.e. $\lambda \in \mathbb{C} \backslash \{ \omega_1^2(k), \omega_2^2(k), \dots \}$, and $g \in L_\#^2(\mathcal{Q}, \mathbb{C}^3)$ there is a unique $v \in V_k$ such that

$$a_k(v, \varphi) - \lambda \langle v, \varphi \rangle = \langle g, \varphi \rangle \quad \forall \, \varphi \in V_k. \tag{2.5}$$

Moreover, there is a constant $c > 0$ such that

$$\|v\|_{H^1(\mathcal{Q})} \leqslant c \, \|g\|_{L^2(\mathcal{Q})}. \tag{2.6}$$

Here the equivalence of the $H(\mathrm{curl})$ and H^1-norms on V_k has been used. In fact, the eigenfunctions automatically satisfy the regularity

$$q_j(\cdot, k) \in H_\#(\mathrm{curl}_\epsilon^2) \cap V_k. \tag{2.7}$$

To show this, it suffices to prove that (2.4) holds for all $\varphi \in C_c^\infty(\mathcal{Q})$. Then the weak curl $\nabla_k' \times$ of $\frac{1}{\epsilon}\nabla_k' \times q_j(\cdot, k)$ equals $\omega_j^2(k)q_j(\cdot, k)$, which is in $L_\#^2(\mathcal{Q}, \mathbb{C}^3)$. Due to the Helmholtz decomposition in lemma A.1 we have

$$H_\#(\mathrm{curl}, \mathcal{Q}) = V_k \oplus \nabla_k' H_\#^1(\mathcal{Q}),$$

so, by a density argument, it remains to show that (2.4) holds for all $\varphi \in \nabla_k' C_\#^\infty(\overline{\mathcal{Q}})$, where as usual the subscript $\#$ denotes periodicity. Substituting $\varphi = \nabla_k' \psi$ with $\psi \in C_\#^\infty(\overline{\mathcal{Q}})$, we clearly have $\nabla_k' \times \nabla_k' \psi = 0$, as well as

$$\int_\mathcal{Q} q_j(x, k) \cdot \overline{\nabla_k'\psi}(x)\mathrm{d}x = \int_{\partial\mathcal{Q}} q_j(x, k)\overline{\psi}(x) \cdot \nu\,\mathrm{d}S(x) - \int_\mathcal{Q} \nabla_k' \cdot q_j(x, k)\overline{\psi}(x)\mathrm{d}x = 0,$$

where $\nu := (\nu_1, \nu_2, 0)^\mathsf{T}$ and $(\nu_1, \nu_2)^\mathsf{T}$ being a.e. defined as the unit outer normal vector of $\partial\mathcal{Q}$. Indeed, first $\nabla_k' \cdot q_j(\cdot, k) = 0$ because $q_j(\cdot, k) \in V_k$. Second, as $V_k \subset H_{\mathrm{loc}}^1(\mathbb{R}^2)$, the boundary term is well-defined and, by periodicity of both $q_j(\cdot, k)$ and ψ, the contributions of the boundary integral on opposite sides of the periodicity cell $\mathcal{Q}$ cancel out.

Due to the regularity in (2.7) we conclude

$$\nabla_k' \times \left(\frac{1}{\epsilon}\nabla_k' \times q_j(\cdot, k)\right) = \omega_j^2(k)q_j(\cdot, k) \quad \text{in } L_\#^2(\mathcal{Q}, \mathbb{C}^3) \text{ for each } j \in \mathbb{N}. \tag{2.8}$$

For the spectrum of the operator $L^{(H)}$ in $L^2(\mathbb{R}^2)$ one has

$$\sigma(L^{(H)}) = \bigcup_{k\in\mathbb{B}}\sigma(L_k^{(H)}) = \bigcup_{k\in\mathbb{B},\, n\in\mathbb{N}}\omega_n(k),$$

see [18].

2.3. Spectral problem for the E−field

Let $k \in \mathbb{B}$ be fixed. As we show now, for each eigenfunction q_j of (2.4) the function (for $\omega_j(k) \neq 0$)

$$p_j(x, k) := \frac{\mathrm{i}}{\epsilon(x)\omega_j(k)}\nabla_k' \times q_j(x, k) \tag{2.9}$$

is an $H_\#(\mathrm{curl}^2)$ eigenfunction of the eigenvalue problem for the E-field. Due to (2.7) we first have $p_j(\cdot, k) \in H_\#(\mathrm{curl})$. Next, (2.8) implies $\nabla_k' \times p_j(\cdot, k) = \mathrm{i}\omega_j(k)q_j(\cdot, k) \in H_\#(\mathrm{curl})$, hence $p_j(\cdot, k) \in H_\#(\mathrm{curl}^2)$ and

$$L_k^{(E)}p_j(\cdot, k) := \nabla_k' \times \nabla_k' \times p_j(\cdot, k) = \epsilon\omega_j^2(k)p_j(\cdot, k) \quad \text{in } L_\#^2(\mathcal{Q}). \tag{2.10}$$

The sequence $(p_j(\cdot, k))_j$ satisfies the orthogonality

$$\langle p_i(\cdot, k), p_j(\cdot, k)\rangle_\epsilon := \langle p_i(\cdot, k), \epsilon p_j(\cdot, k)\rangle = \delta_{ij} \quad \forall\, i, j \in \mathbb{N} \tag{2.11}$$

because

$$\langle p_i(\cdot,k), p_j(\cdot,k)\rangle_\epsilon = \frac{1}{\omega_i(k)\omega_j(k)} \int_Q \frac{1}{\epsilon} \nabla'_k \times q_i(x,k) \cdot \overline{\nabla'_k \times q_j(x,k)} \mathrm{d}x$$
$$= \frac{1}{\omega_i(k)\omega_j(k)} a_k(q_i(\cdot,k), q_j(\cdot,k)) = \delta_{ij}.$$

Besides the periodicity in x the functions $p_n, n \in \mathbb{R}$, are quasiperiodic in k, namely

$$p_n(x, k+K) = p_n(x,k)\mathrm{e}^{-\mathrm{i}K\cdot x} \quad \text{for all } x \in \mathbb{R}^2 \text{ and } K \in \Lambda^*.$$

Two symmetries of the eigenfunctions p_n will be used in the analysis. Firstly, because the eigenvalue problem is invariant under the complex conjugation combined with replacing k by $-k$, one sees that $\omega_n(k) = \omega_n(-k)$ for all $n \in \mathbb{N}$ since they are real. Similarly, one also deduces that $\overline{p_n(x,k)}$ is an eigenfunction of $L^{(E)}_{-k}$ if and only if $p_n(x,k)$ is an eigenfunction of $L^{(E)}_k$. This implies that the eigenfunction $p_n(x,-k)$ can be chosen to agree with $\overline{p_n(x,k)}$ for all $k \in \mathbb{R}^2 \backslash \{0\}$. Notice that at $k=0$ the operator $L^{(E)}_0$ is real, so a real eigenfunction can always be chosen. Hence we have

$$p_n(x,-k) = \overline{p_n(x,k)} \quad \text{for all } x,k \in \mathbb{R}^2, \ n \in \mathbb{N}. \tag{2.12}$$

Secondly, if $\epsilon(x) = \epsilon(-x)$ and if $p_n(x,k)$ is an eigenfunction of $L^{(E)}_k$, it is easy to show that $\overline{p_n(-x,k)}$ is an eigenfunction of $L^{(E)}_k$, too. Therefore, if $\omega_n(k)^2$ is a geometrically simple eigenvalue of (2.4), there is always a choice of the phase of the normalized eigenfunction p_n such that the $\mathcal{PT}-$symmetry

$$p_n(-x,k) = \overline{p_n(x,k)} \quad \text{for all } x \in \mathbb{R}^2 \tag{2.13}$$

holds.

The map $k \mapsto \omega_n(k)$ with $\omega_n \geqslant 0$ is called the *nth eigenvalue* and the map $(k,n) \mapsto \omega_n(k)$ the *band structure*. Clearly, since the spectrum (for each k) is given by $\{\omega_1(k)^2, \omega_2(k)^2, \ldots\}$, there are also the negative eigenvalues $\omega_{-n} := -\omega_n$, but they play no role in our analysis. Notice also that the band structure is the same for both operators $L^{(H)}$ and $L^{(E)}$.

2.4. Inhomogeneous linear equation for the $E-$field

Our asymptotic and nonlinear analysis is performed for the $E-$ field and in the fixed point argument we need to solve the inhomogeneous problem

$$L^{(E)}_k u - \omega^2 \epsilon u = f \tag{2.14}$$

with ω^2 in the resolvent set of $\epsilon^{-1} L^{(E)}_k$, i.e. of $L^{(H)}_k$. In our application we have $f \in H^2_\#(Q)$ and the fixed point argument requires the estimate $\|u\|_{H^2(Q)} \leqslant c\|f\|_{H^2(Q)}$. We prove this estimate next.

Lemma 2.1. *Let* $k \in \mathbb{B}$, $\epsilon \in W^{2,\infty}(Q)$, $\epsilon^{-1} \in L^\infty(Q)$, $\omega^2 \in \mathbb{C}\backslash\{\omega_1^2(k), \omega_2^2(k), \ldots\}$, *and* $f \in H^2_\#(Q)$. *Then* (2.14) *has a unique solution* $u \in H_\#(\mathrm{curl}^2)$ *such that*

$$\|u\|_{H^2(Q)} \leqslant c\|f\|_{H^2(Q)} \tag{2.15}$$

holds.

Remark 8. Note that (2.15) is clearly not optimal as an estimate of the solution of (2.14). An optimal estimate includes just the L^2-norm on the right-hand side. However, as our nonlinear

analysis below employs an estimate of the form $\|u\|_{H^2} \leqslant c\|f\|_{H^2}$, this suboptimality is not of essence.

Proof. For $\omega^2 \in \mathbb{C} \setminus \{\omega_1^2(k), \omega_2^2(k), \dots\}$ we choose ω such that e.g. $\arg \omega \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right]$. We define first

$$\tilde{f} := -\frac{\mathrm{i}}{\omega} \nabla_k' \times \frac{f}{\epsilon}$$

and solve $L_k^{(H)} v - \omega^2 v = \tilde{f}$ in the weak sense, see (2.5). Due to (2.6) we get $\|v\|_{H^1(\mathcal{Q})} \leqslant c \|\tilde{f}\|_{L^2(\mathcal{Q})}$. Since $\epsilon \in W^{1,\infty}(\mathcal{Q})$ and $\epsilon^{-1} \in L^\infty(\mathcal{Q})$, we get

$$\|v\|_{H^1(\mathcal{Q})} \leqslant c \, \|f\|_{H(\mathrm{curl})}. \tag{2.16}$$

Moreover, similarly to (2.7), using the Helmholtz decomposition of lemma A.1, we get $v \in H_\#(\mathrm{curl}_\epsilon^2)$ and

$$L_k^{(H)} v - \omega^2 v = \tilde{f} \quad \text{in } L_\#^2(\mathcal{Q}, \mathbb{C}^3). \tag{2.17}$$

Next, we set

$$u := \frac{\mathrm{i}}{\omega \epsilon} \nabla_k' \times v - \frac{1}{\omega^2 \epsilon} f. \tag{2.18}$$

Then $\nabla_k' \times u = \mathrm{i}\omega v$ and using (2.16) as well as the assumptions on ϵ, we have

$$\|u\|_{H(\mathrm{curl})} \leqslant c\|f\|_{H(\mathrm{curl})}. \tag{2.19}$$

Moreover, since $v \in H_\#(\mathrm{curl})$, we have from $\nabla_k' \times u = \mathrm{i}\omega v$ also $u \in H_\#(\mathrm{curl}^2)$ and thus, applying $\nabla_k' \times$ to (2.18), we obtain that (2.14) holds as an equation in $L_\#^2(\mathcal{Q})$.

Next, we derive the desired H^2-estimate on u. We start with H^1. Because

$$\|u\|_{H^1(\mathcal{Q})} \leqslant c \left(\|u\|_{H(\mathrm{curl})} + \|\nabla_k' \cdot u\|_{L^2(\mathcal{Q})} \right) \tag{2.20}$$

and because of (2.19) it remains to estimate the divergence. Since $\nabla_k' \cdot u = \frac{1}{\epsilon}(\nabla_k' \cdot (\epsilon u) - (\nabla_k' \epsilon) \cdot u)$, from (2.18) we infer

$$\nabla_k' \cdot u = -\frac{1}{\epsilon} \left(\frac{1}{\omega} \nabla_k' \cdot f - (\nabla_k' \epsilon) \cdot u \right). \tag{2.21}$$

By $\epsilon \in W^{1,\infty}$ and $\epsilon^{-1} \in L^\infty$, we get then

$$\|\nabla_k' \cdot u\|_{L^2(\mathcal{Q})} \leqslant c \left(\|f\|_{H^1(\mathcal{Q})} + \|u\|_{L^2(\mathcal{Q})} \right) \leqslant c \, \|f\|_{H^1(\mathcal{Q})}, \tag{2.22}$$

where the last inequality holds by (2.19). Next,

$$\|u\|_{H^2(\mathcal{Q})} \leqslant c \left(\|u\|_{L^2(\mathcal{Q})} + \|\nabla_k' \times u\|_{H^1(\mathcal{Q})} + \|\nabla_k' \cdot u\|_{H^1(\mathcal{Q})} \right)$$

and, using again (2.18) and (2.21), and $\epsilon \in W^{2,\infty}$ we obtain

$$\|u\|_{H^2(\mathcal{Q})} \leqslant c \left(\|u\|_{H^1(\mathcal{Q})} + \left\| \nabla_k' \times \left(\frac{1}{\epsilon} \nabla_k' \times v \right) \right\|_{H^1(\mathcal{Q})} + \|\nabla_k' \times f\|_{H^1(\mathcal{Q})} + \|\nabla_k' \cdot f\|_{H^1(\mathcal{Q})} \right).$$

The estimate (2.15) is finally deduced by (2.16), (2.17) and (2.19), (2.20), (2.22). $\qquad\square$

2.5. Bloch transformation

To take advantage of the fact that the coefficients of our problem (1.6) are periodic, we will work in Bloch variables, i.e. we will employ the Bloch transform to change the problem into a family of problems on the periodicity cell $\mathcal{Q}$, parametrized by the wave vector $k \in \mathbb{B}$. The above discussion (sections 2.2 and 2.3) implies that the resulting equation has (for each k) a linear operator with a discrete spectrum.

The *Bloch transform* $\mathcal{T} : L^2(\mathbb{R}^2) \to L^2(\mathbb{B}, L^2_\#(\mathcal{Q}))$ so that $v \mapsto \widetilde{v}$ and its inverse are formally defined as

$$\widetilde{v}(x, k) = (\mathcal{T}v)(x, k) := \sum_{K \in \Lambda^*} \widehat{v}(k + K)\mathrm{e}^{\mathrm{i}K \cdot x},$$

$$v(x) = (\mathcal{T}^{-1}\widetilde{v})(x) = \int_{\mathbb{B}} \widetilde{v}(x, k)\mathrm{e}^{\mathrm{i}k \cdot x}\, \mathrm{d}k$$

for all $x, k \in \mathbb{R}^2$, see e.g. [38] or [2, chapter 7]. For the domain and range of $\mathcal{T}$ see (2.26). Here $\widehat{v}$ denotes the *Fourier transform* of $v \in L^1(\mathbb{R}^2)$

$$\widehat{v}(k) := \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} v(x)\mathrm{e}^{-\mathrm{i}k \cdot x}\, \mathrm{d}x,$$

which is extended to $L^2(\mathbb{R}^2)$ functions as usual. We also define $\widehat{v}(x) := \int_{\mathbb{R}^2} v(k)e^{\mathrm{i}k \cdot x}\mathrm{d}k$ which can be shown to be the inverse Fourier transform in $L^2(\mathbb{R}^2)$.

The definition of $\tilde{v}$ yields naturally the periodicity in x and the quasi-periodicity in k, i.e.

$$
\begin{aligned}
\widetilde{v}(x + R, k) &= \widetilde{v}(x, k) && \text{for all } R \in \Lambda, \ x \in \mathbb{R}^2, \ k \in \mathbb{R}^2, \\
\widetilde{v}(x, k + K) &= \mathrm{e}^{-\mathrm{i}K \cdot x}\widetilde{v}(x, k) && \text{for all } K \in \Lambda^*, \ x \in \mathbb{R}^2, \ k \in \mathbb{R}^2.
\end{aligned}
\tag{2.23}
$$

Moreover, the product of two functions $f, g \in L^2(\mathbb{R}^2)$ for which also $fg \in L^2(\mathbb{R}^2)$ is transformed by $\mathcal{T}$ into a convolution of the transformed functions:

$$(\mathcal{T}(fg))(x, k) = \int_{\mathbb{B}} \tilde{f}(x, k - l)\tilde{g}(x, l)\, \mathrm{d}l =: (\tilde{f} *_{\mathbb{B}} \tilde{g})(x, k), \tag{2.24}$$

where the quasiperiodicity property in (2.23) is used if $k - l \notin \mathbb{B}$. For the same reason the convolution in $\mathbb{B}$ can be substituted by a convolution on any shifted Brillouin zone, i.e.

$$(\tilde{f} *_{\mathbb{B}} \tilde{g})(x, k) = \int_{\mathbb{B} + k_*} \tilde{f}(x, k - l)\tilde{g}(x, l)\, \mathrm{d}l \quad \text{for all } k_* \in \mathbb{C}^2.$$

If $h \in L^\infty(\mathbb{R}^2)$ enjoys periodicity with respect to the same lattice Λ and $g \in L^2(\mathbb{R}^2)$, then

$$(\mathcal{T}(hg))(x, k) = h(x)(\mathcal{T}g)(x, k) \tag{2.25}$$

for almost all $x \in \mathbb{R}^2$ and $k \in \mathbb{B}$.

The function spaces for the Bloch transform. Let $H^s(\mathbb{R}^n)$ with $s > 0$ be the stankdard (possibly fractional) Sobolev space. The Bloch transform

$$\mathcal{T} : H^s(\mathbb{R}^n, \mathbb{C}) \to \mathcal{X}_s := L^2(\mathbb{B}, H^s_\#(\mathcal{Q}, \mathbb{C})) \tag{2.26}$$

is an isomorphism for $s \geqslant 0$ [35, 38]. The norm in $\mathcal{X}_s$ is defined as

$$\|\widetilde{u}\|_{\mathcal{X}_s} = \left(\int_{\mathbb{B}} \|\widetilde{u}(\cdot, k)\|^2_{H^s(\mathcal{Q})} \mathrm{d}k \right)^{\frac{1}{2}},$$

where $\mathcal{Q}$ is an arbitrary interval in $\mathbb{R}^n$ and $\mathbb{B}$ the corresponding reciprocal periodicity cell. For our application we have $n = 2$ with $\mathcal{Q}$ and $\mathbb{B}$ as defined in section 1. For vector-valued functions $u \in H^s(\mathbb{R}^n, \mathbb{C}^m)$, $m \in \mathbb{R}$, the transform $\mathcal{T}$ is defined componentwise and the space $\mathcal{X}_s$ is $\mathcal{X}_s := L^2(\mathbb{B}, H^s_\#(\mathcal{Q}, \mathbb{C}^m))$ with the norm $\|f\|_{\mathcal{X}_s} := \max_{1 \leqslant j \leqslant m} \|f_j\|_{\mathcal{X}_s}$.

Note that due to the quasi-periodicity of $\tilde{u}$ in k, the $\mathcal{X}_s$-norm is equivalent to

$$\left(\int_{\mathbb{B}+k_*} \|\widetilde{u}(\cdot, k)\|^2_{H^s(\mathcal{Q})} \right)^{\frac{1}{2}}$$

for any $k_* \in \mathbb{R}^2$. We take advantage of this property in our estimates below.

Because of the polynomial nonlinearity in (1.6) and our approach employing a fixed point argument we require our function space to have the algebra property with respect to the point-wise multiplication. We recall that if $s > n/2$, the Sobolev space $H^s(\mathbb{R}^n)$ enjoys this property. Moreover, it embeds into the space of bounded and continuous functions decaying to 0 at ∞.

In the Bloch variables, where multiplication is transformed into a convolution, we need the algebra property with respect to the convolution. Combining the algebra property of $H^s(\mathbb{R}^n), s > n/2$, and (2.24), we get the following algebra property for our working space $\mathcal{X}_s$:

$$\|\widetilde{f} *_\mathbb{B} \widetilde{g}\|_{\mathcal{X}_s} \leqslant c \, \|\widetilde{f}\|_{\mathcal{X}_s} \|\widetilde{g}\|_{\mathcal{X}_s} \quad \text{for any } \widetilde{f}, \widetilde{g} \in \mathcal{X}_s \text{ if } s > n/2. \tag{2.27}$$

We introduce also the weighted spaces $L^2_s(\mathbb{R}^n)$ defined as

$$L^2_s(\mathbb{R}^n) := \left\{ f \in L^2(\mathbb{R}^n) \mid \|f\|^2_{L^2_s(\mathbb{R}^n)} := \int_{\mathbb{R}^n} (1 + |x|)^{2s} |f(x)|^2 \, \mathrm{d}x < \infty \right\}. \tag{2.28}$$

Recall that the Fourier transform is an isomorphism from $H^s(\mathbb{R}^n)$ to $L^2_s(\mathbb{R}^n)$ for $s \geqslant 0$.

3. Assumptions

We start with the following basic assumptions on the coefficients and on the band structure.

(A1) $\epsilon : \mathbb{R}^2 \to \mathbb{R}$ and $\chi^{(3)} : \mathbb{R}^2 \to \mathbb{R}^{3 \times 3 \times 3 \times 3}$ are Λ-periodic and real-valued and $\epsilon > 0$;

(A2) The spectrum $\cup_{n \in \mathbb{N}, k \in \mathbb{B}} \{\omega_n(k)\} \subset \mathbb{R}$ possesses a gap;

(A3) The points $k^{(1)}, \ldots, k^{(N)} \in \mathbb{B}$ are distinct and constitute the level set $W_{\omega_*} \subset \mathbb{B}$ of one of the gap edges, denoted by ω_*, and the eigenvalues at the level ω_* are all geometrically simple, i.e.

$$\omega_* = \omega_n(k), \ n \in \mathbb{N}, \ k \in \mathbb{B} \Rightarrow k \in \{k^{(1)}, \ldots, k^{(N)}\},$$

and

$$\dim \ \ker \left(L^{(E)}_{k^{(j)}} - \epsilon \omega_*^2 I \right) = 1.$$

Hence, due to the monotonicity $\omega_n(k) \leqslant \omega_{n+1}(k)$, we have

$$\exists n_* \in \mathbb{N} : \ \omega_{n_*}(k^{(j)}) = \omega_* \quad \forall \ j \in \{1, \ldots, N\}.$$

(A4) The eigenvalue ω_{n_*} is twice continuously differentiable at $k^{(j)}$ and $\nabla^2 \omega_{n_*}(k^{(j)})$, the Hessian of ω_{n_*} at $k = k^{(j)}$, is definite for each $j \in \{1, \ldots, N\}$.

The formal asymptotic analysis of gap solitons in [43] used assumptions (A1), (A2), and (A4). In assumption (A3) multiple eigenvalues were allowed at the points $k^{(j)}, j = 1, \ldots, N$. Here, in order to be able to prove symmetries of the Bloch waves at $k^{(j)}, j = 1, \ldots, N$, which are needed in the restriction of the nonlinear problem to a symmetric subspace, we require

the geometric simpleness. The unique (up to a phase factor $e^{i\alpha}$) normalized eigenfunction at $k = k^{(j)}$ and $\omega = \omega_*$ is denoted by $p_{n_*}(\cdot, k^{(j)})$.

Note that, according to the mathematical folklore, simple eigenvalues depend smoothly on the coefficients of the operator. Nevertheless, we are not aware of an existing result applicable to our operator $L_k^{(H)}$ or $L_k^{(E)}$, hence we assume the C^2-regularity in (A4). In addition our proof requires the Lipschitz continuity throughout $\mathbb{B}$, which we prove in the appendix, see lemma A.2.

Clearly, ω_* must be the maximum or minimum of the eigenvalue ω_{n_*}. Hence, based on (A4), $\nabla^2 \omega_{n_*}(k^{(j)})$ is either positive definite for all j or negative definite for all j. Note that the assumption that ω_* is attained at the points $k = k^{(j)}$, $j = 1, \ldots, N$ by the same eigenvalue ω_{n_*} is in accordance with the numbering of the eigenvalues $\omega_n(k)$ at each k according to the magnitude.

Remark 9. Assumption (A3) seems relatively restrictive as it does not allow for $\{k^{(1)}, \ldots, k^{(N)}\}$ to be a proper subset of the level set W_{ω_*}. We need this assumption to estimate the correction term, which is supported (in the wave-number k) away from small neighbourhoods of the points $\{k^{(1)}, \ldots, k^{(N)}\}$, i.e. away from the support of the main contribution of the solution. The support of the correction term must not intersect W_{ω_*} because otherwise $(\omega_{n_*}(k) - \omega_*)^{-1}$ blows up on this support. Note that this can be contrasted against the case of the bifurcation of nonlinear Bloch waves in [15], where $\{k^{(1)}, \ldots, k^{(N)}\}$ can be a proper subset of W_{ω_*} provided the k-points generated by (iterations of) the nonlinearity, i.e. the points

$$k \in S_3^n(\{k^{(1)}, \ldots, k^{(N)}\}) \quad \text{for some } n \in \mathbb{N},$$

where

$$S_3(\{k^{(1)}, \ldots, k^{(N)}\}) := \{k^{(\alpha)} + k^{(\beta)} - k^{(\gamma)} : \alpha, \beta, \gamma \in \{1, \ldots, N\}\},$$

lie outside the level set. Unlike in [15] the k-support of the leading order term of the gap solitons contains whole neighbourhoods of the points $\{k^{(1)}, \ldots, k^{(N)}\}$ (and not isolated points) implying iterations of S_3 applied to the union of these neighbourhoods generate all $k \in \mathbb{B}$.

It is however possible for some components of the CME-solutions to be zero, i.e. $A_{m_1} = \cdots = A_{m_M} = 0$ for some $1 \leqslant m_1, \ldots, m_M \leqslant N$ (with $M < N$). This can happen only if the CMEs are consistent with the reduction to the components $A_k, k \in \{1, \ldots, N\} \backslash \{m_1, \ldots, m_M\}$ or equivalently if

$$S_3(\{k^{(m_1)}, \ldots, k^{(m_M)}\}) \cap \{k^{(1)}, \ldots, k^{(N)}\} = \{k^{(m_1)}, \ldots, k^{(m_M)}\}.$$

In that sense assumption (A3) is effectively the same as assuming that $\{k^{(1)}, \ldots, k^{(N)}\}$ is a consistent subset (in the above sense) of the level set W_{ω_*} and that W_{ω_*} is finite.

To rigorously justify the formal approximation via (1.8), we need to assume the following additional conditions:

(A5) the material functions ϵ and $\chi^{(3)}$ satisfy $\epsilon \in W^{2,\infty}(\mathbb{R}^2)$, $\epsilon^{-1} \in L^\infty(\mathbb{R}^2)$, $\chi^{(3)} \in H_{\mathrm{loc}}^2(\mathbb{R}^2)$;
(A6) symmetry of the material: $\epsilon(x) = \epsilon(-x)$, $\chi^{(3)}(x) = \chi^{(3)}(-x)$ for all $x \in \mathbb{R}^2$;
(A7) the eigenvalue $\omega_{n_*}(k)$ is geometrically simple for almost all (w.r.t. the Lebesgue measure) $k \in \mathbb{B}$.

Note that assumption (A7) allows for the touching of eigenvalue graphs $(k, \omega_{n_*}(k))$ and $(k, \omega_m(k))$ with $m \neq n_*$ as long as they touch along a curve; which is the canonical situation. This curve may include the points $k^{(1)}, \ldots, k^{(N)}$, see assumption (A3).

Under the above assumptions theorem 1.1 justifies the use of the effective amplitude equation (1.10) to determine the envelopes A_j in the ansatz (1.8) and constitutes the main result of the paper. It uses the following definition.

Definition 3.1. A solution $A_* \in L^2(\mathbb{R}^2)^N$ of (1.10), denoted by $\mathcal{G}(A) = 0$, is called *non-degenerate* if the kernel of the Jacobian of $\mathcal{G}$ evaluated at A_* is only three dimensional as generated by the two spatial shift invariances and the complex phase invariance of the CMEs, i.e.

$$\ker \begin{pmatrix} \partial_{A_R}\mathcal{G}_R(A_*) & \partial_{A_I}\mathcal{G}_R(A_*) \\ \partial_{A_R}\mathcal{G}_I(A_*) & \partial_{A_I}\mathcal{G}_I(A_*) \end{pmatrix} = \mathrm{span}\left\{ \partial_{y_1}\begin{pmatrix} A_{*,R} \\ A_{*,I} \end{pmatrix}, \partial_{y_2}\begin{pmatrix} A_{*,R} \\ A_{*,I} \end{pmatrix}, \begin{pmatrix} -A_{*,I} \\ A_{*,R} \end{pmatrix} \right\},$$

where $A_{*,R} := \mathrm{Re}(A_*), A_{*,I} := \mathrm{Im}(A_*)$ and analogously for the other variables and functions.

Assumptions (A.6) and (A.7) are used to remove invariances (and thus eliminate non-trivial elements of the kernel) in a perturbed CME-problem by restricting to a symmetric subspace. This perturbed system is obtained in the justification analysis. The symmetric subspace is defined by the $\mathcal{PT}$-symmetry, see definition 1.1. In this subspace the CMEs no longer possess the invariances wrt. the spatial shift and the complex phase. Hence, under the non-degeneracy condition, the linearized operator of the perturbed CME system is invertible. Note that other symmetric subspaces can be used to eliminate the kernel, see [13].

Moreover, the evenness of ϵ and $\chi^{(3)}$ implies that the coefficients $I^j_{\alpha,\beta,\gamma}$ are real as explained at the end of section 4.6.1.

4. Proof of theorem 1.1

From now on, the bifurcation parameter ω is chosen to lie in the spectral gap in a $\mathcal{O}(\varepsilon^2)$-vicinity of the edge ω_*, i.e.

$$\omega = \omega_* + \varepsilon^2\Omega, \tag{4.1}$$

where $\Omega = \pm 1$, the sign being determined by the condition that ω shall lie in the gap. Hence, $\Omega = \pm 1$ if ω_* is the bottom/top edge of a spectral gap, respectively.

4.1. Lyapunov–Schmidt decomposition

We study the problem in Bloch variables in the space $\mathcal{X}_2$. Applying the Bloch transform $\mathcal{T}$ to (1.6), we get

$$L_k\widetilde{u}(x,k) := \nabla'_k \times \nabla'_k \times \widetilde{u}(x,k) - \omega^2\epsilon(x)\widetilde{u}(x,k) = \omega^2\widetilde{F}(\widetilde{u})(x,k), \tag{4.2}$$

where

$$\widetilde{F}_d(\widetilde{u}) = \sum_{a,b,c=1}^3 \underline{\chi}^{(3)}_{a,b,c,d}\widetilde{u}_a *_{\mathbb{B}} \widetilde{u}_b *_{\mathbb{B}} \widetilde{\overline{u}}_c. \tag{4.3}$$

Here properties (2.24) and (2.25) have been used. Recall that ω is fixed as in (4.1). Note that in (4.3) the double convolution equals $(\widetilde{u}_a *_{\mathbb{B}} \widetilde{u}_b) *_{\mathbb{B}} \widetilde{\overline{u}}_c = \widetilde{u}_a *_{\mathbb{B}} (\widetilde{u}_b *_{\mathbb{B}} \widetilde{\overline{u}}_c)$.

Note also that below $f *_\mathbb{B} g$ is understood componentwise for scalar f and vector valued g.

Theorem 1.1 claims that the solution can be approximated by a modulated sum of the Bloch eigenfunction $p_{n_*}(\cdot, k)$ at the chosen points $k^{(1)}, \ldots, k^{(N)}$. Therefore, we decompose $\widetilde{u}$ into the part $\widetilde{v}$ corresponding to the eigenfunction p_{n_*} and the rest $\widetilde{w}$. Next, $\widetilde{v}$ is once more split into a first term which incorporates the behaviour in the vicinity of the points $k^{(1)}, \ldots, k^{(N)}$ and a rest. To this end, we first introduce some projections on $L^2(\mathcal{Q})$ which take into account the presence of the potential $\epsilon(\cdot)$.

Projections. Let P_k denote the standard $L^2(\mathcal{Q})$-projection onto the mode $p_{n_*}(\cdot, k)$, i.e. for $f \in L^2(\mathcal{Q})$

$$(P_k f)(\cdot, k) := \langle f, p_{n_*}(\cdot, k) \rangle \, p_{n_*}(\cdot, k),$$

and let $Q_k := I - P_k$ be its L^2-orthogonal projection. As the normalization of the mode $p_{n_*}(\cdot, k)$ holds in the L^2-norm weighted by the periodic potential $\epsilon(\cdot)$ (see (2.11)), we also introduce

$$(P_k^\epsilon f)(\cdot, k) := \langle f, \epsilon(\cdot) p_{n_*}(\cdot, k) \rangle \, p_{n_*}(\cdot, k) \quad \text{and} \quad Q_k^\epsilon := I - P_k^\epsilon$$

as well as

$$^\epsilon P_k := \epsilon(\cdot) P_k \quad \text{and} \quad ^\epsilon Q_k := I - {}^\epsilon P_k = I - \epsilon(\cdot) P_k.$$

Lemma 4.1. $P_k^\epsilon, Q_k^\epsilon, {}^\epsilon P_k,$ and ${}^\epsilon Q_k$ are projections in $L^2(\mathcal{Q})$ for which the following orthogonality conditions hold:

(a) $P_k^\epsilon L^2 \perp_{L^2} {}^\epsilon Q_k L^2,$
(b) ${}^\epsilon P_k L^2 \perp_{L^2} Q_k^\epsilon L^2,$
(c) $P_k^\epsilon L^2 \perp_{L^2_\epsilon} Q_k^\epsilon L^2,$
(d) ${}^\epsilon P_k L^2 \perp_{L^2_{\epsilon^{-1}}} {}^\epsilon Q_k L^2,$

where L^2 stands for $L^2(\mathcal{Q})$ and L^2_w is the weighted $L^2(\mathcal{Q})$ by the weight $w(\cdot)$, i.e. $f \perp_{L^2_w} g$ means $\int_\mathcal{Q} f \cdot \overline{g} \, w \, dx = 0$.

Proof. We prove just (a), the proof of the claims (b)–(d) being similar.

Let $f \in {}^\epsilon Q_k L^2(\mathcal{Q})$, that is $f \in L^2(\mathcal{Q})$ such that ${}^\epsilon P_k v = 0$. Hence $\langle f, p_{n_*}(\cdot, k) \rangle = 0$. Let moreover $g \in P_k^\epsilon L^2(\mathcal{Q})$, i.e. $g = P_k^\epsilon g$. Then,

$$\langle f, g \rangle = \langle f, \langle g, \epsilon(\cdot) p_{n_*}(\cdot, k) \rangle \, p_{n_*}(\cdot, k) \rangle = \langle g, \epsilon(\cdot) p_{n_*}(\cdot, k) \rangle \langle f, p_{n_*}(\cdot, k) \rangle = 0.$$

$\square$

In the sequel the operator L_k in (4.2) needs to be inverted with the inverse bounded independently of ε. Recall that for $\omega = \omega_*$ the kernel of L_k is non-trivial at $k \in W_{\omega_*}$, i.e. at $k \in \{k^{(1)}, \ldots, k^{(N)}\}$, cf (A3). For $\omega = \omega_* + \varepsilon^2 \Omega$ the bound on the inverse explodes as $\varepsilon \to 0$. The Lyapunov–Schmidt reduction based on the projections introduced above decomposes the problem into a critical and a regular part. In particular, the projections ${}^\epsilon P_k$ and P_k^ϵ are onto the set of modes, the eigenfunctions of which attain in $\mathbb{B}$ the 'critical' value ω_*, cf (A3). This means, we expect that the complementary projections produce an operator with the inverse bounded independently of ε. This is what we prove in the following result.

Lemma 4.2. There exists $\varepsilon_0 > 0$ such that for all $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$ and $\Omega = \pm 1$ the linear operator $\mathcal{L}_k := {}^\epsilon Q_k L_k Q_k^\epsilon : H_\#(\mathrm{curl}^2, \mathcal{Q}) \to {}^\epsilon Q_k L^2_\#(\mathcal{Q})$ with $\omega = \omega_* + \varepsilon^2 \Omega$ is invertible on

$Q_k^\epsilon H_\#(\mathrm{curl}^2, \mathcal{Q})$ *and*

$$\|\mathcal{L}_k^{-1}\|_{^\epsilon Q_k L_\#^2 \to Q_k^\epsilon H_\#(\mathrm{curl}^2)} \leqslant C_\mathcal{L}, \tag{4.4}$$

where the constant $C_\mathcal{L}$ is independent of ε, Ω, and k.

Remark 10. We point out that the operator $\mathcal{L}_k$ depends on ε via the factor ω^2 in L_k.

Proof. **Step 1.** $\mathcal{L}_k$ *is injective.* Due to the linearity of $\mathcal{L}_k$ it is equivalent to show that if $v \in Q_k^\epsilon H_\#(\mathrm{curl}^2, \mathcal{Q})$ such that $\mathcal{L}_k v = 0$, then $v = 0$. Let thus $v \in H_\#(\mathrm{curl}^2, \mathcal{Q})$ be such that $Q_k^\epsilon v = v$. $\mathcal{L}_k v = 0$ means that

$$L_k Q_k^\epsilon v = L_k v \in \ker{}^\epsilon Q_k = \mathrm{span}\{\epsilon p_{n_*}(\cdot, k)\}.$$

In addition, the assumption $Q_k^\epsilon v = v$ implies $P_k^\epsilon v = 0$, i.e. $\langle v, \epsilon p_{n_*}(\cdot, k)\rangle = 0$. Hence

$$L_k v = \epsilon\langle L_k v, p_{n_*}(\cdot, k)\rangle p_{n_*}(\cdot, k) = \epsilon\langle v, L_k p_{n_*}(\cdot, k)\rangle p_{n_*}(\cdot, k)$$

$$= \epsilon(\omega_{n_*}^2(k) - \omega_*^2)\langle v, \epsilon p_{n_*}(\cdot, k)\rangle p_{n_*}(\cdot, k) = 0.$$

This implies $v = 0$ because $\omega \notin \sigma(L) = \bigcup_{k \in \mathbb{B}} \sigma(L_k)$.

Step 2. $\mathcal{L}_k$ *is surjective.* The aim is to show that for any $f \in {}^\epsilon Q_k L_\#^2(\mathcal{Q})$ there exists $u \in Q_k^\epsilon H_\#(\mathrm{curl}^2, \mathcal{Q})$ such that $\mathcal{L}_k u = f$.

First, notice that $f \in {}^\epsilon Q_k L_\#^2(\mathcal{Q})$ if and only if $f \in L_\#^2(\mathcal{Q})$ and $\langle f, p_{n_*}(\cdot, k)\rangle = 0$. By the closed range theorem the equation $\mathcal{L}_k u = f$ is solvable in the domain of $\mathcal{L}_k$, i.e. in $H_\#(\mathrm{curl}^2, \mathcal{Q})$, if and only if $f \perp \ker(\mathcal{L}_k)$. Here we are using that the operator $\mathcal{L}_k$ is self-adjoint and we postpone the proof to the subsequent lemma 4.3. Let thus $v \in \ker\left({}^\epsilon Q_k L_k Q_k^\epsilon\right)$, i.e. $v = {}^\epsilon P_k L_k Q_k^\epsilon v$, then there holds

$$\langle f, v\rangle = \langle {}^\epsilon Q_k f, {}^\epsilon P_k L_k Q_k^\epsilon v\rangle = \langle {}^\epsilon P_k L_k {}^\epsilon Q_k f, Q_k^\epsilon v\rangle,$$

where the symmetry of ${}^\epsilon P_k L_k$ is shown in lemma 4.3. Noticing that by lemma 4.1(b) the spaces ${}^\epsilon P_k L^2$ and $Q_k^\epsilon L^2$ are L^2-orthogonal, we deduce $\langle f, v\rangle = 0$, i.e. $f \perp \ker(\mathcal{L}_k)$. The sought function is then $Q_k^\epsilon u$, with u given by the closed range theorem.

Step 3. The estimate (4.4). Recall first that for a linear self-adjoint operator $\mathcal{A}$ acting on a Hilbert space, the well-known estimate

$$\|\mathcal{A}^{-1}\| \leqslant \mathrm{dist}(0, \sigma(\mathcal{A}))^{-1}$$

holds, where $\sigma(\mathcal{A})$ denotes its spectrum. The self-adjointness of ${}^\epsilon Q_k L_k Q_k^\epsilon$ is shown in lemma 4.3(b), so we need to bound $\mathrm{dist}(0, \sigma({}^\epsilon Q_k L_k Q_k^\epsilon))^{-1}$.

For a fixed ω the spectrum of L_k is given by $\cup_{j \in \mathbb{N}}\{\omega_j^2(k) - \omega^2\}$ as shown in sections 2.2 and 2.3. The application of the projections ${}^\epsilon Q_k$ and Q_k^ϵ yields $\sigma(\mathcal{L}_k) = \bigcup_{j \in \mathbb{N}\setminus\{n_*\}}\{\omega_j^2(k) - \omega^2\}$. By our assumptions on the band structure (A3) we infer that each of the remaining eigenvalues has some positive distance to ω_*, hence for all $j \neq n_*$ we have $\inf_{k \in \mathbb{B}}|\omega_j^2(k) - \omega_*^2| =: m_j > 0$. Since the map $j \mapsto \omega_j(k)$ is increasing for every $k \in \mathbb{B}$ fixed, $m := \min_{j \neq n_1, \dots, n_N} m_j > 0$ is well-defined and we infer

$$\|({}^\epsilon Q_k L_k Q_k^\epsilon)^{-1}\|_{^\epsilon Q_k L_\#^2 \to Q_k^\epsilon H_\#(\mathrm{curl}^2)}$$

$$\leqslant \mathrm{dist}(0, \sigma({}^\epsilon Q_k L_k Q_k^\epsilon))^{-1} \leqslant 2\left(\min_{j \neq n_1, \dots, n_N}\inf_{k \in \mathbb{B}}|\omega_j^2(k) - \omega^2|\right)^{-1} = \frac{2}{m} =: C_\mathcal{L}$$

for all ε small enough. $\qquad\square$

Remark 11. Notice that without the projections estimate (4.4) does not hold even if $\Omega \in \{-1, 1\}$ is chosen such that $\omega = \omega_* + \varepsilon^2\Omega$ lies outside the spectrum. The operator L_k would be invertible but by (4.1) we would only get $\text{dist}(0, \sigma(L_k)) = O(\varepsilon^2)$, whence $\|L_k^{-1}\|$ would not be bounded uniformly in ε.

Lemma 4.3.

(a) *The operator $^\epsilon P_k L_k : H_\#(\text{curl}^2, \mathcal{Q}) \to L^2_\#(\mathcal{Q})$ is symmetric.*
(b) *The operator $\mathcal{L}_k : H_\#(\text{curl}^2, \mathcal{Q}) \to {}^\epsilon Q_k L^2_\#(\mathcal{Q})$ is self-adjoint.*

Proof. (i) Let $v, w \in H_\#(\text{curl}^2, \mathcal{Q})$, then

$$
\begin{aligned}
\langle {}^\epsilon P_k L_k v, w \rangle &= \langle P_k L_k v, \epsilon w \rangle = \langle L_k v, p_{n_*}(\cdot, k) \rangle \langle p_{n_*}(\cdot, k), \epsilon w \rangle \\
&= \langle v, L_k p_{n_*}(\cdot, k) \rangle \langle \epsilon p_{n_*}(\cdot, k), w \rangle \\
&= \langle v, \epsilon p_{n_*}(\cdot, k) \rangle \langle \omega_{n_j}(k_*)^2 \epsilon p_{n_*}(\cdot, k), w \rangle \\
&= \langle v, \epsilon p_{n_*}(\cdot, k) \rangle \langle L_k p_{n_*}(\cdot, k), w \rangle \\
&= \langle v, \epsilon p_{n_j}(\cdot, k) \rangle \langle p_{n_*}(\cdot, k), L_k w \rangle \\
&= \langle v, \epsilon p_{n_*}(\cdot, k) \rangle \overline{\langle L_k w, p_{n_*}(\cdot, k) \rangle} = \langle v, {}^\epsilon P_k L_k w \rangle.
\end{aligned}
$$

(ii) First we claim that for $v \in H_\#(\text{curl}^2, \mathcal{Q})$ one can rewrite $\mathcal{L}_k v$ as

$$
\mathcal{L}_k v = L_k v - (\omega_*^2 - \omega^2) \langle v, \epsilon p_{n_*}(\cdot, k) \rangle \epsilon p_{n_*}(\cdot, k). \tag{4.5}
$$

Indeed, recalling the expression of the projections $^\epsilon Q_k$ and Q_k^ϵ, one finds

$$
\begin{aligned}
\mathcal{L}_k v &= {}^\epsilon Q_k L_k Q_k^\epsilon v = L_k Q_k^\epsilon v - \langle L_k Q_k^\epsilon v, p_{n_*}(\cdot, k) \rangle \epsilon p_{n_*}(\cdot, k) \\
&= L_k \left(v - \langle v, \epsilon p_{n_*}(\cdot, k) \rangle p_{n_*}(\cdot, k) \right) \\
&\quad - \langle \left(v - \langle v, \epsilon p_{n_*}(\cdot, k) \rangle p_{n_*}(\cdot, k) \right), L_k p_{n_*}(\cdot, k) \rangle \epsilon p_{n_*}(\cdot, k) \\
&= L_k v - \langle v, \epsilon p_{n_*}(\cdot, k) \rangle L_k p_{n_*}(\cdot, k) - \langle v, L_k p_{n_*}(\cdot, k) \rangle \epsilon p_{n_*}(\cdot, k) \\
&\quad + \langle v, \epsilon p_{n_*}(\cdot, k) \rangle \langle p_{n_*}(\cdot, k), L_k p_{n_*}(\cdot, k) \rangle \epsilon p_{n_*}(\cdot, k),
\end{aligned}
$$

having used the symmetry of L_k on $H_\#(\text{curl}^2, \mathcal{Q})$. Identity (4.5) is then proven recalling that $p_{n_*}(\cdot, k)$ are normalized as in (2.11) and $L_k p_{n_*}(\cdot, k) = (\omega_*^2 - \omega^2) \epsilon p_{n_*}(\cdot, k)$.

Next, we show that $D(\mathcal{L}_k^*) \subset H_\#(\text{curl}^2, \mathcal{Q})$, where $\mathcal{L}_k^*$ denotes the adjoint operator of $\mathcal{L}_k$. Let $\phi \in D(\mathcal{L}_k^*)$, then there exists $\eta \in L^2_\#(\mathcal{Q})$ so that $\langle \mathcal{L}_k v, \phi \rangle = \langle v, \eta \rangle$ for all $v \in D(\mathcal{L}_k) = H_\#(\text{curl}^2, \mathcal{Q})$. Using (4.5) and reorganising the terms, one obtains

$$
\langle L_k v, \phi \rangle = \langle v, \eta + (\omega_*^2 - \omega^2) \langle \phi, \epsilon p_{n_*}(\cdot, k) \rangle \epsilon p_{n_*}(\cdot, k) \rangle =: \langle v, \tilde{\eta} \rangle \tag{4.6}
$$

with $\tilde{\eta} \in L^2_\#(\mathcal{Q})$. This means that $\phi \in D(L_k^*) = D(L_k) = H_\#(\text{curl}^2, \mathcal{Q})$ since L_k is self-adjoint [8].

Finally, for $\phi \in H_\#(\text{curl}^2, \mathcal{Q})$ one has $\langle L_k v, \phi \rangle = \langle v, L_k \phi \rangle$, therefore from (4.5) and (4.6) one infers

$$
\begin{aligned}
\langle \mathcal{L}_k v, \phi \rangle &= \langle v, L_k \phi \rangle - (\omega_*^2 - \omega^2) \langle v, \epsilon p_{n_*}(\cdot, k) \rangle \langle \epsilon p_{n_*}(\cdot, k), \phi \rangle \\
&= \langle v, L_k \phi - (\omega_*^2 - \omega^2) \langle \phi, \epsilon p_{n_*}(\cdot, k) \rangle \epsilon p_{n_*}(\cdot, k) \rangle = \langle v, \mathcal{L}_k \phi \rangle.
\end{aligned}
$$

This implies $\phi \in D(\mathcal{L}_k^*)$ and that $\mathcal{L}_k$ is symmetric. Having proved that $D(\mathcal{L}_k^*) = H_\#(\mathrm{curl}^2, \mathcal{Q}) = D(\mathcal{L}_k)$ and the symmetry, we conclude that $\mathcal{L}_k$ is self-adjoint. $\qquad\square$

Decomposition of the solution. We decompose the solution $\widetilde{u}$ of (4.2) using the above projections as

$$\widetilde{u}(x,k) = \widetilde{v}(x,k) + \widetilde{w}(x,k), \tag{4.7}$$

where

$$\widetilde{v}(x,k) := P_k^\epsilon \widetilde{u}(x,k) = \langle \widetilde{u}(\cdot,k), \epsilon(\cdot) p_{n_*}(\cdot,k) \rangle\, p_{n_*}(x,k) =: U(k) p_{n_*}(x,k)$$

and

$$\widetilde{w}(x,k) := Q_k^\epsilon \widetilde{u}(x,k),$$

and where we have defined

$$U(k) := \langle \widetilde{u}(\cdot,k), \epsilon(\cdot) p_{n_*}(\cdot,k) \rangle.$$

We note that U is Λ-periodic because $\widetilde{u}(x,\cdot)$ and $p_{n_*}(x,\cdot)$ are quasiperiodic.

Now we suitable project equation (4.2) and find an equivalent system of two equations, the linear part of which is decoupled. On the one hand, applying P_k to (4.2), we find

$$\omega^2 \langle \widetilde{F}(\widetilde{u})(\cdot,k), p_{n_*}(\cdot,k) \rangle = \langle \widetilde{u}(\cdot,k), \left(\nabla_k' \times \nabla_k' \times -\omega^2 \epsilon(\cdot) \right) p_{n_*}(\cdot,k) \rangle$$

$$= \langle \widetilde{u}(\cdot,k), \left(\omega_{n_*}^2(k) - \omega^2 \right) \epsilon(\cdot) p_{n_*}(\cdot,k) \rangle.$$

By the definition of U we have

$$\left(\omega_{n_*}^2(k) - \omega^2 \right) U(k) = \omega^2 \langle \widetilde{F}(\widetilde{u})(\cdot,k), p_{n_*}(\cdot,k) \rangle. \tag{4.8}$$

On the other hand, we get

$${}^\epsilon Q_k L_k Q_k^\epsilon \widetilde{w}(x,k) = \omega^2\, {}^\epsilon Q_k \widetilde{f}(\widetilde{u})(x,k). \tag{4.9}$$

Indeed, $L_k \widetilde{u}(x,k) = L_k\left(\widetilde{v}(x,k) + \widetilde{w}(x,k) \right)$, where $\widetilde{w}(x,k) = Q_k^\epsilon \widetilde{w}(x,k)$ and

$$\begin{aligned}
{}^\epsilon Q_k L_k \widetilde{v}(x,k) &= {}^\epsilon Q_k \left(U(k) \left(\omega_{n_*}^2(k) - \omega^2 \right) \epsilon(\cdot) p_{n_*}(\cdot,k) \right) \\
&= U(k) \left(\omega_{n_*}^2(k) - \omega^2 \right) (I - \epsilon(\cdot)P_k) \left(\epsilon(\cdot) p_{n_*}(\cdot,k) \right) \\
&= U(k) \left(\omega_{n_*}^2(k) - \omega^2 \right) \big(\epsilon(\cdot) p_{n_*}(\cdot,k) \\
&\quad - \epsilon(\cdot) \langle \epsilon(\cdot) p_{n_*}(\cdot,k), p_{n_*}(\cdot,k) \rangle p_{n_*}(\cdot,k) \big) \\
&= 0
\end{aligned}$$

by the normalization (2.11) of the Bloch eigenfunctions.

Next, we decompose $\widetilde{w}$ further into

$$\widetilde{w}(x,k) = \widetilde{w}_0(x,k) + \widetilde{w}_R(x,k),$$

where $\widetilde{w}_0$ and $\widetilde{w}_R$ solve the equations

$${}^\epsilon Q_k L_k Q_k^\epsilon \widetilde{w}_0(x,k) = \omega^2\, {}^\epsilon Q_k \widetilde{F}(\widetilde{v})(x,k), \tag{4.10}$$

$$^{\epsilon}Q_k L_k Q_k^{\epsilon} \tilde{w}_R(x,k) = \omega^{2\epsilon} Q_k(\tilde{F}(\tilde{u}) - \tilde{F}(\tilde{v}))(x,k). \tag{4.11}$$

The system (4.8), (4.10), and (4.11) is an equivalent reformulation of equation (4.2). We search for a solution $\tilde{u}$ (represented by the variables $U, \tilde{w}_0, \tilde{w}_R$) which is close to the Bloch transformation of the formal ansatz u_{ans}. In detail, for the sought solution the components $\tilde{w}_0$ and $\tilde{w}_R$ are small and U is concentrated at the points $k^{(j)}$, $j = 1, \ldots, N$ and near the concentration points $k^{(j)}$ it approximates $\hat{A}_j$, where A is a solution of the CMEs. The Bloch transformation of the formal ansatz is

$$\tilde{u}_{\text{ans}}(x,k) = \frac{1}{\varepsilon} \sum_{j=1}^{N} \sum_{K \in \Lambda^*} \hat{A}_j \left(\frac{k - k^{(j)} + K}{\varepsilon} \right) p_{n_*}(x, k^{(j)}) e^{iK \cdot x} \tag{4.12}$$

using (2.25) and the fact that $\left(A(\varepsilon \cdot) e^{ik^{(j)} \cdot} \right)^{\wedge}(k) = \varepsilon^{-2} \hat{A}(\varepsilon^{-1}(k - k^{(j)}))$. Since $\hat{A}_j(\varepsilon^{-1}(k - k^{(j)}))$ is concentrated near $k = k^{(j)}$, we decompose U on $\mathbb{B}$ into $N + 1$ parts with the first N being compactly supported in the vicinity of one of the points $k^{(j)}$ and the last one supported away from all $k^{(j)}$. This is then extended Λ^*-periodically onto $\mathbb{R}^2$. We write

$$U(k) = \sum_{k=1}^{N} \frac{1}{\varepsilon} \hat{B}_j \left(\frac{k - k^{(j)}}{\varepsilon} \right) + \hat{C}(k), \quad k \in \mathbb{R}^2,$$

where

$$\text{supp} \left(\hat{B}_j \left(\frac{\cdot - k^{(j)}}{\varepsilon} \right) \right) \cap \mathbb{B} = B_{\varepsilon^r}(k^{(j)}) \quad \text{and} \quad \text{supp}(\hat{C}) \cap \mathbb{B} = \mathbb{B} \setminus \bigcup_{j=1}^{N} B_{\varepsilon^r}(k^{(j)}) \tag{4.13}$$

and where $\hat{C}$ is Λ^*-periodic and $\hat{B}_j$ is $\varepsilon^{-1}\Lambda^*$-periodic on $\mathbb{R}^2$. That means

$$\hat{C}(k + K) = \hat{C}(k), \qquad \hat{B}_j(k + \varepsilon^{-1}K) = \hat{B}_j(k) \quad \text{for all } K \in \Lambda^* \text{ and } k \in \mathbb{R}^2.$$

Here $r \in (0,1)$ is a parameter to be specified to suit the nonlinear estimates. Moreover we define $\hat{B}_j^*$ and $\hat{C}^*$ as the restrictions of such functions to the periodicity cell, i.e.

$$\hat{B}_j^* := \chi_{\varepsilon^{-1}\mathbb{B}} \hat{B}_j \quad \text{and} \quad \hat{C}^* := \chi_{\mathbb{B}} \hat{C}.$$

We point out that in $\hat{B}_j$ and $\hat{C}$ the $\hat{\ }$ notation does not refer to the Fourier transform of given functions B_j, C; it just stresses out the connection between $\hat{B}_j$ and $\hat{A}_j$, the latter of course being the Fourier transform of A_j. With this further decomposition the sought solution has the components $\hat{B}_j^*$ close to $\hat{A}_j$ and the component $\hat{C}^*$ small. Note that $\hat{A}_j$, $j = 1, \ldots, N$, satisfies equation (1.10) transformed in Fourier variables, i.e.

$$\Omega \hat{A}_j(k') - \frac{1}{2} \left((k_1')^2 \partial_{k_1}^2 \omega_{n_*}(k^{(j)}) + (k_2')^2 \partial_{k_2}^2 \omega_{n_*}(k^{(j)}) + 2k_1' k_2 \partial_{k_1 k_2}^2 \omega_{n_*}(k^{(j)}) \right) \hat{A}_j(k') + \hat{\mathcal{N}}_j(k') = 0, \; k' \in \mathbb{R}^2.$$

$$\tag{4.14}$$

The aim now is to apply a fixed point argument to solve system (4.8), (4.10), and (4.11), which is of course coupled in the components $(\hat{B}_j)_{j=1}^{N}, \hat{C}, \tilde{w}_0, \tilde{w}_R$. The equations for the components $\tilde{w}_0$ and $\tilde{w}_R$ both involve the linear operator $^{\epsilon}Q_k L_k Q_k^{\epsilon}$, see (4.10) and (4.11). This operator

is boundedly invertible on its image by lemma 4.2, and the bound on the inverse is independent of ε. This is thanks to the fact that the projection Q_k^ϵ projects the Bloch eigenfunction p_{n_*} out.

We will also make use of the notation

$$\widetilde{v}_B(x,k) := \sum_{j=1}^{N} \widetilde{v}_{B_j}(x,k) \quad \text{where} \quad \widetilde{v}_{B_j}(x,k) := \frac{1}{\varepsilon}\widehat{B}_j\left(\frac{k-k^{(j)}}{\varepsilon}\right)p_{n_*}(x,k) \quad (4.15)$$

and

$$\widetilde{v}_C(x,k) := \widehat{C}(k)p_{n_*}(x,k).$$

We can thus write

$$\widetilde{v}(x,k) = \widetilde{v}_B(x,k) + \widetilde{v}_C(x,k).$$

Inspired by the strategy of [9, 12, 13], our algorithm to construct a solution $\widetilde{u} \in \mathcal{X}_2$ of problem (4.2) is the following nested fixed point argument.

(a) For any given $\widetilde{v}$ bounded, determine the unique small solution $\widetilde{w}_0$ of the linear problem (4.10) by means of lemma 4.2;

(b) For any given $\widetilde{v}$ bounded and $\widetilde{w}_0$ from step 1, apply the Banach fixed point theorem to (4.11) in a neighbourhood of zero to find a small solution $\widetilde{w}_R$;

(c) For any given N-tuple $(\widehat{B}_j)_{j=1}^N$ with $(\widehat{B}_j^*)_{j=1}^N$ decaying sufficiently fast, find a small $\widehat{C}$ with support as in (4.13) applying the Banach fixed point theorem to (4.8) on this support;

(d) For $\widehat{C}$ given by step 3 prove the existence of such solutions $(\widehat{B}_j^*)_{j=1}^N$ to (4.8) which are close to $(\widehat{A}_j)_{j=1}^N$ and have the support as in (4.13). It is here where the restriction to a $\mathcal{PT}$-symmetric subspace is used.

The rest of the section carries this algorithm out.

4.2. Preliminary estimates

We define for convenience $L^2 := L^2(\mathbb{R}^2)$ and $L_{s_B}^2 := L_{s_B}^2(\mathbb{R}^2)$ (the weighted space $L_s^2(\mathbb{R}^2)$ defined in (2.28)).

Lemma 4.4. *The following inequalities hold*

$$\|\widetilde{v}_B\|_{\mathcal{X}_2} \leqslant c\sum_{j=1}^{N}\|\widehat{B}_j\|_{L^2(\varepsilon^{-1}\mathbb{B})}, \qquad \|\widetilde{v}_C\|_{\mathcal{X}_2} \leqslant c\,\|\widehat{C}\|_{L^2(\mathbb{B})}.$$

Proof. Using the regularity result in lemma A.3, we have $\|p_{n_*}(\cdot,k)\|_{H^2(\mathcal{Q})}$ uniformly bounded in $k \in \mathbb{B}$. Therefore,

$$\|\widetilde{v}_C\|_{\mathcal{X}_2} \leqslant \operatorname{ess\,sup}_{k\in\mathbb{B}}\|p_{n_*}(\cdot,k)\|_{H^2(\mathcal{Q})}\|\widehat{C}\|_{L^2(\mathbb{B})} \leqslant c\,\|\widehat{C}\|_{L^2(\mathbb{B})},$$

$$\|\widetilde{v}_B\|_{\mathcal{X}_2} \leqslant \operatorname{ess\,sup}_{k\in\mathbb{B}}\|p_{n_*}(\cdot,k)\|_{H^2(\mathcal{Q})}\frac{1}{\varepsilon}\sum_{j=1}^{N}\left\|\widehat{B}_j\left(\frac{\cdot - k^{(j)}}{\varepsilon}\right)\right\|_{L^2(\mathbb{B}+k^{(j)})}$$

$$\leqslant c\sum_{j=1}^{N}\left(\int_{\varepsilon^{-1}\mathbb{B}}|\widehat{B}_j(t)|^2\,\mathrm{d}t\right)^{\frac{1}{2}} = c\sum_{j=1}^{N}\|\widehat{B}_j\|_{L^2(\varepsilon^{-1}\mathbb{B})}.$$

Note that in the first inequality for $\|\widetilde{v}_B\|_{\mathcal{X}_2}$ the $k^{(j)}$-shift of the k-integral is allowed due to the periodicity of $\widehat{B}_j$. $\qquad\square$

Clearly, the estimate in lemma 4.4 is $O(1)$ in ε. In estimating the residual below, it will be necessary to show the smallness of the nonlinearity in the $\mathcal{X}_2$-norm for ε small. Inspired by (33) in [14], the next lemma produces this smallness for the components $\widetilde{v}_B$ if $s_B \geqslant 2$.

Lemma 4.5. *Let $s_B \geqslant 2$. For $a, b \in \{1, 2, 3\}$ and $\widetilde{v}_B$ defined in (4.15) there holds*

$$\|\widetilde{v}_{B,a} *_{\mathbb{B}} \widetilde{v}_{B,b}\|_{\mathcal{X}_2} \leqslant c\varepsilon \sum_{i,j=1}^{N} \|\widehat{B}_i^*\|_{L_{s_B}^2} \|\widehat{B}_j^*\|_{L_{s_B}^2}.$$

Proof. For $a \in \{1, 2, 3\}$ we define $v_{B,a} := \mathcal{T}^{-1}\widetilde{v}_{B,a}$. First,

$$\|\widetilde{v}_{B,a} *_{\mathbb{B}} \widetilde{v}_{B,b}\|_{\mathcal{X}_2}^2 = \|\widetilde{v_{B,a} v_{B,b}}\|_{\mathcal{X}_2}^2 \leqslant c\|v_{B,a} v_{B,b}\|_{H^2(\mathbb{R}^2)}^2, \tag{4.16}$$

by the isomorphism (2.26). We now claim that

$$\|v_{B,a} v_{B,b}\|_{H^2(\mathbb{R}^2)}^2 \leqslant c\varepsilon^4 \sum_{i,j=1}^{N} \|B_i^*(\varepsilon\cdot)B_j^*(\varepsilon\cdot)\|_{H^2(\mathbb{R}^2)}^2, \tag{4.17}$$

where $B_i^* := (\widehat{B}_i^*)^\vee$ for $i \in \{1, \ldots, N\}$. Indeed, for $i, j \in \{1, \ldots, N\}$ one has

$$\|v_{B_i,a} v_{B_j,b}\|_{L^2(\mathbb{R}^2)}^2 \leqslant c\varepsilon^4 \mathrm{ess}\sup_{k \in \mathbb{B}} \|p_{n_*}(\cdot, k)\|_\infty^2 \int_{\mathbb{C}^2} \prod_{l=i,j} \left| \int_{\varepsilon^{-1}\mathbb{B}} \widehat{B}_l(\kappa) \mathrm{e}^{\mathrm{i}\varepsilon\kappa\cdot x} \, \mathrm{d}\kappa \right|^2 \mathrm{d}x$$

$$\leqslant c\varepsilon^4 \|B_i^*(\varepsilon\cdot)B_j^*(\varepsilon\cdot)\|_{L^2(\mathbb{C}^2)}^2,$$

$$\|\nabla\left(v_{B_i,a} v_{B_j,b}\right)\|_{L^2(\mathbb{R}^2)}^2 \leqslant c\varepsilon^4 \mathrm{ess}\sup_{k \in \mathbb{B}} \|p_{n_*}(\cdot, k)\|_{W^{1,\infty}}^2$$

$$\times \int_{\mathbb{R}^2} \prod_{l=i,j} \left| \int_{\varepsilon^{-1}\mathbb{B}} (1 + \mathrm{i}\varepsilon\kappa)\widehat{B}_l(\kappa) \mathrm{e}^{\mathrm{i}\varepsilon\kappa\cdot x} \, \mathrm{d}\kappa \right|^2 \mathrm{d}x$$

$$\leqslant \varepsilon^4 \|B_i^*(\varepsilon\cdot)B_j^*(\varepsilon\cdot)\|_{H^1(\mathbb{R}^2)}^2$$

and analogously one infers

$$\|\nabla^2\left(v_{B_i,a} v_{B_j,b}\right)\|_{L^2(\mathbb{R}^2)}^2 \leqslant \varepsilon^4 \|B_i^*(\varepsilon\cdot)B_j^*(\varepsilon\cdot)\|_{H^2(\mathbb{R}^2)}^2.$$

Notice that here we used the regularity estimate for the eigenfunctions given by lemma A.4.

Therefore, combining (4.16) and (4.17) one finds

$$\|\widetilde{v}_{B,a}*_{\mathbb{B}}\widetilde{v}_{B,b}\|_{\mathcal{X}_2}^2 \leqslant c\varepsilon^4 \sum_{i,j=1}^{N} \|B_i^*(\varepsilon\cdot)B_j^*(\varepsilon\cdot)\|_{H^2(\mathbb{R}^2)}^2$$

$$= c\varepsilon^4 \sum_{i,j=1}^{N} \int_{\mathbb{R}^2} (1+|k|^2)^2 \left|\left(B_i^*(\varepsilon\cdot)B_j^*(\varepsilon\cdot)\right)^{\wedge}(k)\right|^2 \mathrm{d}k$$

$$\leqslant c\varepsilon^4 \sum_{i,j=1}^{N} \int_{\mathbb{R}^2} \left|\left(B_i^*(\varepsilon\cdot)B_j^*(\varepsilon\cdot)\right)^{\wedge}(k)\right|^2 \mathrm{d}k + c\varepsilon^4 \sum_{i,j=1}^{N} \int_{\mathbb{R}^2} |k|^4 \left|\left(B_i^*(\varepsilon\cdot)B_j^*(\varepsilon\cdot)\right)^{\wedge}(k)\right|^2 \mathrm{d}k$$

$$= c\varepsilon^4 \sum_{i,j=1}^{N} \|B_i^*(\varepsilon\cdot)B_j^*(\varepsilon\cdot)\|_{L^2(\mathbb{R}^2)}^2 + c\varepsilon^4 \sum_{i,j=1}^{N} \int_{\mathbb{R}^2} \left|\varepsilon^{-1}k\right|^4 \left|\left(B_i^*B_j^*\right)^{\wedge}\left(\varepsilon^{-1}k\right)\right|^2 \mathrm{d}k$$

$$\leqslant c\varepsilon^2 \sum_{i,j=1}^{N} \|B_i^*\|_{L^\infty}^2 \|B_j^*\|_{L^2(\mathbb{R}^2)}^2 + c\varepsilon^6 \sum_{i,j=1}^{N} \|\left(B_i^*B_j^*\right)^{\wedge}\|_{L_2^2(\mathbb{R}^2)}^2$$

$$\leqslant c\varepsilon^2 \sum_{i,j=1}^{N} \|B_i^*\|_{H^2(\mathbb{R}^2)}^2 \|B_j^*\|_{H^2(\mathbb{R}^2)}^2$$

$$\leqslant c\varepsilon^2 \sum_{i,j=1}^{N} \|\widehat{B}_i^*\|_{L_2^2(\mathbb{R}^2)}^2 \|\widehat{B}_j^*\|_{L_2^2(\mathbb{R}^2)}^2.$$

Notice that we have used the embedding $H^2(\mathbb{R}^2) \hookrightarrow L^\infty(\mathbb{R}^2)$ and, to conclude, the isomorphism property of the Fourier transform between $H^2(\mathbb{R}^2)$ and $L_2^2(\mathbb{R}^2)$. $\qquad\square$

For the analysis of the nonlinearity we need to calculate the double convolutions $\widetilde{v}_a*_{\mathbb{B}}\widetilde{v}_b*_{\mathbb{B}}\overline{\widetilde{v}}_c$ (for $a,b,c \in \{1,2,3\}$) appearing in $\widetilde{F}(\widetilde{v})$. We have

$$(\widetilde{v}_a*_{\mathbb{B}}\widetilde{v}_b*_{\mathbb{B}}\overline{\widetilde{v}}_c)(x,k) = \int_{\mathbb{B}} \left(\int_{\mathbb{B}} \widetilde{v}_a(k-s-t,x)\widetilde{v}_b(t,x)\mathrm{d}t\right) \overline{\widetilde{v}}_c(s,x)\mathrm{d}s$$

$$= \int_{2\mathbb{B}}\int_{\mathbb{B}} \widetilde{v}_a(k-l,x)\widetilde{v}_b(l-s,x)\overline{\widetilde{v}}_c(s,x)\mathrm{d}s\,\mathrm{d}l \qquad (4.18)$$

using the transformation $l = s + t$. Also note that the integration domains $2\mathbb{B}$ and $\mathbb{B}$ can be both shifted by an arbitrary $k_* \in \mathbb{R}^2$ due to the quasi-periodicity of $\widetilde{v}_a, \widetilde{v}_b$, and $\widetilde{v}_c$ with respect to the variable k.

Lemma 4.6. *Let $\widehat{B}_j \in L_{s_B}^2(\varepsilon^{-1}\mathbb{B})$ with $s_B > 1$ and $\widehat{C} \in L^2(\mathbb{B})$ have the supports as in (4.13). Then*

$$\|\widetilde{F}(\widetilde{v})\|_{\mathcal{X}_2} \leqslant c\varepsilon^2 \left(\sum_{\alpha,\beta,\gamma=1}^{N} \|\widehat{B}_\alpha^*\|_{L_{s_B}^2}\|\widehat{B}_\beta^*\|_{L_{s_B}^2}\|\widehat{B}_\gamma^*\|_{L_{s_B}^2} + \sum_{\alpha,\beta=1}^{N} \|\widehat{B}_\alpha^*\|_{L_{s_B}^2}\|\widehat{B}_\beta^*\|_{L_{s_B}^2}\|\widehat{C}^*\|_{L^2}\right)$$

$$+ c\varepsilon \sum_{\alpha=1}^{N} \|\widehat{B}_\alpha^*\|_{L_{s_B}^2}\|\widehat{C}^*\|_{L^2}^2 + c\|\widehat{C}^*\|_{L^2}^3. \qquad (4.19)$$

For $\widetilde{F}(\widetilde{v}_B)$ we have for each $d = 1, 2, 3$

$$
\tilde{F}_d(\tilde{v}_B)(x, k) = \varepsilon \sum_{a,b,c=1}^{3} \underline{\chi}_{a,b,c,d}^{(3)}(x) \sum_{\alpha,\beta,\gamma=1}^{N} \int_{2\varepsilon^{-1}\mathbb{B}} \int_{\varepsilon^{-1}\mathbb{B}} \widehat{B}_\alpha \left(\frac{k - (k^{(\alpha)} + k^{(\beta)} - k^{(\gamma)})}{\varepsilon} - l \right)
$$
$$
\cdot \widehat{B}_\beta(l - t) \overline{\widehat{B}}_\gamma(t) \, p_{n_*,a}(x, k - k^{(\beta)} + k^{(\gamma)} - \varepsilon l)
$$
$$
\cdot p_{n_*,b}(x, k^{(\beta)} + \varepsilon(l - t)) \, p_{n_*,c}(x, -k^{(\gamma)} + \varepsilon t) \, \mathrm{d}t \, \mathrm{d}l \tag{4.20}
$$

and the estimate

$$
\|\widetilde{F}(\tilde{v}_B)\|_{\mathcal{X}_2} \leqslant c\varepsilon^2 \sum_{\alpha,\beta,\gamma=1}^{N} \|\widehat{B}_\alpha^*\|_{L_{s_B}^2} \|\widehat{B}_\beta^*\|_{L_{s_B}^2} \|\widehat{B}_\gamma^*\|_{L_{s_B}^2}. \tag{4.21}
$$

Proof.   First, notice that

$$
\widetilde{v}_B(x, k) = \frac{1}{\varepsilon} \sum_{j=1}^{N} \widehat{B}_j \left(\frac{k + k^{(j)}}{\varepsilon} \right) p_{n_*}(x, k).
$$

Indeed,

$$
\widetilde{v}_B(x) = \frac{1}{\varepsilon} \sum_{j=1}^{N} \int_{\mathbb{B}} \overline{\widehat{B}}_j \left(\frac{k - k^{(j)}}{\varepsilon} \right) \overline{p_{n_*}(x, k)} \mathrm{e}^{-\mathrm{i}k \cdot x} \, \mathrm{d}k
$$
$$
= \frac{1}{\varepsilon} \sum_{j=1}^{N} \int_{\mathbb{B}} \widehat{\overline{B}}_j \left(\frac{-k + k^{(j)}}{\varepsilon} \right) \overline{p_{n_*}(x, k)} \mathrm{e}^{-\mathrm{i}k \cdot x} \, \mathrm{d}k
$$
$$
= \frac{1}{\varepsilon} \sum_{j=1}^{N} \int_{\mathbb{B}} \widehat{\overline{B}}_j \left(\frac{k + k^{(j)}}{\varepsilon} \right) \overline{p_{n_*}(x, -k)} \mathrm{e}^{\mathrm{i}k \cdot x} \, \mathrm{d}k
$$

and finally we use the symmetry $\overline{p_{n_*}(x, -k)} = p_{n_*}(x, k)$, see (2.12).

Similarly,

$$
\widetilde{v}_C(x, k) = \widehat{\overline{C}}(k) p_{n_*}(x, k).
$$

We start with formula (4.20). Using (4.18), we have

$$
\tilde{F}_d(\tilde{v}_B)(x, k) = \varepsilon \sum_{a,b,c=1}^{3} \underline{\chi}_{a,b,c,d}^{(3)}(x) \sum_{\alpha,\beta,\gamma=1}^{N} \int_{2\mathbb{B}+k^{(\beta)}-k^{(\gamma)}} \int_{\mathbb{B}-k^{(\gamma)}} \widehat{B}_\alpha \left(\frac{k - l - k^{(\alpha)}}{\varepsilon} \right)
$$
$$
\cdot \widehat{B}_\beta \left(\frac{l - t - k^{(\beta)}}{\varepsilon} \right) \widehat{\overline{B}}_\gamma \left(\frac{t + k^{(\gamma)}}{\varepsilon} \right) p_{n_*,a}(x, k - l) p_{n_*,b}(x, l - t) p_{n_*,c}(x, t) \, \mathrm{d}t \, \mathrm{d}l,
$$

where we have used (4.18) and the fact that due to the quasi-periodicity in k the convolution domains can be shifted by arbitrary $k_* \in \mathbb{R}^2$. The transformations $t' = \varepsilon^{-1}(t + k^{(\gamma)})$, $l' = \varepsilon^{-1}(l - k^{(\beta)} + k^{(\gamma)})$ produce (4.20).

The estimates (4.19) and (4.21) are proved next. Let us denote $\|\chi^{(3)}\|_{H^2(\mathcal{Q})} := \max_{a,b,c,d\in\{1,2,3\}}\|\chi^{(3)}_{a,b,c,d}\|_{H^2(\mathcal{Q})}$. Using the assumption on $\chi^{(3)}$ in (A5),

$$\|\widetilde{F}(\widetilde{v})\|_{\mathcal{X}_2} \leqslant c\, \|\chi^{(3)}\|_{H^2(\mathcal{Q})} \max_{a\in\{1,2,3\}} \|p_{n_*,a}\|^3_{L^\infty(\mathbb{B},H^2(\mathcal{Q}))}$$

$$\cdot \left[\varepsilon^{-6} \sum_{\alpha,\beta,\gamma=1}^{N} \int_{\mathbb{B}} \left| \int_{2\mathbb{B}}\int_{\mathbb{B}} \widehat{B}_\alpha\left(\frac{k-l-k^{(\alpha)}}{\varepsilon}\right) \widehat{B}_\beta\left(\frac{l-t-k^{(\beta)}}{\varepsilon}\right) \right.\right.$$

$$\times \overline{\widehat{B}}_\gamma\left(\frac{t+k^{(\gamma)}}{\varepsilon}\right) \mathrm{d}t\,\mathrm{d}l \Bigg|^2 \mathrm{d}k + \varepsilon^{-4}\sum_{\alpha,\beta=1}^{N} \int_{\mathbb{B}}\left| \int_{2\mathbb{B}}\int_{\mathbb{B}} \right.$$

$$\times \widehat{B}_\alpha\left(\frac{k-l-k^{(\alpha)}}{\varepsilon}\right) \widehat{B}_\beta\left(\frac{l-t-k^{(\beta)}}{\varepsilon}\right) \overline{\widehat{C}}(t)\mathrm{d}t\,\mathrm{d}l \Bigg|^2 \mathrm{d}k$$

$$+ \varepsilon^{-2}\sum_{\alpha=1}^{N} \int_{\mathbb{B}}\left| \int_{2\mathbb{B}}\int_{\mathbb{B}} \widehat{B}_\alpha\left(\frac{k-l-k^{(\alpha)}}{\varepsilon}\right) \widehat{C}(l-t)\overline{\widehat{C}}(t)\mathrm{d}t\,\mathrm{d}l \right|^2 \mathrm{d}k$$

$$+ \int_{\mathbb{B}}\left| \int_{2\mathbb{B}}\int_{\mathbb{B}} \widehat{C}(k-l)\widehat{C}(l-t)\overline{\widehat{C}}(t)\mathrm{d}t\,\mathrm{d}l \right|^2 \mathrm{d}k \Bigg]^{\frac{1}{2}}$$

$$\leqslant c\left[\varepsilon^4 \sum_{\alpha,\beta,\gamma=1}^{N} \|\widehat{B}_\alpha *_{\varepsilon^{-1}\mathbb{B}} \widehat{B}_\beta *_{\varepsilon^{-1}\mathbb{B}} \overline{\widehat{B}}_\gamma\|^2_{L^2(\varepsilon^{-1}\mathbb{B})} + \varepsilon^4 \right.$$

$$\times \sum_{\alpha,\beta=1}^{N} \int_{\mathbb{B}}\left| \int_{2\varepsilon^{-1}\mathbb{B}}\int_{\varepsilon^{-1}\mathbb{B}} \widehat{B}_\alpha\left(\frac{k-k^{(\alpha)}-k^{(\beta)}}{\varepsilon}-l'\right) \right.$$

$$\times \widehat{B}_\beta\left(l'-t'\right) \overline{\widehat{C}}(\varepsilon t')\mathrm{d}t'\,\mathrm{d}l' \Bigg|^2 \mathrm{d}k + \varepsilon^6\sum_{\alpha=1}^{N} \int_{\mathbb{B}}\left| \int_{2\varepsilon^{-1}\mathbb{B}}\int_{\varepsilon^{-1}\mathbb{B}} \right.$$

$$\times \widehat{B}_\alpha\left(\frac{k-k^{(\alpha)}}{\varepsilon}-l'\right) \widehat{C}(\varepsilon(l'-t'))\overline{\widehat{C}}(\varepsilon t')\mathrm{d}t'\,\mathrm{d}l' \Bigg|^2 \mathrm{d}k$$

$$+ \int_{\mathbb{B}}\left| \int_{2\mathbb{B}}\int_{\mathbb{B}} \widehat{C}(k-l)\widehat{C}(l-t)\overline{\widehat{C}}(t)\mathrm{d}t\,\mathrm{d}l \right|^2 \mathrm{d}k \Bigg]^{\frac{1}{2}},$$

where we have used again transformations of the type $t' = \varepsilon^{-1}(t+k^{(\gamma)})$ and the fact that the convolution domains can be shifted by arbitrary $k_* \in \mathbb{R}^2$.

Next, for $j \in \{1,\ldots,N\}$ and $m \in \mathbb{N}$ we introduce notation for $\widehat{B}$ and $\widehat{C}$ restricted to m^2 periodicity cells. In detail, let

$$\widehat{B}_j^{(m)} := \chi_{m\varepsilon^{-1}\mathbb{B}} \widehat{B}_j, \qquad \widehat{C}^{(m)} := \chi_{m\mathbb{B}} \widehat{C}.$$

Note that $\widehat{B}_j^{(1)} = \widehat{B}_j^*$ and $\widehat{C}^{(1)} = \widehat{C}^*$. With this notation we have

$$\|\widetilde{F}(\widetilde{v})\|_{\mathcal{X}_2} \leqslant c \left[\varepsilon^4 \sum_{\alpha,\beta,\gamma=1}^{N} \|\widehat{B}_\alpha^{(3)} * \widehat{B}_\beta^{(2)} * \widehat{\overline{B}}_\gamma^*\|_{L^2(\mathbb{R}^2)}^2 + \varepsilon^6 \sum_{\alpha,\beta=1}^{N} \|\widehat{B}_\alpha^{(3)} * \widehat{B}_\beta^{(2)} * \widehat{\overline{C}}^*(\varepsilon\cdot)\|_{L^2(\mathbb{R}^2)}^2 \right.$$

$$\left. + \varepsilon^8 \sum_{\alpha=1}^{N} \|\widehat{B}_\alpha^{(3)} * \widehat{C}^{(2)}(\varepsilon\cdot)*\widehat{\overline{C}}^*(\varepsilon\cdot)\|_{L^2(\mathbb{R}^2)}^2 + \|\widehat{C}^{(3)} * \widehat{C}^{(2)} * \widehat{\overline{C}}^*\|_{L^2(\mathbb{R}^2)}^2 \right]^{\frac{1}{2}},$$

where $*$ denotes the convolution over the full $\mathbb{R}^2$. Using Young's inequality for convolutions and the fact that (due to the periodicity) $\|\widehat{B}_j^{(m)}\|_{L^p(\mathbb{C}^2)} \leqslant c \|\widehat{B}_j^*\|_{L^p(\mathbb{C}^2)}$ and $\|\widehat{C}^{(m)}\|_{L^p(\mathbb{C}^2)} \leqslant c \|\widehat{C}^*\|_{L^p(\mathbb{B})}$ for all $m \in \mathbb{N}$, $p \geqslant 1$, and $j = 1, \ldots, N$, we estimate

$$\|\widehat{B}_\alpha^{(3)} * \widehat{B}_\beta^{(2)} * \widehat{\overline{B}}_\gamma^*\|_{L^2(\mathbb{R}^2)} \leqslant c \|\widehat{B}_\alpha^*\|_{L^2(\mathbb{R}^2)} \|\widehat{B}_\beta^*\|_{L^1(\mathbb{R}^2)} \|\widehat{B}_\gamma^*\|_{L^1(\mathbb{R}^2)},$$

$$\|\widehat{B}_\alpha^{(3)} * \widehat{B}_\beta^{(2)} * \widehat{\overline{C}}^*(\varepsilon\cdot)\|_{L^2} \leqslant c \|\widehat{B}_\alpha^*\|_{L^1(\mathbb{R}^2)} \|\widehat{B}_\beta^*\|_{L^1(\mathbb{R}^2)} \|\widehat{C}^*(\varepsilon\cdot)\|_{L^2(\mathbb{R}^2)},$$

$$\|\widehat{B}_\alpha^{(3)} * \widehat{C}^{(2)}(\varepsilon\cdot) * \widehat{\overline{C}}^*(\varepsilon\cdot)\|_{L^2(\mathbb{R}^2)} \leqslant c \|\widehat{B}_\alpha^*\|_{L^1(\mathbb{R}^2)} \|\widehat{C}^*(\varepsilon\cdot)\|_{L^1(\mathbb{R}^2)} \|\widehat{C}^*(\varepsilon\cdot)\|_{L^2(\mathbb{R}^2)},$$

$$\|\widehat{C}^{(3)} * \widehat{C}^{(2)} * \widehat{\overline{C}}^*\|_{L^2(\mathbb{R}^2)} \leqslant c \|\widehat{C}^*\|_{L^2(\mathbb{B})} \|\widehat{C}^*\|_{L^1(\mathbb{B})}^2.$$

Finally, we arrive at (4.19) by using $\|\widehat{\overline{C}}^*\|_{L^1(\mathbb{B})} \leqslant c \|\widehat{\overline{C}}^*\|_{L^2(\mathbb{B})}$,

$$\|\widehat{C}^*(\varepsilon\cdot)\|_{L^2(\mathbb{R}^2)} = \varepsilon^{-1}\|\widehat{C}\|_{L^2(\mathbb{B})}, \qquad \|\widehat{C}^*(\varepsilon\cdot)\|_{L^1(\mathbb{R}^2)} = \varepsilon^{-2}\|\widehat{C}\|_{L^1(\mathbb{B})},$$

and

$$\|\widehat{B}_j^*\|_{L^1(\mathbb{R}^2)} \leqslant c \|\widehat{B}_j^*\|_{L_{s_B}^2(\mathbb{R}^2)} \quad \text{if } s_B > 1. \tag{4.22}$$

The last inequality follows from

$$\|\widehat{B}_j^*\|_{L^1(\mathbb{R}^2)} = \int_{\mathbb{R}^2} (1+|t|)^{-s_B} (1+|t|)^{s_B} |\widehat{B}_j^*(t)| \mathrm{d}t$$

$$\leqslant \left(\int_{\mathbb{R}^2} (1+|t|)^{-2s_B} \mathrm{d}t \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^2} (1+|t|)^{2s_B} |\widehat{B}_j^*(t)|^2 \mathrm{d}t \right)^{\frac{1}{2}},$$

where the first factor is bounded provided $s_B > 1$.

Estimate (4.21) is clearly the first part of (4.19). □

For the whole analysis of the components $\widetilde{w}_0$ and $\widetilde{w}_R$ we assume that for all sufficiently small $\varepsilon > 0$ and some $s_B > 1$ we have

$$\|\widehat{C}^*\|_{L^2} = \|\widehat{C}\|_{L^2(\mathbb{B})} \leqslant c_0 \varepsilon^{2-2r}$$

and

$$\|\widehat{B}_j^*\|_{L_{s_B}^2} = \|\widehat{B}_j\|_{L_{s_B}^2(\varepsilon^{-1}\mathbb{B})} \leqslant c \quad \text{for all } j \in \{1, \ldots, N\} \tag{4.23}$$

for some constants $c_0, c > 0$. The regularity parameter s_B is chosen below in order for the required estimates to work. Notice that under such an assumption the terms involving $\widehat{C}$ in the estimate (4.19) are $o(\varepsilon^2)$, provided $r \in \left(0, \frac{1}{2}\right]$.

4.3. Component $\widetilde{w}_0$

Under assumption (4.23) with $s_B \geqslant 2$ we solve the (linear) equation (4.10) for $\widetilde{w}_0$ and derive an $\mathcal{X}_2$-estimate on $\widetilde{w}_0$.

Lemma 4.7. *Let $s_B \geqslant 2$ and assume* (4.23), *where $\widehat{B}_j$ and $\widehat{C}$ have the supports as in* (4.13) *with $r \in (0, \frac{2}{3}]$. Then* (4.10) *has a unique solution $\widetilde{w}_0 \in \mathcal{X}_2$ with*

$$\|\widetilde{w}_0\|_{\mathcal{X}_2} \leqslant c_1 \|\widetilde{F}(\widetilde{v})\|_{\mathcal{X}_2} \leqslant c_2 \varepsilon^2, \tag{4.24}$$

where $c_1, c_2 > 0$ and c_2 depends polynomially on $\|\widehat{B}_j^*\|_{L^2_{s_B}}, \|\widehat{C}^*\|_{L^2}$ for all $j \in \{1, \ldots, N\}$.

Proof. By lemma 4.6 and assumption (4.23), we get for the right-hand side of (4.10)

$$\|{}^\epsilon Q_k \widetilde{F}(\tilde{v})\|_{\mathcal{X}_2} \leqslant c \, \|\tilde{F}(\tilde{v})\|_{\mathcal{X}_2} \leqslant \tilde{c}\varepsilon^2$$

if $5 - 4r \geqslant 2$ and $6 - 6r \geqslant 2$, which hold for $r \in (0, \frac{2}{3}]$. Here the constant $\tilde{c}$ depends polynomially on $\|\widehat{B}_j^*\|_{L^2_{s_B}}$, $j \in \{1, \ldots, N\}$, and $\|\widehat{C}^*\|_{L^2}$. Lemma 4.2 produces a solution $\widetilde{w}_0(\cdot, k) \in Q_k^\epsilon H_\#(\mathrm{curl}^2, \mathcal{Q})$ with

$$\|\widetilde{w}_0\|_{L^2(\mathbb{B}, H_\#(\mathrm{curl}^2, \mathcal{Q}))} \leqslant c \, \|\widetilde{F}(\tilde{v})\|_{L^2(\mathbb{B}, L^2(\mathcal{Q}))} \leqslant c \, \|\widetilde{F}(\tilde{v})\|_{\mathcal{X}_2}. \tag{4.25}$$

For an $\mathcal{X}_1$-estimate we need to control also the divergence $\nabla'_k \cdot \widetilde{w}_0$. First note that

$${}^\epsilon Q_k \left(\epsilon\widetilde{w}_0\right)(\cdot, k) = \epsilon\widetilde{w}_0(\cdot, k) - \langle\epsilon\widetilde{w}_0(\cdot, k), p_{n_*}(\cdot, k)\rangle\epsilon p_{n_*}(\cdot, k).$$

Taking the divergence $\nabla'_k \cdot$ of equation (4.10) produces

$$\begin{aligned}
\nabla'_k \cdot \widetilde{w}_0 = \frac{1}{\epsilon} \Bigg(&-\nabla'_k \cdot {}^\epsilon Q_k \widetilde{F}(\tilde{v}) - \nabla'_k \epsilon \cdot \widetilde{w}_0 - \Bigg\langle \frac{1}{\omega^2}\nabla'_k \times \nabla'_k \times \widetilde{w}_0(\cdot, k) \\
&- \epsilon\widetilde{w}_0(\cdot, k), p_{n_*}(\cdot, k) \Bigg\rangle \nabla'_k \cdot \left(\epsilon p_{n_*}(\cdot, k)\right) \Bigg).
\end{aligned}$$

Due to (A5) and the regularity $p_{n_*}(\cdot, k) \in H^2(\mathcal{Q})$ for all $k \in \mathbb{B}$ given by lemma A.3, we have

$$\nabla'_k \cdot \tilde{w}_0(\cdot, k) \in L^2(\mathcal{Q}) \quad \text{and hence} \quad \tilde{w}_0(\cdot, k) \in H^1(\mathcal{Q}).$$

This allows us to estimate the $\mathcal{X}_1$-norm

$$\|\widetilde{w}_0\|_{\mathcal{X}_1} \leqslant c \left(\|\widetilde{w}_0\|_{L^2(\mathbb{B}, H_\#(\text{curl}, \mathcal{Q}))} + \|\nabla'_k \cdot \widetilde{w}_0\|_{\mathcal{X}_0} \right)$$

$$\leqslant c \left(\|\widetilde{F}(\widetilde{v})\|_{\mathcal{X}_2} + \|\nabla'_k \cdot {}^\epsilon Q_k \widetilde{F}(\widetilde{v})\|_{\mathcal{X}_0} + \left\| \left\langle \frac{1}{\omega^2} \nabla'_k \times \nabla'_k \times \widetilde{w}_0(\cdot, k) \right. \right. \right.$$

$$\left. \left. \left. - \epsilon \widetilde{w}_0(\cdot, k), p_{n_*}(\cdot, k) \right\rangle \right\|_{L^2(\mathbb{B})} + \|\widetilde{w}_0\|_{\mathcal{X}_0} \right)$$

$$\leqslant c \left(\|\widetilde{F}(\widetilde{v})\|_{\mathcal{X}_2} + \|\nabla'_k \cdot \widetilde{F}(\widetilde{v})\|_{\mathcal{X}_0} + \|\widetilde{F}(\widetilde{v})\|_{\mathcal{X}_0} \right.$$

$$\left. + \|\widetilde{w}_0\|_{L^2(\mathbb{B}, H_\#(\text{curl}^2, \mathcal{Q}))} + \|\widetilde{w}_0\|_{\mathcal{X}_0} \right)$$

$$\leqslant c \|\widetilde{F}(\widetilde{v})\|_{\mathcal{X}_2},$$

where in the second and the last step we used (4.25). The $\mathcal{X}_2$-estimate is analogous:

$$\|\widetilde{w}_0\|_{\mathcal{X}_2} \leqslant c \left(\|\widetilde{w}_0\|_{\mathcal{X}_1} + \|\nabla'_k \times \widetilde{w}_0\|_{\mathcal{X}_1} + \|\nabla'_k \cdot \widetilde{w}_0\|_{\mathcal{X}_1} \right)$$

$$\leqslant c \left(\|\widetilde{F}(\widetilde{v})\|_{\mathcal{X}_2} + \|\nabla'_k \times \widetilde{w}_0\|_{L^2(\mathbb{B}, L^2(\mathcal{Q}))} \right.$$

$$\left. + \|\nabla'_k \times \nabla'_k \times \widetilde{w}_0\|_{L^2(\mathbb{B}, L^2(\mathcal{Q}))} + \|\nabla'_k \cdot \widetilde{w}_0\|_{\mathcal{X}_1} \right)$$

$$\leqslant c \|\widetilde{F}(\widetilde{v})\|_{\mathcal{X}_2}.$$

Note that in the second inequality the divergence $\nabla'_k \cdot$ of $\nabla'_k \times \widetilde{w}_0$ vanishes and in the last inequality the estimate of $\|\nabla'_k \cdot \widetilde{w}_0\|_{\mathcal{X}_1}$ is analogous to the $\mathcal{X}_1$ estimate above. $\qquad \square$

4.4. Component $\widetilde{w}_R$

Next, we keep assumption (4.23) and solve equation (4.11) for $\widetilde{w}_R$ via a Banach fixed point argument with $\widetilde{v}$ satisfying (4.8) and $\widetilde{w}_0$ as just obtained in section 4.3. We show that for $r \in \left(0, \frac{1}{2}\right]$ a solution of order $O(\varepsilon^3)$ (in the $\mathcal{X}_2$-norm) exists. We write

$$\widetilde{w}_R = \left({}^\epsilon Q_k L_k Q_k^\epsilon \right)^{-1} \omega^2 \, {}^\epsilon Q_k \left(\tilde{F}(\widetilde{v} + \widetilde{w}_0 + \widetilde{w}_R) - \tilde{F}(\widetilde{v}) \right) =: G(\widetilde{w}_R) \qquad (4.26)$$

and we aim to show the contraction property of the map G in the ball

$$\mathcal{B}^{\mathcal{X}_2}_{K\varepsilon^\eta} := \{ f \in \mathcal{X}_2 \mid \|f\|_{\mathcal{X}_2} < K\varepsilon^\eta \}$$

for suitable values of $K, \eta > 0$. Applying the algebra property (2.27) of $\mathcal{X}_2$, we get

$$\|G(\widetilde{w}_R)\|_{\mathcal{X}_2} \leqslant c \left[\sum_{a,b \in \{1,2,3\}} \left(\|\widetilde{v}_a *_{\mathbb{B}} \widetilde{v}_b\|_{\mathcal{X}_2} + \|\widetilde{v}_a *_{\mathbb{B}} \overline{\widetilde{v}}_b\|_{\mathcal{X}_2} \right) \right.$$

$$\times \left(\|\widetilde{w}_0\|_{\mathcal{X}_2} + \|\widetilde{w}_R\|_{\mathcal{X}_2} \right) + \|\widetilde{w}_0\|^2_{\mathcal{X}_2} \left(\|\widetilde{v}\|_{\mathcal{X}_2} + \|\widetilde{w}_R\|_{\mathcal{X}_2} \right)$$

$$\left. + \|\widetilde{w}_R\|^2_{\mathcal{X}_2} \left(\|\widetilde{v}\|_{\mathcal{X}_2} + \|\widetilde{w}_0\|_{\mathcal{X}_2} \right) + \|\widetilde{v}\|_{\mathcal{X}_2} \|\widetilde{w}_0\|_{\mathcal{X}_2} \|\widetilde{w}_R\|_{\mathcal{X}_2} + \|\widetilde{w}_0\|^3_{\mathcal{X}_2} + \|\widetilde{w}_R\|^3_{\mathcal{X}_2} \right].$$

Recalling lemmas 4.4 and 4.5 we first have for any $a, b = 1, 2, 3$

$$\|\widetilde{v}_a *_{\mathbb{B}} \widetilde{v}_b\|_{\mathcal{X}_2} \leqslant \|\widetilde{v}_{B,a} *_{\mathbb{B}} \widetilde{v}_{B,b}\|_{\mathcal{X}_2} + \|\widetilde{v}_{B,a}\|_{\mathcal{X}_2}\|\widetilde{v}_{C,b}\|_{\mathcal{X}_2} + \|\widetilde{v}_{C,a}\|_{\mathcal{X}_2}\|\widetilde{v}_{B,b}\|_{\mathcal{X}_2} + \|\widetilde{v}_{C,a}\|_{\mathcal{X}_2}\|\widetilde{v}_{C,b}\|_{\mathcal{X}_2}$$

$$\leqslant c\varepsilon \sum_{i,j=1}^{N} \|\widehat{B}_i^*\|_{L_{s_B}^2}\|\widehat{B}_j^*\|_{L_{s_B}^2} + c\sum_{i=1}^{N} \|\widehat{B}_i^*\|_{L_{s_B}^2}\|\widehat{C}\|_{L^2(\mathbb{B})} + c\|\widehat{C}\|_{L^2(\mathbb{B})}^2$$

$$\leqslant c\left(\varepsilon + \varepsilon^{2-2r} + \varepsilon^{4-4r}\right).$$

Therefore, recalling (4.24), we obtain for $\widetilde{w}_R \in \mathcal{B}_{K\varepsilon^{\eta}}^{\mathcal{X}_2}$

$$\|G(\widetilde{w}_R)\|_{\mathcal{X}_2} \leqslant \tilde{c}\left[(\varepsilon + \varepsilon^{2-2r} + \varepsilon^{4-4r})(\varepsilon^2 + K\varepsilon^{\eta}) + \varepsilon^4(1 + K\varepsilon^{\eta})\right.$$
$$\left. + K^2\varepsilon^{2\eta}(1 + \varepsilon^2) + K\varepsilon^{2+\eta} + \varepsilon^6 + K^3\varepsilon^{3\eta}\right]$$
$$\leqslant (\tilde{c} + 1)\left(\varepsilon^3 + K\varepsilon^{1+\eta} + K^2\varepsilon^{2\eta}\right),$$

where the second inequality holds for all $r \in \left(0, \frac{1}{2}\right]$ and $\tilde{c}$ is a constant depending just on the norms $\|\widehat{B}_j^*\|_{L_{s_B}^2}, j \in \{1, \ldots, N\}$, and c_0. Choosing $\eta = 3$ and $K = \tilde{c} + 2$, then $G : \mathcal{B}_{K\varepsilon^{\eta}}^{\mathcal{X}_2} \to \mathcal{B}_{K\varepsilon^{\eta}}^{\mathcal{X}_2}$, i.e.

$$\|G(\widetilde{w}_R)\|_{\mathcal{X}_2} \leqslant K\varepsilon^3. \tag{4.27}$$

Next, we address the contraction property. For $\widetilde{u}^{(j)} := \widetilde{v} + \widetilde{w}_0 + \widetilde{w}_R^{(j)}, \ j \in \{1, 2\}$ we have

$$\|G(\widetilde{w}_R^{(1)}) - G(\widetilde{w}_R^{(2)})\|_{\mathcal{X}_2}$$

$$\leqslant c\left[\max_{a,b\in\{1,2,3\}}\left(\|\widetilde{v}_a *_{\mathbb{B}} \widetilde{v}_b\|_{\mathcal{X}_2} + \|\widetilde{v}_a *_{\mathbb{B}} \overline{\widetilde{v}}_b\|_{\mathcal{X}_2}\right) + \|w_0\|_{\mathcal{X}_2}^2 + \|\widetilde{w}_R^{(1)}\|_{\mathcal{X}_2}^2\right.$$

$$\left. + \|\widetilde{w}_R^{(2)}\|_{\mathcal{X}_2}^2 + \|\widetilde{v}\|_{\mathcal{X}_2}\|\widetilde{w}_R^{(1)}\|_{\mathcal{X}_2} + \|\widetilde{v}\|_{\mathcal{X}_2}\|\widetilde{w}_R^{(2)}\|_{\mathcal{X}_2}\right] \|\widetilde{w}_R^{(1)} - \widetilde{w}_R^{(2)}\|_{\mathcal{X}_2}$$

$$\leqslant c\varepsilon\|\widetilde{w}_R^{(1)} - \widetilde{w}_R^{(1)}\|_{\mathcal{X}_2}. \tag{4.28}$$

The contraction thus follows provided ε is small enough. By the Banach fixed point theorem there exists a unique solution to equation (4.26) for $\widetilde{w}_R$ which satisfies the estimate

$$\|\widetilde{w}_R\|_{\mathcal{X}_2} \leqslant K\varepsilon^3. \tag{4.29}$$

For later use, we need also to show the Lipschitz dependence of $\widetilde{w}$ on $\widetilde{v}_B$ and $\widetilde{v}_C$.

Lemma 4.8. *The map $\mathcal{B}_{\rho}^{\mathcal{X}_2}(0) \times \mathcal{B}_{\rho\varepsilon^{2-2r}}^{\mathcal{X}_2}(0) \ni (\widetilde{v}_B, \widetilde{v}_C) \mapsto \widetilde{w}(\widetilde{v}_B, \widetilde{v}_C) \in \mathcal{B}_{\rho\varepsilon^2}^{\mathcal{X}_2}(0)$ is Lipschitz-continuous for any $\rho > 0$ and $\varepsilon > 0$ small enough. The Lipschitz constant C_L satisfies $C_L = \mathcal{O}(\varepsilon)$ as $\varepsilon \to 0$.*

Proof. Let $\widetilde{v}^{(1)} = (\widetilde{v}_B^{(1)}, \widetilde{v}_C^{(1)}), \ \widetilde{v}^{(2)} = (\widetilde{v}_B^{(2)}, \widetilde{v}_C^{(2)}) \in \mathcal{B}_{\rho}^{\mathcal{X}_2}(0) \times \mathcal{B}_{\rho\varepsilon^{2-2r}}^{\mathcal{X}_2}(0)$ and for $i \in \{1, 2\}$ define $\widetilde{w}_0^{(i)} = \widetilde{w}_0(\widetilde{v}^{(i)})$ as solutions of (4.10) and $\widetilde{w}_R^{(i)} = \widetilde{w}_R(\widetilde{v}^{(i)})$ as solutions of (4.11) with $\widetilde{w}_0$ replaced by $\widetilde{w}_0^{(i)}$. Such functions are well-defined since $\widetilde{v}^{(i)} \in \mathcal{B}_{\rho}^{\mathcal{X}_2}(0) \times \mathcal{B}_{\rho\varepsilon^{2-2r}}^{\mathcal{X}_2}(0)$ implies that the respective coefficients $\widehat{B}_j^{(i)}, \widehat{C}^{(i)}$ fulfill assumption (4.23). Since $\|\widetilde{w}_0^{(1)} - \widetilde{w}_0^{(2)}\|_{\mathcal{X}_2} \leqslant c_1\|\widetilde{F}(\widetilde{v}^{(1)}) - \widetilde{F}(\widetilde{v}^{(2)})\|_{\mathcal{X}_2}$, we obtain similarly to (4.28)

$$\|\widetilde{w}_0^{(1)} - \widetilde{w}_0^{(2)}\|_{\mathcal{X}_2} \leqslant c\varepsilon\|\widetilde{v}^{(1)} - \widetilde{v}^{(2)}\|_{\mathcal{X}_2}.$$

For $\widetilde{w}_R^{(i)}$ we have

$$
{}^{\epsilon}Q_k L_k Q_k^{\epsilon} \tilde{w}_R^{(i)} = \omega^2 \, {}^{\epsilon}Q_k \left(\tilde{F}(\tilde{v}^{(i)} + \tilde{w}_0^{(i)} + \tilde{w}_R^{(i)}) - \tilde{F}(\tilde{v}^{(i)}) \right)
$$

and again with analogous computations as in (4.28) we then get

$$
\|\widetilde{w}_R^{(1)} - \widetilde{w}_R^{(2)}\|_{\mathcal{X}_2} \leqslant c\varepsilon \left(\|\widetilde{v}^{(1)} - \widetilde{v}^{(2)}\|_{\mathcal{X}_2} + \|\widetilde{w}_R^{(1)} - \widetilde{w}_R^{(2)}\|_{\mathcal{X}_2} \right),
$$

which leads to the desired estimate. □

4.5. Component $\widehat{C}$

So far, we have completed the first two steps of the initial program: under assumption (4.23) we inferred the existence of a small solution $\widetilde{w}$ of (4.9). Now we have to deal with the component $\widetilde{v}$, i.e. the projection of $\widetilde{u}$ onto the mode p_{n_*}. Recall that our aim is to find solutions $\widetilde{v}$ with $\widehat{B}_j^*$ close to the coefficients $\widehat{A}_j$ from the ansatz (4.12) and with $\widehat{C}$ small.

In this section we assume $r \in \left(0, \frac{1}{2}\right]$, choose an arbitrary $(\widehat{B}_j)_{j=1}^N$ with $\|\widehat{B}_j^*\|_{L_{s_B}^2(\mathbb{R}^2)} \leqslant c$ for all $j \in \{1, \dots, N\}$, and seek a small $\widehat{C}$. Recall that $\widetilde{u} = \widetilde{v} + \widetilde{w} = \widetilde{v}_B + \widetilde{v}_C + \widetilde{w}$, where $\widetilde{v}_B$ is now fixed and $\widetilde{w}$ is determined by sections 4.3 and 4.4. Hence, we write $\widetilde{w} = \widetilde{w}(\widetilde{v}_C)$ suppressing the dependence on $\widetilde{v}_B$.

Since the support of $\widehat{C}^*$ within the Brillouin zone $\mathbb{B}$ is in $\mathbb{B} \backslash \cup_{j=1}^N B_{\varepsilon^r}(k^{(j)})$, we introduce the characteristic function

$$
\chi^C(k) := 1 - \sum_{j=1}^N \chi_{B_{\varepsilon^r}(k^{(j)})}(k).
$$

The equation for $\widehat{C}$ then reads

$$
\begin{aligned}
\widehat{C}(k) = {}&(\omega_{n_*}(k)^2 - \omega^2)^{-1} \omega^2 \chi^C(k) \langle \widehat{F}(\widetilde{v}_B)(\cdot, k), p_{n_*}(\cdot, k) \rangle \\
&+ (\omega_{n_*}(k)^2 - \omega^2)^{-1} \omega^2 \chi^C(k) \\
&\times \left\langle \left(\widehat{F}(\widetilde{v}_B + \widetilde{v}_C + \widetilde{w}(\widetilde{v}_C)) - \widehat{F}(\widetilde{v}_B) \right)(\cdot, k), p_{n_*}(\cdot, k) \right\rangle \\
=:{}&T_1(k) + T_2(k).
\end{aligned} \tag{4.30}
$$

In order to enjoy the algebra property of $\mathcal{X}_s, s > 1$, we multiply both sides of (4.30) by $p_{n_*}(\cdot, k)$ and apply a fixed point approach to the resulting equation for $\widetilde{v}_C$ in a small ball in $\mathcal{X}_2$.

$$
\widetilde{v}_C(x, k) = T_1(k) p_{n_*}(x, k) + T_2(k) p_{n_*}(x, k) =: \mathcal{H}(\widetilde{v}_C)(x, k), \tag{4.31}
$$

which we aim to solve in the ball

$$
\mathcal{B}_{K\varepsilon^{2-2r}}^{\mathcal{X}_2} := \{ f \in \mathcal{X}_2 \mid \|f\|_{\mathcal{X}_2} \leqslant K\varepsilon^{2-2r} \}
$$

for some $K > 0$ and all $\varepsilon > 0$ small enough.

We have isolated the leading order part $\widetilde{F}(\widetilde{v}_B)$ of the nonlinearity in the term T_1. Note that $\widetilde{F}(\widetilde{v}_B)$ is concentrated only near a finite number of k-points, namely

$$\operatorname{supp}(\widetilde{F}(\widetilde{v}_B)(x,\cdot)) \cap \mathbb{B} \subset B_{3\varepsilon^r}(S),$$

where

$$S := \left\{ k \in \mathbb{B} : k = k^{(\alpha)} + k^{(\beta)} - k^{(\gamma)} + K \text{ with } \alpha, \beta, \gamma \in \{1, \ldots, N\}, \ K \in \Lambda^* \right\}.$$

We write

$$S = S_c \cup S_{nc}, \quad \text{where} \quad S_c := \{k^{(1)}, \ldots, k^{(N)}\}.$$

S_c is the critical set as it lies in the level set W_{ω_*}. Note that $\chi^C(k) = 0$ for each $k \in B_{\varepsilon^r}(S_c)$. We split T_1 accordingly

$$T_1 = T_{1,c} + T_{1,nc}, \quad \text{where} \quad \operatorname{supp}(T_{1,c}) \subset B_{3\varepsilon^r}(S_c) \backslash B_{\varepsilon^r}(S_c) \text{ and } \operatorname{supp}(T_{1,nc}) \subset B_{3\varepsilon^r}(S_{nc}).$$

We estimate these components separately.

First, we estimate the factor $(\omega^2 - \omega_{n_*}(k)^2)^{-1}$. Due to assumption (A3) we have

$$|(\omega^2 - \omega_{n_*}(k)^2)^{-1}| \leqslant c \quad \text{for } k \in \operatorname{supp}(T_{1,nc}). \tag{4.32}$$

On $\operatorname{supp}(T_{1,c})$ we use the locally quadratic nature of $\omega_{n_*}(k)$ near $k = k^{(j)}$, $j = 1, \ldots, N$. Indeed, as ω_{n_*} has an extremum at each $k^{(j)}$, we have

$$\nabla\left(\omega_{n_*}(k)^2\right)\big|_{k=k^{(j)}} = 2\omega_* \nabla \omega_{n_*}(k^{(j)}) = 0.$$

Moreover,

$$\nabla^2 \left(\omega_{n_*}(k)^2\right)\big|_{k=k^{(j)}} = 2\omega_* \nabla^2 \omega_{n_*}(k^{(j)}) + \mathcal{K}(\omega_{n_*})(k^{(j)}),$$

where

$$\mathcal{K}(\omega_{n_*}) := \begin{pmatrix} \left(\partial_1 \omega_{n_*}\right)^2 & \left(\partial_1 \omega_{n_*}\right)\left(\partial_2 \omega_{n_*}\right) \\ \left(\partial_1 \omega_{n_*}\right)\left(\partial_2 \omega_{n_*}\right) & \left(\partial_2 \omega_{n_*}\right)^2 \end{pmatrix},$$

the determinant of which evidently vanishes. Using (A4), we deduce then that the Hessian of $k \mapsto \left(\omega_{n_*}(k)\right)^2$ is definite at $k^{(j)}$. This in turns implies that

$$|\omega_{n_*}(k)^2 - \omega^2|^{-1} \leqslant c\varepsilon^{-2r} \quad \text{for } k \in \operatorname{supp}(T_{1,c}) \tag{4.33}$$

for ε small enough. It is mainly here where assumption (A3) is used. If $\{k^{(1)}, \ldots, k^{(N)}\}$ was a proper subset of the level set W_{ω_*}, then $\operatorname{supp}(T_{1,c})$ would intersect W_{ω_*} and $|\omega_{n_*}(k)^2 - \omega^2|^{-1}$ would blow up on $\operatorname{supp}(T_{1,c})$.

The estimate of $T_{1,nc}$ follows directly from (4.32) and (4.21). We get

$$\|T_{1,nc} p_{n_*}\|_{\mathcal{X}_2} \leqslant c\varepsilon^2 \sum_{\alpha,\beta,\gamma=1}^N \|\widehat{B}_\alpha^*\|_{L_{s_B}^2} \|\widehat{B}_\beta^*\|_{L_{s_B}^2} \|\widehat{B}_\gamma^*\|_{L_{s_B}^2} \leqslant c\varepsilon^2 \tag{4.34}$$

using (4.23).

For $T_{1,c}$ the estimate (4.33) causes loss of powers of ε but we gain some powers by assuming a fast decay of $\widehat{B}_j$. In detail, define the weight $h_j(k) := \left(1 + \frac{|k - k^{(j)}|}{\varepsilon}\right)^{s_B}$. Then analogously to the proof of (4.21) in lemma 4.6 we have, due to (4.33),

$$
\|T_{1,c} p_{n_*}\|_{\mathcal{X}_2} \leqslant c\varepsilon^{2-2r} \sum_{j=1}^{N} \sum_{\alpha,\beta,\gamma=1}^{N} \sup_{k \in B_{\varepsilon^r}(k^{(j)})} h_j(k)^{-1} \|h_j(k^{(j)} + \varepsilon \cdot) \widehat{B}_\alpha^* * \widehat{B}_\beta^* * \overline{\widehat{B}}_\gamma^*\|_{L^2}
$$

$$
\leqslant c\varepsilon^{2-2r+(1-r)s_B} \sum_{\alpha,\beta,\gamma=1}^{N} \|\widehat{B}_\alpha^*\|_{L_{s_B}^2} \|\widehat{B}_\beta^*\|_{L_{s_B}^2} \|\widehat{B}_\gamma^*\|_{L_{s_B}^2}
$$

$$
\leqslant c\varepsilon^{2-2r+(1-r)s_B}. \tag{4.35}
$$

The estimate of the term T_2 is more delicate as T_2 is nonlinear in $\widetilde{v}_C$. Since $\|\widetilde{v}_B\|_{\mathcal{X}_2} \leqslant c$ by assumption, we get $\widetilde{w} \in \mathcal{B}_{K\varepsilon^2}^{\mathcal{X}_2}$ as a solution of (4.9) with K dependent on $\left(\|\widehat{B}_j^*\|_{L_{s_B}^2}\right)_{j=1}^N$, provided $\|\tilde{v}_C\|_{\mathcal{X}_2} \leqslant K\varepsilon^{2-2r}$. We show now that such $\tilde{v}_C$ exists. First, similarly to the map G in section 4.4 and taking into account (4.33),

$$
\|T_2 p_{n_*}\|_{\mathcal{X}_2} \leqslant c\varepsilon^{-2r} \|\widetilde{F}(\widetilde{u}) - \widetilde{F}(\widetilde{v}_B)\|_{\mathcal{X}_2} \leqslant c\varepsilon^{-2r} \Bigg[\|\widetilde{v}_B\|_{\mathcal{X}_2} \|\widetilde{v}_C\|_{\mathcal{X}_2} \|\widetilde{w}\|_{\mathcal{X}_2}
$$

$$
+ \|\widetilde{w}\|_{\mathcal{X}_2}^3 + \|\widetilde{w}\|_{\mathcal{X}_2}^2 (\|\widetilde{v}_B\|_{\mathcal{X}_2} + \|\widetilde{v}_C\|_{\mathcal{X}_2}) + \|\widetilde{v}_C\|_{\mathcal{X}_2}^2 \|\widetilde{w}\|_{\mathcal{X}_2}
$$

$$
+ \sum_{a,b \in \{1,2,3\}} \left(\|\widetilde{v}_{B,a} *_{\mathbb{B}} \widetilde{v}_{B,b}\|_{\mathcal{X}_2} + \|\widetilde{v}_{B,a} *_{\mathbb{B}} \overline{\widetilde{v}}_{B,b}\|_{\mathcal{X}_2} \right) \|\widetilde{w}\|_{\mathcal{X}_2}
$$

$$
+ \sum_{a,b,c \in \{1,2,3\}} \sum_{\substack{\mu_1,\mu_2,\mu_3 \in \{B,C\} \\ (\mu_1,\mu_2,\mu_3) \neq (B,B,B)}} \|\widetilde{v}_{\mu_1,a} *_{\mathbb{B}} \widetilde{v}_{\mu_2,b} *_{\mathbb{B}} \overline{\widetilde{v}}_{\mu_3,c}\|_{\mathcal{X}_2} \Bigg]. \tag{4.36}
$$

We may apply the algebra property of $\mathcal{X}_2$ and lemma 4.5 to treat the convolution terms. For $\widetilde{v}_C \in \mathcal{B}_{K\varepsilon^{2-2r}}^{\mathcal{X}_2}$ we obtain

$$
\|\widetilde{v}_{B_i,a} *_{\mathbb{B}} \widetilde{v}_{B_j,b} *_{\mathbb{B}} \overline{\widetilde{v}}_{C_m,c}\|_{\mathcal{X}_2} \leqslant \|\widetilde{v}_{B_i,a} *_{\mathbb{B}} \widetilde{v}_{B_j,b}\|_{\mathcal{X}_2} \|\widetilde{v}_{C_m,c}\|_{\mathcal{X}_2} \leqslant c\varepsilon^{3-2r},
$$

$$
\|\widetilde{v}_{B_i,a} *_{\mathbb{B}} \widetilde{v}_{C_j,b} *_{\mathbb{B}} \overline{\widetilde{v}}_{C_m,c}\|_{\mathcal{X}_2} \leqslant \|\widetilde{v}_{B_i,a}\|_{\mathcal{X}_2} \|\widetilde{v}_{C_j,b}\|_{\mathcal{X}_2} \|\widetilde{v}_{C_m,c}\|_{\mathcal{X}_2} \leqslant c\varepsilon^{4-4r},
$$

$$
\|\widetilde{v}_{C_i,a} *_{\mathbb{B}} \widetilde{v}_{C_j,b} *_{\mathbb{B}} \overline{\widetilde{v}}_{C_m,c}\|_{\mathcal{X}_2} \leqslant \|\widetilde{v}_{C_i,a}\|_{\mathcal{X}_2} \|\widetilde{v}_{C_j,b}\|_{\mathcal{X}_2} \|\widetilde{v}_{C_m,c}\|_{\mathcal{X}_2} \leqslant c\varepsilon^{6-6r}.
$$

Analogous estimates hold when the complex conjugation is moved onto another term. From (4.36) we then get

$$
\|T_2 p_{n_*}\|_{\mathcal{X}_2} \leqslant c\varepsilon^{-2r} \left(\varepsilon^{4-2r} + \varepsilon^6 + \varepsilon^4(1 + \varepsilon^{2-2r}) \right.
$$

$$
\left. + \varepsilon^{6-4r} + \varepsilon^3 + \varepsilon^{3-2r} + \varepsilon^{4-4r} + \varepsilon^{6-6r} \right)
$$

$$
\leqslant c\varepsilon^{3-4r}, \tag{4.37}
$$

since $r \in \left(0, \frac{1}{2}\right]$.

Combining then (4.34), (4.35), and (4.37), we obtain

$$\|\mathcal{H}(\widetilde{v}_C)\|_{\mathcal{X}_2} \leqslant c\left(\varepsilon^2 + \varepsilon^{(1-r)s_B+2-2r} + \varepsilon^{3-4r}\right).$$

Since for $r \in \left(0, \frac{1}{2}\right]$ all exponents are greater than or equal to $2 - 2r$, we get

$$\mathcal{H} : \mathcal{B}_{K\varepsilon^{2-2r}}^{\mathcal{X}_2} \to \mathcal{B}_{K\varepsilon^{2-2r}}^{\mathcal{X}_2}$$

with some $K = K\left(\|\widehat{B}_1^*\|_{L_{s_B}^2}, \ldots, \|\widehat{B}_N^*\|_{L_{s_B}^2}\right)$.

We address now the contraction property of the map $\mathcal{H}$. Take $\widetilde{v}_C^{(1)}, \widetilde{v}_C^{(2)} \in \mathcal{B}_{K\varepsilon^{2-2r}}^{\mathcal{X}_2}$ and consider $\widetilde{w}^{(i)} := \widetilde{w}(\widetilde{v}_C^{(i)})$ for $i \in \{1, 2\}$ as given by sections 4.3 and 4.4. We aim to estimate $\|\mathcal{H}(\widetilde{v}_C^{(1)}) - \mathcal{H}(\widetilde{v}_C^{(2)})\|_{\mathcal{X}_2}$. Clearly, T_1 is independent of $\widetilde{v}_C$ and similarly to (4.36) we infer

$$\left\| \left(T_2(\widetilde{v}_C^{(1)}) - T_2(\widetilde{v}_C^{(2)})\right) p_{n_*} \right\|_{\mathcal{X}_2}$$

$$\leqslant c\varepsilon^{-2r} \Big[\|\widetilde{v}_B\|_{\mathcal{X}_2} \|\widetilde{w}^{(1)} *_{\mathbb{B}} \widetilde{v}_C^{(1)} - \widetilde{w}^{(2)} *_{\mathbb{B}} \widetilde{v}_C^{(2)}\|_{\mathcal{X}_2}$$

$$+ \|\widetilde{w}^{(1)} *_{\mathbb{B}} \widetilde{w}^{(1)} *_{\mathbb{B}} \overline{\widetilde{v}}_C^{(1)} - \widetilde{w}^{(2)} *_{\mathbb{B}} \widetilde{w}^{(2)} *_{\mathbb{B}} \overline{\widetilde{v}}_C^{(2)}\|_{\mathcal{X}_2}$$

$$+ \|\widetilde{w}^{(1)} *_{\mathbb{B}} \widetilde{v}_C^{(1)} *_{\mathbb{B}} \overline{\widetilde{v}}_C^{(1)} - \widetilde{w}^{(2)} *_{\mathbb{B}} \widetilde{v}_C^{(2)} *_{\mathbb{B}} \overline{\widetilde{v}}_C^{(2)}\|_{\mathcal{X}_2}$$

$$+ \sum_{a,b\in\{1,2,3\}} \left(\|\widetilde{v}_{B,a} *_{\mathbb{B}} \widetilde{v}_{B,b}\|_{\mathcal{X}_2} + \|\widetilde{v}_{B,a} *_{\mathbb{B}} \overline{\widetilde{v}}_{B,b}\|_{\mathcal{X}_2} \right) \|\widetilde{v}_C^{(1)} - \widetilde{v}_C^{(2)}\|_{\mathcal{X}_2}$$

$$+ \|\widetilde{v}_B\|_{\mathcal{X}_2} \|\widetilde{v}_C^{(1)} *_{\mathbb{B}} \widetilde{v}_C^{(1)} - \widetilde{v}_C^{(2)} *_{\mathbb{B}} \widetilde{v}_C^{(2)}\|_{\mathcal{X}_2}$$

$$+ \|\widetilde{v}_C^{(1)} *_{\mathbb{B}} \widetilde{v}_C^{(1)} *_{\mathbb{B}} \overline{\widetilde{v}}_C^{(1)} - \widetilde{v}_C^{(2)} *_{\mathbb{B}} \widetilde{v}_C^{(2)} *_{\mathbb{B}} \overline{\widetilde{v}}_C^{(2)}\|_{\mathcal{X}_2} \Big].$$

All the terms are then estimated in a similar way, e.g.

$$\|\widetilde{w}^{(1)} *_{\mathbb{B}} \widetilde{v}_C^{(1)} - \widetilde{w}^{(2)} *_{\mathbb{B}} \widetilde{v}_C^{(2)}\|_{\mathcal{X}_2}$$

$$\leqslant \|\widetilde{w}^{(1)}\|_{\mathcal{X}_2} \|\widetilde{v}_C^{(1)} - \widetilde{v}_C^{(2)}\|_{\mathcal{X}_2} + \|\widetilde{v}_C^{(2)}\|_{\mathcal{X}_2} \|\widetilde{w}^{(1)} - \widetilde{w}^{(2)}\|_{\mathcal{X}_2}$$

$$\leqslant c\varepsilon^2 \|\widetilde{v}_C^{(1)} - \widetilde{v}_C^{(2)}\|_{\mathcal{X}_2} + K\varepsilon^{2-2r} \|\widetilde{w}^{(1)} - \widetilde{w}^{(2)}\|_{\mathcal{X}_2}$$

$$\leqslant c\left(\varepsilon^2 + \varepsilon^{3-2r}\right) \|\widetilde{v}_C^{(1)} - \widetilde{v}_C^{(2)}\|_{\mathcal{X}_2},$$

where we applied lemma 4.8 in the last step. Hence

$$\|\mathcal{H}(\widetilde{v}_C^{(1)}) - \mathcal{H}(\widetilde{v}_C^{(2)})\|_{\mathcal{X}_2} \leqslant c(\varepsilon^{2-2r} + \varepsilon^{3-4r})\|\widetilde{v}_C^{(1)} - \widetilde{v}_C^{(2)}\|_{\mathcal{X}_2},$$

i.e. a contraction for $\varepsilon > 0$ small enough due to $r \in (0, \frac{1}{2}]$.

As a result, we obtain a solution $\widetilde{v}_C \in \mathcal{B}_{K\varepsilon^{2-2r}}^{\mathcal{X}_2}$ of (4.31), where $K = K\left(\|\widehat{B}_1\|_{L_{s_B}^2}, \ldots, \|\widehat{B}_N\|_{L_{s_B}^2}\right)$, and in turn the estimates

$$\|\widetilde{v}_C\|_{\mathcal{X}_2} \leqslant c\varepsilon^{2-2r} \tag{4.38}$$

and

$$\|\widehat{C}\|_{L^2(\mathbb{B})} \leqslant c\varepsilon^{2-2r}. \tag{4.39}$$

Indeed, (4.39) follows from (4.38):

$$\|\widehat{C}\|_{L^2(\mathbb{B})} \leqslant \left(\min_{k \in \mathbb{B}} \|p_{n_*}(\cdot, k)\|_{H^2(\mathcal{Q})}\right)^{-1} \left(\int_{\mathbb{B}} \left|\widehat{C}(k)\right|^2 \|p_{n_*}(\cdot, k)\|_{H^2(\mathcal{Q})}^2 \mathrm{d}k\right)^{\frac{1}{2}}$$

$$\leqslant c \|\widetilde{v}_C\|_{\mathcal{X}_2} \leqslant c\varepsilon^{2-2r}.$$

4.6. Components $\widehat{B}_j$

We finally address the component $\widetilde{v}_B$ of the solution $\widetilde{u}$ and with $\widetilde{w}$ and $\widetilde{v}_C$ found above we solve for such $\widetilde{v}_B$, for which $\left(\widehat{B}_j\right)_{j=1}^N$ is close to the solutions $\left(\widehat{A}_j\right)_{j=1}^N$ of the CMEs (4.14). As a result the component $\widetilde{v}_B$ is the dominant part of the solution $\widetilde{u}$.

Equation (4.8) on the compact support of $\widehat{B}_j^*$ can be rewritten as

$$(\omega_{n_*}^2(k) - \omega^2)\frac{1}{\varepsilon}\widehat{B}_j^*\left(\frac{k - k^{(j)}}{\varepsilon}\right) = \omega^2 \langle \widetilde{F}(\widetilde{u})(\cdot, k), p_{n_*}(\cdot, k)\rangle, \quad k \in B_{\varepsilon^r}(k^{(j)})$$

or equivalently

$$\frac{\omega_{n_*}^2(k) - \omega^2}{\varepsilon}\widehat{B}_j^*\left(\frac{k - k^{(j)}}{\varepsilon}\right) = \omega^2 \chi_{\varepsilon,r}(k - k^{(j)}) \langle \widetilde{F}(\widetilde{u})(\cdot, k), p_{n_*}(\cdot, k)\rangle, \quad k \in \mathbb{C}^2, \quad (4.40)$$

where we define $\chi_{\varepsilon,r} \coloneqq \chi_{B_{\varepsilon^r}(0)}$.

Expanding the eigenvalue ω_{n_*} near $k^{(j)}$ by assumptions (A3) and (A4) as

$$\omega_{n_*}(k) = \omega_* + \frac{1}{2}(k - k^{(j)})^\mathsf{T}\nabla^2 \omega_{n_*}(k^{(j)})(k - k^{(j)}) + \omega_r(k),$$

where $|\omega_r(k)| \leqslant C|k - k^{(j)}|^3$, and then recalling (4.1), we obtain

$$\omega_{n_*}^2(k) - \omega^2 = \varepsilon^2 \omega_* \left(\left(\frac{k - k^{(j)}}{\varepsilon}\right)^T \nabla^2 \omega_{n_*}(k^{(j)})\left(\frac{k - k^{(j)}}{\varepsilon}\right) - 2\Omega\right) - \varepsilon^4 \Omega^2 + \omega_R(k)$$

with

$$|\omega_R(k)| \leqslant C|k - k^{(j)}|^3. \tag{4.41}$$

Inserting this into equation (4.40) and defining $k' \coloneqq \frac{k - k^{(j)}}{\varepsilon}$, we obtain for $k' \in B_{\varepsilon^{r-1}}(0)$

$$\omega_* \left((k')^\mathsf{T}\nabla^2 \omega_{n_*}(k^{(j)})k' - 2\Omega\right) \widehat{B}_j^*(k')$$

$$= \frac{\omega^2}{\varepsilon}\chi_{\varepsilon,r-1}(k') \langle \widetilde{F}(\widetilde{v})(\cdot, k^{(j)} + \varepsilon k'), p_{n_*}(\cdot, k^{(j)} + \varepsilon k')\rangle$$

$$+ \frac{\omega^2}{\varepsilon}\chi_{\varepsilon,r-1}(k') \langle \left(\widetilde{F}(\widetilde{u}) - \widetilde{F}(\widetilde{v})\right)(\cdot, k^{(j)} + \varepsilon k'), p_{n_*}(\cdot, k^{(j)} + \varepsilon k')\rangle$$

$$- \frac{1}{\varepsilon}\left(\omega_R(k^{(j)} + \varepsilon k') - \varepsilon^4 \Omega^2\right)\widehat{B}_j^*(k'). \tag{4.42}$$

Now we estimate separately the terms on the right. We will see that the second and third terms are small, while the first one recovers the right-hand side of the CMEs (4.14), so that (4.42) may be interpreted as a perturbed CME system.

First we deal with the third term of (4.42). By (4.41),

$$
\left\| \frac{\omega_R(k^{(j)} + \varepsilon \cdot)}{\varepsilon} \widehat{B}_j \right\|_{L^2(\varepsilon^{-1}\mathbb{B})}
$$

$$
= \frac{1}{\varepsilon} \left(\int_{\varepsilon^{-1}\mathbb{B}} |\omega_R(k^{(j)} + \varepsilon k')|^2 |\widehat{B}_j(k')|^2 \, \mathrm{d}k' \right)^{\frac{1}{2}}
$$

$$
\leqslant c\varepsilon^2 \left(\int_{B_{\varepsilon^{r-1}}(0)} \frac{|k'|^6}{(1 + |k'|)^{2s_B}} (1 + |k'|)^{2s_B} |\widehat{B}_j(k')|^2 \, \mathrm{d}k' \right)^{\frac{1}{2}}
$$

$$
\leqslant c\varepsilon^2 \sup_{|k'|<\varepsilon^{r-1}} \frac{|k'|^3}{(1 + |k'|)^{s_B}} \|\widehat{B}_j^*\|_{L^2_{s_B}}
$$

$$
\leqslant c\varepsilon^{2-\max\{0,(1-r)(3-s_B)\}} \|\widehat{B}_j^*\|_{L^2_{s_B}}. \tag{4.43}
$$

To make this term $o(1)$, we need that $2 > \max\{0, (1 - r)(3 - s_B)\}$. This is ensured for all $r \in \left(0, \frac{1}{2}\right]$ as long as we take $s_B > 1$.

The second term in (4.42) is estimated similarly as in section 4.4. Indeed,

$$
\left\| \frac{\omega^2}{\varepsilon}, \left\langle \left(\widetilde{F}(\widetilde{u}) - \widetilde{F}(\widetilde{v}) \right)(\cdot, k^{(j)} + \varepsilon k'), p_{n_*}(\cdot, k^{(j)} + \varepsilon k') \right\rangle \right\|_{L^2(\varepsilon^{-1}\mathbb{B})}
$$

$$
\leqslant \omega^2 \varepsilon^{-1} \mathrm{ess\,sup}_{k\in\mathbb{B}} \|p_{n_*}(\cdot, k)\|_{L^2(\mathcal{Q})}
$$

$$
\times \left(\int_{\varepsilon^{-1}\mathbb{B}} \| \left(\widetilde{F}(\widetilde{u}) - \widetilde{F}(\widetilde{v}) \right)(\cdot, k^{(j)} + \varepsilon k')\|^2_{L^2(\mathcal{Q})} \mathrm{d}k' \right)^{\frac{1}{2}}
$$

$$
\leqslant c\varepsilon^{-1} \left(\int_{\mathbb{B}+k^{(j)}} \| \left(\widetilde{F}(\widetilde{u}) - \widetilde{F}(\widetilde{v}) \right)(\cdot, k)\|^2_{H^2(\mathcal{Q})} \varepsilon^{-2} \, \mathrm{d}k \right)^{\frac{1}{2}}
$$

$$
= c\varepsilon^{-2} \|\widetilde{F}(\widetilde{u}) - \widetilde{F}(\widetilde{v})\|_{\mathcal{X}_2} \leqslant c\varepsilon, \tag{4.44}
$$

where the constant c depends just on $\left(\|\widehat{B}_j^*\|_{L^2_{s_B}} \right)_{j=1}^N$. The last inequality is given by (4.27).

Let us now address the first term in (4.42) on its support $k' \in B_{\varepsilon^{r-1}}(0)$. Equivalently we consider $k \in B_{\varepsilon^r}(k^{(j)})$ and split the term as follows:

$$
\langle \widetilde{F}(\widetilde{v})(\cdot, k), p_{n_*}(\cdot, k) \rangle
$$

$$
= \langle (\widetilde{F}(\widetilde{v}) - \widetilde{F}(\widetilde{v}_B))(\cdot, k), p_{n_*}(\cdot, k) \rangle
$$

$$
+ \varepsilon^{-3} \sum_{\alpha,\beta,\gamma=1}^N \int_{2\mathbb{B}} \int_{\mathbb{B}} \widehat{B}_\alpha \left(\frac{k-l-k^{(\alpha)}}{\varepsilon} \right) \widehat{B}_\beta \left(\frac{l-t-k^{(\beta)}}{\varepsilon} \right) \widehat{\overline{B}}_\gamma \left(\frac{t+k^{(\gamma)}}{\varepsilon} \right)
$$

$$
\cdot \sum_{a,b,c,d=1}^3 \left(\mathsf{B}^{a,b,c,d}(k, k-l, l-t, t) - \Theta^{a,b,c,d}_{\alpha,\beta,\gamma,j} \right) \mathrm{d}t \, \mathrm{d}l
$$

$$+ \varepsilon^{-3} \sum_{a,b,c,d=1}^{3} \sum_{\alpha,\beta,\gamma=1}^{N} \Theta_{\alpha,\beta,\gamma,j}^{a,b,c,d} \int_{2\mathbb{B}} \int_{\mathbb{B}} \widehat{B}_\alpha \left(\frac{k-l-k^{(\alpha)}}{\varepsilon} \right) \widehat{B}_\beta \left(\frac{l-t-k^{(\beta)}}{\varepsilon} \right) \overline{\widehat{B}_\gamma} \left(\frac{t+k^{(\gamma)}}{\varepsilon} \right) \, \mathrm{d}t \, \mathrm{d}l$$

$$:= W_1(k) + W_2(k) + W_3(k), \tag{4.45}$$

where

$$Б^{a,b,c,d}(k, k-l, l-t, t)$$

$$:= \left\langle \underline{\chi}_{a,b,c,d}^{(3)} p_{n_*,a}(\cdot, k-l) p_{n_*,b}(\cdot, l-t) \overline{p_{n_*,c}(\cdot, -t)}, p_{n_*,d}(\cdot, k) \right\rangle$$

and

$$\Theta_{\alpha,\beta,\gamma,j}^{a,b,c,d} := \left\langle \underline{\chi}_{a,b,c,d}^{(3)} u_{n_*,a}(\cdot, k^{(\alpha)}) u_{n_*,b}(\cdot, k^{(\beta)}) \times \overline{u_{n_*,c}(\cdot, k^{(\gamma)})}, u_{n_*,d}(\cdot, k^{(j)}) \right\rangle$$

$$= \left\langle \underline{\chi}_{a,b,c,d}^{(3)} p_{n_*,a}(\cdot, k^{(\alpha)}) p_{n_*,b}(\cdot, k^{(\beta)}) \times \overline{p_{n_*,c}(\cdot, k^{(\gamma)})} \mathrm{e}^{\mathrm{i} K_{\alpha,\beta,\gamma,j} \cdot}, p_{n_*,d}(\cdot, k^{(j)}) \right\rangle$$

$$= Б_{\alpha,\beta,\gamma,j}^{a,b,c,d}(k^{(j)}, k^{(\alpha)} - K_{\alpha,\beta,\gamma,j}, k^{(\beta)}, -k^{(\gamma)}) \tag{4.46}$$

with $K_{\alpha,\beta,\gamma,j} := k^{(\alpha)} + k^{(\beta)} - k^{(\gamma)} - k^{(j)}$.

The aim is to show that $\varepsilon^{-1} W_1$ and $\varepsilon^{-1} W_2$ are small and that $\varepsilon^{-1} \frac{\omega_*}{2} \chi_{\varepsilon,r}(\cdot - k^{(j)}) W_3$ is the Fourier transform of the nonlinear term $\mathcal{N}_j$ in the CMEs applied to $(\widehat{B}_1^*, \ldots, \widehat{B}_N^*)$ and evaluated at $\varepsilon^{-1}(\cdot - k^{(j)})$.

First, W_1 is estimated analogously to the term T_2 in section 4.5 producing

$$\|W_1\|_{L^2(\mathbb{B})} \leqslant c\varepsilon^{3-2r}. \tag{4.47}$$

For W_2 we take advantage of the Lipschitz continuity of $Б$ and of the asymptotically small support of the double convolution of the $\widehat{B}$'s. We rewrite

$$W_2(k^{(j)} + \varepsilon k') = \varepsilon \sum_{a,b,c,d=1}^{3} \sum_{\alpha,\beta,\gamma=1}^{N} \int_{\varepsilon^{-1}(2\mathbb{B}+k^{(\gamma)}-k^{(\beta)})} \int_{\varepsilon^{-1}(\mathbb{B}+k^{(\gamma)})}$$

$$\times \widehat{B}_\alpha(k'-l') \widehat{B}_\beta(l'-t') \overline{\widehat{B}_\gamma}(t') \cdot \Big(Б^{a,b,c,d} \big(k^{(j)} + \varepsilon k', k^{(\alpha)}$$

$$- K_{\alpha,\beta,\gamma,j} + \varepsilon(k'-l'), k^{(\beta)} + \varepsilon(l'-t'), -k^{(\gamma)} + \varepsilon t' \big)$$

$$- Б^{a,b,c,d}(k^{(j)}, k^{(\alpha)} - K_{\alpha,\beta,\gamma,j}, k^{(\beta)}, -k^{(\gamma)}) \Big) \, \mathrm{d}t' \, \mathrm{d}l'$$

using the obvious changes of variables and the $\varepsilon^{-1} \Lambda^*$-periodicity of $\widehat{B}_\alpha, \alpha \in \{1, \ldots, N\}$. Next, we exploit the fact that the map $(k_1, k_2, k_3, k_4) \mapsto Б^{a,b,c,d}(k_1, k_2, k_3, k_4)$ is Lipschitz continuous with respect to all variables, i.e. there is $C_Б > 0$ such that for all $k', l', t' \in \varepsilon^{-1} \mathbb{B}$ *

$$\Big| Б(k^{(j)} + \varepsilon k', k^{(\alpha)} - K_{\alpha,\beta,\gamma,j} + \varepsilon(k'-l'), k^{(\beta)} + \varepsilon(l'-t'), -k^{(\gamma)} + \varepsilon t')$$

$$- Б(k^{(j)}, k^{(\alpha)} - K_{\alpha,\beta,\gamma,j}, k^{(\beta)}, -k^{(\gamma)}) \Big|$$

$$\leqslant C_Б \varepsilon \left(|k'| + |k'-l'| + |l'-t'| + |t'| \right)$$

$$\leqslant 2 C_Б \varepsilon \left(|k'-l'| + |l'-t'| + |t'| \right),$$

where we have omitted the indices of Б for brevity. Therefore

$$\|W_2(k^{(j)} + \varepsilon \cdot)\|_{L^2(\varepsilon^{-1}\mathbb{B})}$$

$$\leqslant c\varepsilon^2 \sum_{\alpha,\beta,\gamma=1}^{N} \left(2\||\tau\widehat{B}_\alpha^*| * |\widehat{B}_\beta^*| * |\overline{\widehat{B}}_\gamma^*|\|_{L^2} + \||\widehat{B}_\alpha^*| * |\widehat{B}_\beta^*| * |\tau\overline{\widehat{B}}_\gamma^*|\|_{L^2} \right)$$

$$\leqslant c\varepsilon^2 \sum_{\alpha,\beta,\gamma=1}^{N} \|\widehat{B}_\alpha^*\|_{L_1^2} \|\widehat{B}_\beta^*\|_{L^1} \|\widehat{B}_\gamma^*\|_{L^1}$$

$$\leqslant c\varepsilon^2 \sum_{\alpha,\beta,\gamma=1}^{N} \|\widehat{B}_\alpha^*\|_{L_s^2} \|\widehat{B}_\beta^*\|_{L_s^2} \|\widehat{B}_\gamma^*\|_{L_s^2} \tag{4.48}$$

for any $s > 1$, where $\tau(k) := k$. Because $\|W_2\|_{L^2(\mathbb{B})} = \varepsilon\|W_2(k^{(j)} + \varepsilon \cdot)\|_{L^2(\varepsilon^{-1}\mathbb{B})}$, we get from (4.47) and (4.48)

$$\varepsilon^{-1}\|W_1 + W_2\|_{L^2(\mathbb{B})} \leqslant c(\varepsilon^{2-2r} + \varepsilon) \leqslant c\varepsilon \tag{4.49}$$

as $r \in \left(0, \frac{1}{2}\right]$.

4.6.1. Perturbed CMEs. We return to equation (4.42). By (4.43)–(4.45) and (4.49) we get for each $j \in \{1, \ldots, N\}$ and $k' \in B_{\varepsilon^{r-1}}(0)$

$$\omega_* \left((k')^\mathsf{T} \nabla^2 \omega_{n_*}(k^{(j)}) k' - 2\Omega \right) \widehat{B}_j^*(k')$$

$$= \frac{\omega^2}{\varepsilon} \chi_{\varepsilon,r-1}(k') \sum_{a,b,c,d=1}^{3} \sum_{\alpha,\beta,\gamma=1}^{N} \Theta_{\alpha,\beta,\gamma,j}^{a,b,c,d}$$

$$\cdot \frac{1}{\varepsilon^3} \left(\widehat{B}_\alpha \left(\frac{\cdot - k^{(\alpha)}}{\varepsilon} \right) *_\mathbb{B} \widehat{B}_\beta \left(\frac{\cdot - k^{(\beta)}}{\varepsilon} \right) *_\mathbb{B} \overline{\widehat{B}}_\gamma \left(\frac{\cdot - k^{(\gamma)}}{\varepsilon} \right) \right) (k^{(j)} + \varepsilon k')$$

$$+ Ж_j(k^{(j)} + \varepsilon k'), \tag{4.50}$$

where $Ж_j$ collects all the perturbations. Since $r \in \left(0, \frac{1}{2}\right]$, it is

$$\|Ж_j(k^{(j)} + \varepsilon \cdot)\|_{L^2(\varepsilon^{-1}\mathbb{B})} \leqslant c \left(\varepsilon^{2-\max\{0,(1-r)(3-s_B)\}} + \varepsilon + \varepsilon^{2-2r} \right)$$

$$\leqslant c \left(\varepsilon^{2-\max\{0,(1-r)(3-s_B)\}} + 2\varepsilon \right). \tag{4.51}$$

Prescribing now

$$s_B \geqslant 3 - \frac{1}{1-r}, \tag{4.52}$$

we see that the first exponent in (4.51) is greater than or equal to 1. Therefore, under condition (4.52) we get

$$\|Ж_j(k^{(j)} + \varepsilon \cdot)\|_{L^2(\varepsilon^{-1}\mathbb{B})} \leqslant c\varepsilon. \tag{4.53}$$

Note that (4.52) is satisfied e.g. by $s_B = 2$ since $r > 0$.

Next, a direct calculation shows that

$$\chi_{\varepsilon,r-1}(k')\frac{1}{\varepsilon^3}\left(\widehat{B}_\alpha\left(\frac{\cdot-k^{(\alpha)}}{\varepsilon}\right)*_\mathbb{B}\widehat{B}_\beta\left(\frac{\cdot-k^{(\beta)}}{\varepsilon}\right)*_\mathbb{B}\widehat{\overline{B}}_\gamma\left(\frac{\cdot-k^{(\gamma)}}{\varepsilon}\right)\right)(k)$$

$$=\begin{cases}\varepsilon(\widehat{B}_\alpha^**_{B_{\varepsilon^{r-1}}(0)}\widehat{B}_\beta^**_{B_{\varepsilon^{r-1}}(0)}\widehat{\overline{B}}_\gamma^*)(k')=\varepsilon(\widehat{B}_\alpha^**\widehat{B}_\beta^**\widehat{\overline{B}}_\gamma^*)(k')&\text{if }(\alpha,\beta,\gamma)\in\sigma_j\\0&\text{if }(\alpha,\beta,\gamma)\notin\sigma_j.\end{cases}$$

$$(4.54)$$

In detail: due to the periodicity of $\widehat{B}_\alpha$ the convolution $*_\mathbb{B}$ can be replaced by $*_{\mathbb{B}+k_*}$ for $k_*\in\mathbb{R}^2$ arbitrary. This implies by the obvious change of variables that the left-hand side equals

$$\frac{1}{\varepsilon^3}\int_{\mathbb{B}-k^{(\gamma)}}\left(\widehat{B}_\alpha\left(\frac{\cdot-k^{(\alpha)}}{\varepsilon}\right)*_{\mathbb{B}+k^{(\beta)}}\widehat{B}_\beta\left(\frac{\cdot-k^{(\beta)}}{\varepsilon}\right)\right)(k-l)\widehat{\overline{B}}_\gamma\left(\frac{l+k^{(\gamma)}}{\varepsilon}\right)\mathrm{d}l$$

$$=\varepsilon^{-1}\int_{B_{\varepsilon^{r-1}}(0)}\int_{\mathbb{B}+k^{(\beta)}}\widehat{B}_\alpha\left(\frac{k-\varepsilon l'-k^{(\alpha)}+k^{(\gamma)}-s}{\varepsilon}\right)\widehat{B}_\beta\left(\frac{s-k^{(\beta)}}{\varepsilon}\right)\widehat{\overline{B}}_\gamma(l')\mathrm{d}s\,\mathrm{d}l'$$

$$=\varepsilon\int_{B_{\varepsilon^{r-1}}(0)}\int_{B_{\varepsilon^{r-1}}(0)}\widehat{B}_\alpha\left(\frac{k-(k^{(\alpha)}+k^{(\beta)}-k^{(\gamma)})}{\varepsilon}-l'-s'\right)\widehat{B}_\beta(s')\widehat{\overline{B}}_\gamma(l')\mathrm{d}s'\,\mathrm{d}l'.$$

With the further transformation $l'+s'=:t'\in B_{2\varepsilon^{r-1}}(0)$, we infer

$$=\varepsilon\int_{B_{2\varepsilon^{r-1}}(0)}\int_{B_{\varepsilon^{r-1}}(0)}\widehat{B}_\alpha\left(\frac{k-k^{(j)}-K_{\alpha,\beta,\gamma,j}}{\varepsilon}-t'\right)\widehat{B}_\beta^*\left(t'-l'\right)\widehat{\overline{B}}_\gamma^*\left(l'\right)\mathrm{d}l'\,\mathrm{d}t'.$$

Now recall that $k\in B_{\varepsilon^r}(k^{(j)})$ and that $\mathrm{supp}(\widehat{B}_\alpha^*)\subset B_{\varepsilon^{r-1}}(0)$. Due to $t'\in B_{2\varepsilon^{r-1}}(0)$ and the $\varepsilon^{-1}\Lambda^*$-periodicity of $\widehat{B}_\alpha$, the function $\widehat{B}_\alpha\left(\frac{k-k^{(j)}-K_{\alpha,\beta,\gamma,j}}{\varepsilon}-t'\right)$ is nonzero if and only if $K_{\alpha,\beta,\gamma,j}\in\Lambda^*$, i.e. if $(\alpha,\beta,\gamma)\in\sigma_j$, with σ_j defined in (1.12). The periodicity allows then for dropping the shift $\varepsilon^{-1}K_{\alpha,\beta,\gamma,j}$ in the argument of $\widehat{B}_\alpha$. Moreover, for $k\in B_{\varepsilon^r}(k^{(j)})$ and $t'\in B_{2\varepsilon^{r-1}}(0)$ it is $\widehat{B}_\alpha\left(\frac{k-k^{(j)}}{\varepsilon}-t'\right)=\widehat{B}_\alpha^*\left(\frac{k-k^{(j)}}{\varepsilon}-t'\right)$. We get

$$\chi_{\varepsilon,r-1}(k')\frac{1}{\varepsilon^3}\left(\widehat{B}_\alpha\left(\frac{\cdot-k^{(\alpha)}}{\varepsilon}\right)*_\mathbb{B}\widehat{B}_\beta\left(\frac{\cdot-k^{(\beta)}}{\varepsilon}\right)*_\mathbb{B}\widehat{\overline{B}}_\gamma\left(\frac{\cdot-k^{(\gamma)}}{\varepsilon}\right)\right)(k^{(j)}+\varepsilon k')$$

$$=\begin{cases}\varepsilon\chi_{\varepsilon,r-1}(k')\int_{B_{2\varepsilon^{r-1}}(0)}\int_{B_{\varepsilon^{r-1}}(0)}\widehat{B}_\alpha\left(k'-t'\right)\widehat{B}_\beta^*\left(t'-l'\right)\widehat{\overline{B}}_\gamma^*\left(l'\right)\mathrm{d}l'\,\mathrm{d}t'&\text{if }(\alpha,\beta,\gamma)\in\sigma_j\\0&\text{if }(\alpha,\beta,\gamma)\notin\sigma_j,\end{cases}$$

from which (4.54) follows.

Hence, by (4.50), (4.53), and (4.54) we deduce that $\left(\widehat{B}_j^*\right)_{j=1}^N$ satisfy the perturbed CMEs

$$\mathcal{G}_j(\widehat{B}^*)(k'):=\left(\frac{1}{2}(k')^T\nabla^2\omega_{n_*}(k^{(j)})k'-\Omega\right)\widehat{B}_j^*(k')-\widehat{\mathcal{N}}_j(\widehat{B}^*)(k')=\widehat{R}_j(\widehat{B}^*)(k'),$$

$$(4.55)$$

for $j \in \{1, \ldots, N\}$ and $k' \in B_{\varepsilon^{r-1}}(0)$, where we recall that (cf (1.11) and (4.46))

$$
\widehat{\mathcal{N}}_j(\widehat{B}^*)(k') = \frac{\omega_*}{2} \sum_{a,b,c,d=1}^{3} \sum_{\alpha,\beta,\gamma=1}^{N} \Theta_{\alpha,\beta,\gamma,j}^{a,b,c,d} \times \left(\widehat{B}_\alpha^* *_{B_{\varepsilon^{r-1}}(0)} \widehat{B}_\beta^* *_{B_{\varepsilon^{r-1}}(0)} \overline{\widehat{B}}_\gamma^* \right)(k')
$$

$$
= \frac{\omega_*}{2} \sum_{(\alpha,\beta,\gamma) \in \sigma_j} I_{\alpha,\beta,\gamma}^{j} \left(\widehat{B}_\alpha^* * \widehat{B}_\beta^* * \overline{\widehat{B}}_\gamma^* \right)(k'),
$$

the coefficients $I_{\alpha,\beta,\gamma}^{j}$ being defined in (1.13). The remainder term $\widehat{R}_j(\widehat{B}^*) = \widehat{R}_j\left(\widehat{B}_1^*, \ldots, \widehat{B}_N^*\right)$ is defined via

$$
\widehat{R}_j(\widehat{B}^*)(k') := \frac{1}{2\omega_*} \mathpzc{K}_j(k^{(j)} + \varepsilon k'), \quad k' \in B_{\varepsilon^{r-1}}(0) \tag{4.56}
$$

and satisfies

$$
\|\widehat{R}_j(\widehat{B}^*)\|_{L^2(\varepsilon^{-1}\mathbb{B})} < c\varepsilon. \tag{4.57}
$$

Notice that (4.55) is therefore an ε-perturbation of the CMEs in Fourier variables on the compact support $k' \in B_{\varepsilon^{r-1}}(0)$. In what follows, we prove the existence of solutions $\widehat{B}^* := \left(\widehat{B}_j^*\right)_{j=1}^{N}$ of (4.55) close to $\chi_{\varepsilon,r-1}\widehat{A}$, where $\widehat{A} := \left(\widehat{A}_j\right)_{j=1}^{N}$ is the Fourier transform of the solution of the CMEs (1.10). We follow the approach of [9, 17].

To this aim, for $j \in \{1, \ldots, N\}$ we define $\widehat{A}_j^\varepsilon := \chi_{\varepsilon,r-1}\widehat{A}_j$ and write

$$
\widehat{B}_j^* = \widehat{A}_j^\varepsilon + \widehat{b}_j
$$

with $\operatorname{supp}(\widehat{b}_j) \subset B_{\varepsilon^{r-1}}(0)$. In order to expand $\mathcal{G} = (\mathcal{G}_1, \ldots, \mathcal{G}_N)$ around the vector $\widehat{A}^\varepsilon$ and use the Jacobian of the CMEs, we write $\mathcal{G}_j$ in the real variables. Indices R and I denote hereafter the real and the imaginary part respectively, e.g. $A_j = A_{j,R} + iA_{j,I}$. We define (cf (1.10))

$$
\phi_j(A) := -\frac{1}{2}\nabla^{\mathsf{T}}(\nabla^2 \omega_{n_*}(k^{(j)})\nabla A_j) - \Omega A_j - \mathcal{N}_j(A), \quad j \in \{1, \ldots, N\}
$$

so that $(\phi(A))^\wedge = \mathcal{G}(\widehat{A})$, which in real variables becomes

$$
\Phi_j(A_R, A_I) := \begin{pmatrix} \operatorname{Re}(\phi_j(A_R + iA_I)) \\ \operatorname{Im}(\phi_j(A_R + iA_I)) \end{pmatrix}, \quad j \in \{1, \ldots, N\}.
$$

We denote its $2N \times 2N$ Jacobian by $D\Phi(A_R, A_I)$, its Fourier counterpart by

$$
D_{\widehat{A}}\mathcal{G}(\widehat{A}) := (D\Phi(A_R, A_I))^\wedge,
$$

as well as its Fourier-truncation

$$
\chi_{\varepsilon,r-1}D_{\widehat{A}}\mathcal{G}(\widehat{A}^\varepsilon) = \chi_{\varepsilon,r-1}\left(D\Phi(A_R^\varepsilon, A_I^\varepsilon)\right)^\wedge,
$$

with $A_R^\varepsilon := (\widehat{A}_R^\varepsilon)^\vee$ and $A_I^\varepsilon := (\widehat{A}_I^\varepsilon)^\vee$. Thus here $D_{\widehat{A}}\mathcal{G}_j(\widehat{A})$ is just a symbolic notation. Recalling the definition of $\mathcal{N}_j$ in (1.11), we have

$$
\operatorname{Re}(\mathcal{N}_j(A)) = \sum_{(\alpha,\beta,\gamma) \in \sigma_j} I_{\alpha,\beta,\gamma}^{j} \left(A_{\alpha,R}A_{\beta,R}A_{\gamma,R} + A_{\alpha,R}A_{\beta,I}A_{\gamma,I} + A_{\alpha,I}A_{\beta,R}A_{\gamma,I} - A_{\alpha,I}A_{\beta,I}A_{\gamma,R} \right),
$$

$$\mathrm{Im}(\mathcal{N}_j(A)) = \sum_{(\alpha,\beta,\gamma)\in\sigma_j} I^j_{\alpha,\beta,\gamma} \left(A_{\alpha,R}A_{\beta,I}A_{\gamma,R} - A_{\alpha,R}A_{\beta,R}A_{\gamma,I} + A_{\alpha,I}A_{\beta,R}A_{\gamma,R} + A_{\alpha,I}A_{\beta,I}A_{\gamma,I} \right).$$

Therefore, for $m \in \{1, \ldots, N\}$

$$\partial_{A_{m,R}}\mathrm{Re}(\mathcal{N}_j(A)) = \sum_{\alpha,\beta=1}^{N} \nu^{11,R}_{\alpha\beta jm} A_{\alpha,R}A_{\beta,R} + \sum_{\alpha,\beta=1}^{N} \nu^{11,I}_{\alpha\beta jm} A_{\alpha,I}A_{\beta,I},$$

$$\partial_{A_{m,I}}\mathrm{Re}(\mathcal{N}_j(A)) = \partial_{A_{m,R}}\mathrm{Im}(\mathcal{N}_j(A)) = \sum_{\alpha,\beta=1}^{N} \nu^{12}_{\alpha\beta jm} A_{\alpha,R}A_{\beta,I},$$

$$\partial_{A_{m,I}}\mathrm{Im}(\mathcal{N}_j(A)) = \sum_{\alpha,\beta=1}^{N} \nu^{22,R}_{\alpha\beta jm} A_{\alpha,R}A_{\beta,R} + \sum_{\alpha,\beta=1}^{N} \nu^{22,I}_{\alpha\beta jm} A_{\alpha,I}A_{\beta,I},$$

where the coefficients $\nu^{11,R}_{\alpha\beta jm}, \ldots, \nu^{22,I}_{\alpha\beta jm}$ are linear combinations of $I^j_{\alpha,\beta,m}$ for all $\alpha, \beta \in \{1, \ldots, N\}$. Hence we may write

$$D\Phi(A_R, A_I) = L - D_A\mathcal{N}(A),$$

where

$$L := -\frac{1}{2} \begin{pmatrix} \nabla^{\mathsf{T}}\left(\nabla^2\omega_{n_*}(k^{(1)})\nabla\right)Id_{2\times 2} & & \\ & \ddots & \\ & & \nabla^{\mathsf{T}}\left(\nabla^2\omega_{n_*}(k^{(N)})\nabla\right)Id_{2\times 2} \end{pmatrix}$$
$$- \Omega Id_{2N\times 2N}$$

and $D_A\mathcal{N}(A)$ is a block matrix with the (j,k)-th block $(j,k \in \{1, \ldots, N\})$ being

$$M^{\mathcal{N}}(A)_{j,m} := \begin{pmatrix} \partial_{A_{m,R}}\mathrm{Re}(\mathcal{N}_j(A)) & \partial_{A_{m,I}}\mathrm{Re}(\mathcal{N}_j(A)) \\ \partial_{A_{m,R}}\mathrm{Im}(\mathcal{N}_j(A)) & \partial_{A_{m,I}}\mathrm{Im}(\mathcal{N}_j(A)) \end{pmatrix}$$

$$= \begin{pmatrix} \displaystyle\sum_{\alpha,\beta=1}^{N}\left(\nu^{11,R}_{\alpha\beta jm}A_{\alpha,R}A_{\beta,R} + \nu^{11,I}_{\alpha\beta jm}A_{\alpha,I}A_{\beta,I}\right) & \displaystyle\sum_{\alpha,\beta=1}^{N}\nu^{12}_{\alpha\beta jm}A_{\alpha,R}A_{\beta,I} \\ \displaystyle\sum_{\alpha,\beta=1}^{N}\nu^{21}_{\alpha\beta jm}A_{\alpha,R}A_{\beta,I} & \displaystyle\sum_{\alpha,\beta=1}^{N}\left(\nu^{22,R}_{\alpha\beta jm}A_{\alpha,R}A_{\beta,R} + \nu^{22,I}_{\alpha\beta jm}A_{\alpha,I}A_{\beta,I}\right) \end{pmatrix}.$$

In Fourier variables this rewrites as

$$D_{\widehat{A}}\mathcal{G}(\widehat{A}) = \widehat{L} - D_{\widehat{A}}\widehat{\mathcal{N}}(\widehat{A}), \tag{4.58}$$

where $\widehat{L}$ is a block-diagonal matrix with N blocks of size 2×2, where the jth block is $(\frac{1}{2}(k')^{\mathsf{T}}\nabla^2\omega_{n_*}(k^{(j)})k' - \Omega)Id_{2\times 2}$ and $D_{\widehat{A}}\widehat{\mathcal{N}}(\widehat{A})$ is a block matrix with the (j,m)-th block $(j, m \in$

$\{1, \ldots, N\}$) being

$$\hat{M}^{\mathcal{N}}(\widehat{A})_{j,m} = \begin{pmatrix} \sum_{\alpha,\beta=1}^{N} \left(\nu_{\alpha\beta jm}^{11,R} \widehat{A}_{\alpha,R} * \widehat{A}_{\beta,R} + \nu_{\alpha\beta jm}^{11,I} \widehat{A}_{\alpha,I} * \widehat{A}_{\beta,I} \right) & \sum_{\alpha,\beta=1}^{N} \nu_{\alpha\beta jm}^{12} \widehat{A}_{\alpha,R} * \widehat{A}_{\beta,I} \\ \sum_{\alpha,\beta=1}^{N} \nu_{\alpha\beta jm}^{21} \widehat{A}_{\alpha,R} * \widehat{A}_{\beta,I} & \sum_{\alpha,\beta=1}^{N} \left(\nu_{\alpha\beta jm}^{22,R} \widehat{A}_{\alpha,R} * \widehat{A}_{\beta,R} + \nu_{\alpha\beta jm}^{22,I} \widehat{A}_{\alpha,I} * \widehat{A}_{\beta,I} \right) \end{pmatrix}.$$

The action of $\widehat{L}$ is multiplicative but $\hat{M}^{\mathcal{N}}(\widehat{A})$ acts as a convolution operator, e.g. $(\widehat{A}_{\alpha,R} * \widehat{A}_{\beta,R})(\widehat{b}_{\alpha,R}) = \widehat{A}_{\alpha,R} * \widehat{A}_{\beta,R} * \widehat{b}_{\alpha,R}$. If $\widehat{A} \in L_{s_A}^2(\mathbb{R}^2)$ with $s_A > 1$, then

$$D_{\widehat{A}} \mathcal{G}(\widehat{A}) : L_2^2(\mathbb{R}^2) \to L^2(\mathbb{R}^2).$$

For $\widehat{L}$ this follows from the second order property of L. For $\hat{M}^{\mathcal{N}}(\widehat{A})$ we have, e.g.

$$\|\widehat{A}_{\alpha,R} * \widehat{A}_{\beta,R} * \widehat{b}_{\alpha,R}\|_{L^2} \leqslant \|\widehat{A}_{\alpha,R}\|_{L^1} \|\widehat{A}_{\beta,R}\|_{L^1} \|\widehat{b}_{\alpha,R}\|_{L^2}$$

$$\leqslant \|\widehat{A}_{\alpha,R}\|_{L_{s_A}^2} \|\widehat{A}_{\beta,R}\|_{L_{s_A}^2} \|\widehat{b}_{\alpha,R}\|_{L_2^2}$$

using Young's inequality for convolutions and (4.22).

From (4.55) and using a Taylor expansion of $\mathcal{G}(\widehat{A}^\varepsilon + \widehat{b})$, we deduce then the following system of equations for the error term $\widehat{b}$,

$$\chi_{\varepsilon,r-1} D_{\widehat{A}} \mathcal{G}(\widehat{A}) \widehat{b} = \chi_{\varepsilon,r-1} \widehat{R}(\widehat{A}^\varepsilon + \widehat{b}) - \chi_{\varepsilon,r-1} \left(\mathcal{G}(\widehat{A}^\varepsilon + \widehat{b}) - D_{\widehat{A}} \mathcal{G}(\widehat{A}) \widehat{b} \right)$$

$$= \chi_{\varepsilon,r-1} \widehat{R}(\widehat{A}^\varepsilon + \widehat{b}) - \chi_{\varepsilon,r-1}$$

$$\times \left(\mathcal{G}(\widehat{A}^\varepsilon) + \left(D_{\widehat{A}} \mathcal{G}(\widehat{A}^\varepsilon) - D_{\widehat{A}} \mathcal{G}(\widehat{A}) \right) \widehat{b} + g(\widehat{b}) \right)$$

$$=: \mathcal{W}(\widehat{b}), \tag{4.59}$$

where g is quadratic in $\widehat{b}$. Once more, we want to apply a fixed point argument to (4.59) on a small ball around the origin in $\left(L_{s_B}^2(\mathbb{R}^2) \right)^N$. Hence we need to estimate the terms in $\mathcal{W}$. First, using the assumption that $\left(\widehat{A}_j \right)_{j=1}^N$ solves the CMEs (4.14), for $k' \in B_{\varepsilon^{r-1}}(0)$ we have

$$\mathcal{G}_j(\widehat{A}^\varepsilon)(k') = \left(\frac{1}{2}(k')^\mathsf{T} \nabla^2 \omega_{n_*}(k^{(j)}) k' - \Omega \right) \widehat{A}_j(k') - \widehat{\mathcal{N}}_j(\widehat{A}^\varepsilon)(k')$$

$$= \widehat{\mathcal{N}}_j(\widehat{A})(k') - \widehat{\mathcal{N}}_j(\widehat{A}^\varepsilon)(k')$$

$$= \frac{\omega_*}{2} \sum_{(\alpha,\beta,\gamma) \in \sigma_j} I_{\alpha,\beta,\gamma}^j \left(\widehat{A}_\alpha * \widehat{A}_\beta * \overline{\widehat{A}}_\gamma - \widehat{A}_\alpha^\varepsilon * \widehat{A}_\beta^\varepsilon * \overline{\widehat{A}}_\gamma^\varepsilon \right)(k').$$

Notice that the right-hand side includes terms which are double convolutions between $\widehat{A}_j^\varepsilon$ and $\widehat{a}_j^\varepsilon := \widehat{A}_j - \widehat{A}_j^\varepsilon = \left(1 - \chi_{\varepsilon,r-1} \right) \widehat{A}_j$ with at least one occurrence of $\widehat{a}_j^\varepsilon$. Since for $k' \in \mathbb{R}^2 \backslash B_{\varepsilon^{r-1}}(0)$ there holds

$$|\widehat{a}_j^\varepsilon(k')| \leqslant (1 + |k'|)^{s_A} |\widehat{a}_j^\varepsilon(k')| \sup_{|k'| > \varepsilon^{r-1}} (1 + |k'|)^{-s_A}$$

$$\leqslant c\varepsilon^{s_A(1-r)} (1 + |k'|)^{s_A} |\widehat{A}_j(k')|,$$

we have by Young's inequality for convolutions and (4.22)

$$\|\widehat{a}_\alpha^\varepsilon * \widehat{A}_\beta^\varepsilon * \overline{\widehat{A}_\gamma^\varepsilon}\|_{L^2(\mathbb{R}^2)} \leqslant \|\widehat{a}_\alpha^\varepsilon\|_{L^2(\mathbb{R}^2)} \|\widehat{A}_\beta^\varepsilon\|_{L^1(\mathbb{R}^2)} \|\widehat{A}_\gamma^\varepsilon\|_{L^1(\mathbb{R}^2)}$$

$$\leqslant c\varepsilon^{s_A(1-r)} \|\widehat{A}_\alpha\|_{L^2_{s_A}(\mathbb{R}^2)} \|\widehat{A}_\beta\|_{L^2_{s_A}(\mathbb{R}^2)} \|\widehat{A}_\gamma\|_{L^2_{s_A}(\mathbb{R}^2)}$$

for $s_A > 1$, and similarly one may handle all other terms, because again by (4.22) one has $\|\widehat{a}_j^\varepsilon\|_{L^1} \leqslant \|\widehat{a}_j^\varepsilon\|_{L^2_{s_A}} \leqslant \|\widehat{A}_j\|_{L^2_{s_A}}$. Hence

$$\|\mathcal{G}_j(\widehat{A}^\varepsilon)\|_{L^2(\mathbb{R}^2)}$$

$$\leqslant c \sum_{\alpha,\beta,\gamma=1}^N \varepsilon^{s_A(1-r)} \|\widehat{A}_\alpha\|_{L^2_{s_A}(\mathbb{R}^2)} \|\widehat{A}_\beta\|_{L^2_{s_A}(\mathbb{R}^2)} \|\widehat{A}_\gamma\|_{L^2_{s_A}(\mathbb{R}^2)}. \tag{4.60}$$

Next, we estimate the difference of the Jacobians in (4.59). Since the linear part of them (cf(4.58)) is the same for $\widehat{A}^\varepsilon$ and $\widehat{A}$, we get

$$\|\chi_{\varepsilon,r-1} \left(D_{\widehat{A}}\mathcal{G}(\widehat{A}^\varepsilon) - D_{\widehat{A}}\mathcal{G}(\widehat{A}) \right) \widehat{b}\|_{L^2(\mathbb{R}^2)}$$

$$\leqslant \sum_{j,m=1}^N \|\chi_{\varepsilon,r-1} \left((\hat{M}^\mathcal{N}(\widehat{A}^\varepsilon))_{j,m} - (\hat{M}^\mathcal{N}(\widehat{A}))_{j,m} \right) \widehat{b}_m\|_{L^2(\mathbb{R}^2)}$$

$$\leqslant \sum_{\alpha,\beta,j,m=1}^N \left[\|\nu_{\alpha\beta jm}^{11,R} \left(\widehat{A}_{\alpha,R}^\varepsilon * \widehat{A}_{\beta,R}^\varepsilon - \widehat{A}_{\alpha,R} * \widehat{A}_{\beta,R} \right) * \widehat{b}_{m,R}\|_2 \right.$$

$$+ \|\nu_{\alpha\beta jm}^{11,I} \left(\widehat{A}_{\alpha,I}^\varepsilon * \widehat{A}_{\beta,I}^\varepsilon - \widehat{A}_{\alpha,I} * \widehat{A}_{\beta,I} \right) * \widehat{b}_{m,R}\|_2$$

$$+ \|\nu_{\alpha\beta jm}^{12} \left(\widehat{A}_{\alpha,R}^\varepsilon * \widehat{A}_{\beta,I}^\varepsilon - \widehat{A}_{\alpha,R} * \widehat{A}_{\beta,I} \right) * \widehat{b}_{m,I}\|_2$$

$$+ \|\nu_{\alpha\beta jm}^{12} \left(\widehat{A}_{\alpha,R}^\varepsilon * \widehat{A}_{\beta,I}^\varepsilon - \widehat{A}_{\alpha,R} * \widehat{A}_{\beta,I} \right) * \widehat{b}_{m,R}\|_2$$

$$+ \|\nu_{\alpha\beta jm}^{22,R} \left(\widehat{A}_{\alpha,R}^\varepsilon * \widehat{A}_{\beta,R}^\varepsilon - \widehat{A}_{\alpha,R} * \widehat{A}_{\beta,R} \right) * \widehat{b}_{m,I}\|_2$$

$$\left. + \|\nu_{\alpha\beta jm}^{22,I} \left(\widehat{A}_{\alpha,I}^\varepsilon * \widehat{A}_{\beta,I}^\varepsilon - \widehat{A}_{\alpha,I} * \widehat{A}_{\beta,I} \right) * \widehat{b}_{m,I}\|_2 \right].$$

We see that all terms are of the same kind and, moreover, are linear in $\widehat{b}$ and either linear or quadratic in $\widehat{a}^\varepsilon := \widehat{A}^\varepsilon - \widehat{A}$. Applying then estimates analogous to the ones used to deduce (4.60), we infer

$$\|\chi_{\varepsilon,r-1} \left(D_{\widehat{A}}\mathcal{G}(\widehat{A}^\varepsilon) - D_{\widehat{A}}\mathcal{G}(\widehat{A}) \right) \widehat{b}\|_{L^2(\mathbb{R}^2)}$$

$$\leqslant c\varepsilon^{s_A(1-r)} \sum_{\alpha,\beta,\gamma=1}^N \|\widehat{A}_\alpha\|_{L^2_{s_A}(\mathbb{R}^2)} \|\widehat{A}_\beta\|_{L^2_{s_A}(\mathbb{R}^2)} \|\widehat{b}\|_{L^2_{s_B}(\mathbb{R}^2)} \tag{4.61}$$

if $s_A > 1$. Combining (4.57) (note that the dependence of c on $\|\widehat{B}\|_{L^2_{s_B}}$ —and in turn on $\|\widehat{b}\|_{L^2_{s_B}}$ —is polynomial), (4.60), and (4.61), we can thus conclude from (4.59) that

$$\|\mathcal{W}(\widehat{b})\|_{L^2(\mathbb{R}^2)} \leqslant c_A \left(\varepsilon + \varepsilon^{s_A(1-r)} + (\varepsilon + \varepsilon^{s_A(1-r)})\|\widehat{b}\|_{L^2_{s_B}} + \|\widehat{b}\|_{L^2_{s_B}}^2 + \|\widehat{b}\|_{L^2_{s_B}}^3 \right). \tag{4.62}$$

In order to solve (4.59) for $\widehat{b}$ by a fixed point argument, we would need the invertibility of the Jacobian $D_{\widehat{A}}\mathcal{G}(\widehat{A}) : L_2^2(\mathbb{R}^2) \to L^2(\mathbb{R}^2)$. Indeed, from this it would follow that $\chi_{\varepsilon,r-1}D_{\widehat{A}}\mathcal{G}(\widehat{A})\chi_{\varepsilon,r-1} : L_2^2(\mathbb{R}^2) \to L^2(\mathbb{R}^2)$ is uniformly invertible, see [21, theorem IV.3.17]. However, this is not the case, because of the presence of the three zero eigenvalues of $D\Phi(A_R, A_I)$ produced by the two spacial shift invariances and the complex phase invariance of the CMEs (1.10). To eliminate the zero eigenvalues, we assume the non-degeneracy of A, see definition 3.1, and work (in Fourier variables) in a subspace of $L_2^2(\mathbb{R}^2)$ in which the invariances do not hold. A natural subspace is the one generated by the $\mathcal{PT}$-symmetry, i.e. we work with $\widehat{A}$ and $\widehat{b}$ such that

$$A(-x) = \overline{A(x)}, \quad \text{and} \quad b(-x) = \overline{b(x)}$$

or equivalently,

$$\widehat{A} : \mathbb{R}^2 \to \mathbb{R}^N, \qquad \widehat{b} : \mathbb{R}^2 \to \mathbb{R}^N.$$

Under the non-degeneracy condition, the Jacobian $D_{\widehat{A}}\mathcal{G}(\widehat{A})$ is invertible in such a subspace and we can apply a fixed point argument to equation (4.59). In detail, assuming $\widehat{A} : \mathbb{R}^2 \to \mathbb{R}^N$, we look for a solution of

$$\widehat{b} = \left(\chi_{\varepsilon,r-1}D_{\widehat{A}}\mathcal{G}(\widehat{A}) \right)^{-1} \mathcal{W}(\widehat{b}) \tag{4.63}$$

in the space

$$\boldsymbol{L}_{2,\mathrm{sym}}^2 := \left\{ \widehat{b} \in L_2^2(\mathbb{R}^2)^N \,|\, \mathrm{supp}(\widehat{b}) \subset B_{\varepsilon^{r-1}}, \, \widehat{b}(\cdot) \text{ is real} \right\}.$$

However, we need to make sure that the $\mathcal{PT}$-symmetry is preserved by the maps $\mathcal{W}$ and $\left(\chi_{\varepsilon,r-1}D_{\widehat{A}}\mathcal{G}(\widehat{A}) \right)^{-1}$. This is proved at the end of the section. Now we address the application of the fixed point argument to (4.63) in the ball

$$\mathcal{B}_{c\varepsilon^\rho}^{2,\mathrm{sym}} := \left\{ \widehat{b} \in \boldsymbol{L}_{2,\mathrm{sym}}^2 \,|\, \|\widehat{b}\|_{L_2^2} \leqslant c\varepsilon^\rho \right\},$$

where $c, \rho > 0$ have to be found. For $\widehat{b} \in \mathcal{B}_{c\varepsilon^\rho}^{2,\mathrm{sym}}$ we deduce from (4.62) that

$$\|\mathcal{W}(\widehat{b})\|_{L^2(\mathbb{R}^2)} \leqslant c_A \left((\varepsilon + \varepsilon^{s_A(1-r)})(1 + c\varepsilon^\rho) + c^2\varepsilon^{2\rho} + c^3\varepsilon^{3\rho} \right).$$

Choosing

$$\rho = \min\{1, s_A(1-r)\} \quad \text{and} \quad c = 2c_A, \tag{4.64}$$

we infer $\mathcal{W}(\widehat{b}) \in \mathcal{B}_{c\varepsilon^\rho}^{2,\mathrm{sym}}$. Moreover, the map $\widehat{b} \mapsto \mathcal{W}(\widehat{b})$ is contractive in such a ball. Indeed, for $\widehat{b}_1, \widehat{b}_2 \in \mathcal{B}_{c\varepsilon^\rho}^{2,\mathrm{sym}}$,

$$\|\mathcal{W}(\widehat{b}^{(1)}) - \mathcal{W}(\widehat{b}^{(2)})\|_{L^2(\mathbb{R}^2)}$$

$$\leqslant \|g(\widehat{b}^{(1)}) - g(\widehat{b}^{(2)})\|_2 + \|\chi_{\varepsilon,r-1} \left(\widehat{R}(\widehat{A}^\varepsilon + \widehat{b}^{(1)}) - \widehat{R}(\widehat{A}^\varepsilon + \widehat{b}^{(2)}) \right) \|_2$$

$$+ \left\| \chi_{\varepsilon,r-1} \left(D_{\widehat{A}}\mathcal{G}(\widehat{A}^\varepsilon) - D_{\widehat{A}}\mathcal{G}(\widehat{A}) \right) (\widehat{b}^{(1)} - \widehat{b}^{(2)}) \right\|_2$$

$$\leqslant c(\varepsilon + \varepsilon^{s_A(1-r)})\|\widehat{b}^{(1)} - \widehat{b}^{(2)}\|_2,$$

because of (4.57), (4.61), and of the quadratic nature of g. Since $\chi_{\varepsilon,r-1} D_{\widehat{A}} \mathcal{G}(\widehat{A}) : L_{2,\text{sym}}^2 \to L^2(\mathbb{R}^2)$ is boundedly invertible, the existence of a $\mathcal{PT}$-symmetric solution $\widehat{b} \in L_{2,\text{sym}}^2$ of equation (4.59) so that

$$\|\widehat{b}\|_{L_2^2} \leqslant 2c_A \varepsilon^{\min\{1, s_A(1-r)\}} \tag{4.65}$$

follows from the Banach fixed point theorem. Notice that the optimal estimate $\|\widehat{b}\|_{L_2^2} \leqslant 2c_A \varepsilon$ can be obtained for any $s_A > 1$ as $r \in \left(0, \frac{1}{2}\right]$ can be chosen arbitrarily small.

To conclude the argument, it remains to be proved that the $\mathcal{PT}$-symmetry is preserved by the maps $\mathcal{W}$ and $\left(\chi_{\varepsilon,r-1} D_{\widehat{A}} \mathcal{G}(\widehat{A})\right)^{-1}$, i.e. that they map real valued functions $\widehat{b}$ to real valued functions. First, note that u is $\mathcal{PT}$-symmetric if and only if $\widetilde{u}(\cdot, k)$ is so for almost all $k \in \mathbb{B}$. Hence, we can check the inheritance of the property in the Bloch setting. We now need to make sure that all the components in which we decomposed our solution, and which now depend just on $\widehat{b}$, inherit the $\mathcal{PT}$-symmetry. If so, then the residual term $\widehat{R}$ in (4.56) is real. To complete this step, analyzing the equations that $\widetilde{w}_0$, $\widetilde{w}_R$, and $\widehat{C}$ have to fulfill, namely (4.10), (4.11), and (4.30), we see that we just need that our operator L_k, the projections ${}^\epsilon Q_k$ and Q_k^ϵ, and the nonlinear map $\widetilde{F}$ commute with $\mathcal{PT}$. In detail:

- L_k is $\mathcal{PT}$-symmetric since it involves only derivatives of order 2 and 0 and by assumption (A6).
- By the simpleness assumption (A7), the Bloch eigenfunctions $p_{n_*}(\cdot, k)$ are $\mathcal{PT}$-symmetric for almost all $k \in \mathbb{B}$, see (2.13). This, together with (A6), implies that the projections ${}^\epsilon P_k, {}^\epsilon Q_k, P_k^\epsilon, Q_k^\epsilon$ commute with $\mathcal{PT}$. E.g.,

$$
\begin{aligned}
P_k^\epsilon &\left(\mathcal{PT}(\widetilde{u})\right)(x, k) \\
&= \sum_{j=1}^N \langle \mathcal{PT}(\widetilde{u})(\cdot, k), \epsilon p_{n_*}(\cdot, k) \rangle p_{n_*}(x, k) \\
&= \sum_{j=1}^N \langle \mathcal{PT}(\widetilde{u})(\cdot, k), \epsilon \mathcal{PT}(p_{n_*})(\cdot, k) \rangle \mathcal{PT}(p_{n_*})(x, k) \\
&= \mathcal{PT}\left(\sum_{j=1}^N \overline{\langle \mathcal{PT}(\widetilde{u})(\cdot, k), \epsilon \mathcal{PT}(p_{n_*})(\cdot, k) \rangle} \, p_{n_*}(x, k) \right) \\
&= \mathcal{PT}\left(\sum_{j=1}^N \langle \widetilde{u}(\cdot, k), \epsilon p_{n_*}(\cdot, k) \rangle p_{n_*}(x, k) \right) = \mathcal{PT}(P_k^\epsilon \widetilde{u})(x, k),
\end{aligned}
$$

since

$$
\begin{aligned}
\int_{\mathbb{R}^2} \overline{\widetilde{u}}(-x, k) \cdot \epsilon(x) p_{n_*}(-x, k) \mathrm{d}x &= \int_{\mathbb{R}^2} \overline{\widetilde{u}}(y, k) \cdot \epsilon(-y) p_{n_*}(y, k) \mathrm{d}y \\
&= \overline{\int_{\mathbb{R}^2} \widetilde{u}(y, k) \cdot \epsilon(y) \overline{p_{n_*}(y, k)} \mathrm{d}y} \\
&= \overline{\langle \widetilde{u}(\cdot, k), \epsilon p_{n_*}(\cdot, k) \rangle}.
\end{aligned}
$$

- $\widetilde{F}$ only involves convolutions in $\mathbb{B}$ (cf (4.3)), hence the $\mathcal{PT}$-symmetry is trivially preserved using the evenness of $\chi^{(3)}$, see assumption (A6).

Consequently, if we start with a $\mathcal{PT}$-symmetric solution $\widehat{A}$ of the CMEs (1.10) and consider $\widehat{b} \in \mathcal{B}_{c\varepsilon\rho}^{2,\mathrm{sym}}$, then all components $\widehat{B}, \widehat{C}, \widetilde{w}_R, \widetilde{w}_0$ inherit the same symmetry. This implies that the term $\widehat{R}(\widehat{B})$ is real. Moreover, exploiting the $\mathcal{PT}$-symmetry of the mode $p_{n_*}(\cdot, k)$, see (2.13), it is easy to show that the coefficients $\Theta_{\alpha,\beta,\gamma,j}^{a,b,c,d}$ defined in (4.46) (or equivalently the coefficients $I_{\alpha,\beta,\gamma}^{j}$ defined in (1.13)) are real. Hence also $\mathcal{G}(\widehat{B})$ and $D_{\widehat{A}}\mathcal{G}(\widehat{A})$ are real. We are able to conclude that $\left(\chi_{\varepsilon,r-1}D_{\widehat{A}}\mathcal{G}(\widehat{A})\right)^{-1}\mathcal{W}: \mathcal{B}_{c\varepsilon\rho}^{2,\mathrm{sym}} \to \mathcal{B}_{c\varepsilon\rho}^{2,\mathrm{sym}}$ with the former choices of ρ and c in (4.64), and therefore we find a real solution $\widehat{b}$ to (4.63) satisfying (4.65).

This shows that the function $\widetilde{u}$ in (4.7) constructed in sections 4.3–4.6 is a $\mathcal{PT}$-symmetric H^2 solution of (4.2).

4.7. Approximation error of u_{ans}

In order to complete the proof of theorem 1.1 we need to show that the initial ansatz u_{ans} defined in (1.8) is actually a good approximation of the solution u of (1.6) which we constructed in sections 4.3–4.6. Recalling that

$$\widetilde{u}(x,k) = \widetilde{v}(x,k) + \widetilde{w}(x,k) = \widetilde{v}_B(x,k) + \widetilde{v}_C(x,k) + \widetilde{w}_0(x,k) + \widetilde{w}_R(x,k),$$

and in virtue of the estimates (4.24), (4.29), and (4.38), we have

$$\|\widetilde{u}_{\mathrm{ans}} - \widetilde{u}\|_{\mathcal{X}_2} \leqslant \|\widetilde{u}_{\mathrm{ans}} - \widetilde{v}_B\|_{\mathcal{X}_2} + c_A(\varepsilon^{2-2r} + \varepsilon^2 + \varepsilon^3) \leqslant \|\widetilde{u}_{\mathrm{ans}} - \widetilde{v}_B\|_{\mathcal{X}_2} + c\varepsilon^{2-2r}. \quad (4.66)$$

We split now the first term as follows (cf (4.12)):

$$
\begin{aligned}
\widetilde{u}_{\mathrm{ans}}&(x,k) - \widetilde{v}_B(x,k) \\
&= \frac{1}{\varepsilon}\sum_{j=1}^{N}\left\{ \sum_{K\in\Lambda^*}\widehat{A}_j\left(\frac{k-k^{(j)}+K}{\varepsilon}\right)p_{n_*}(x,k^{(j)})e^{iK\cdot x} - \widehat{B}_j\left(\frac{k-k^{(j)}}{\varepsilon}\right)p_{n_*}(x,k) \right\} \\
&= \frac{1}{\varepsilon}\sum_{j=1}^{N}\left\{ -\widehat{b}_j\left(\frac{k-k^{(j)}}{\varepsilon}\right)p_{n_*}(x,k) + (\chi_{\varepsilon,r-1}\widehat{A}_j)\left(\frac{k-k^{(j)}}{\varepsilon}\right) \right. \\
&\quad \cdot \left(p_{n_*}(x,k^{(j)}) - p_{n_*}(x,k)\right) + ((1-\chi_{\varepsilon,r-1})\widehat{A}_j)\left(\frac{k-k^{(j)}}{\varepsilon}\right) \\
&\quad \cdot p_{n_*}(x,k^{(j)}) + \left. \sum_{0\neq K\in\Lambda^*}\widehat{A}_j\left(\frac{k-k^{(j)}+K}{\varepsilon}\right)p_{n_*}(x,k^{(j)})e^{iK\cdot x} \right\}
\end{aligned}
$$

and we estimate term by term. First, by (4.65) one gets

$$
\left\|\widehat{b}_j\left(\frac{\cdot - k^{(j)}}{\varepsilon}\right)p_{n_*}\right\|_{\mathcal{X}_2}^2 \leqslant c\int_{\mathbb{B}}\left|\widehat{b}_j\left(\frac{k-k^{(j)}}{\varepsilon}\right)\right|^2 dk = c\varepsilon^2\|\widehat{b}_j\|_{L^2(\varepsilon^{-1}\mathbb{B})}^2
$$

$$
\leqslant c\varepsilon^{2\left(1+\min\{1, s_A(1-r)\}\right)}. \quad (4.67)
$$

Second, using the Lipschitz continuity of the map $k \mapsto p_{n_*}(\cdot, k) \in H^2(\mathcal{Q})$ for k in a vicinity of $k^{(j)}$ given by lemma A.7, we get

$$\int_{\mathbb{B}} \left| (\chi_{\varepsilon, r-1} \widehat{A}_j) \left(\frac{k - k^{(j)}}{\varepsilon} \right) \right|^2 \| p_{n_*}(\cdot, k) - p_{n_*}(\cdot, k^{(j)}) \|_{H^2(\mathcal{Q})}^2 \, \mathrm{d}k$$

$$\leqslant c\varepsilon^2 \int_{B_{\varepsilon^r}(k^{(j)})} \left| \widehat{A}_j \left(\frac{k - k^{(j)}}{\varepsilon} \right) \right|^2 \left| \frac{k - k^{(j)}}{\varepsilon} \right|^2 \, \mathrm{d}k$$

$$\leqslant c\varepsilon^4 \int_{B_{\varepsilon^{r-1}}(0)} |z|^2 |\widehat{A}_j(z)|^2 \, \mathrm{d}z = c\varepsilon^4 \| \widehat{A}_j \|_{L_1^2}^2 \leqslant c\varepsilon^4 \| \widehat{A}_j \|_{L_{s_A}^2}^2 . \tag{4.68}$$

Next,

$$\int_{\mathbb{B}} \left| (1 - \chi_{\varepsilon, r-1} \widehat{A}_j) \left(\frac{k - k^{(j)}}{\varepsilon} \right) \right|^2 \| p_{n_*}(\cdot, k^{(j)}) \|_{H^2(\mathcal{Q})}^2 \, \mathrm{d}k$$

$$\leqslant c\varepsilon^2 \sup_{z \notin B_{\varepsilon^{r-1}}(0)} (1 + |z|)^{-2s_A} \int_{\mathbb{R}^2 \setminus B_{\varepsilon^{r-1}}(0)} (1 + |z|)^{2s_A} |\widehat{A}_j(z)|^2 \, \mathrm{d}z$$

$$\leqslant c\varepsilon^2 (1 + \varepsilon^{r-1})^{-2s_A} \| \widehat{A}_j \|_{L_{s_A}^2}^2 \leqslant c\varepsilon^{2(1 + s_A(1-r))} \| \widehat{A}_j \|_{L_{s_A}^2}^2 . \tag{4.69}$$

Finally we consider the term involving the translated Brillouin zones:

$$\sum_{0 \neq K \in \Lambda^*} \int_{\mathbb{B}} \left| \widehat{A}_j \left(\frac{k - k^{(j)} + K}{\varepsilon} \right) \right|^2 \| p_{n_*}(\cdot, k^{(j)}) \mathrm{e}^{\mathrm{i}K \cdot} \|_{H^2(\mathcal{Q})}^2 \, \mathrm{d}k$$

$$\leqslant c\varepsilon^2 \sup_{0 \neq \tilde{K} \in \Lambda^*} \sup_{z \in \varepsilon^{-1}(\mathbb{B} - k^{(j)} + \tilde{K})} (1 + |z|)^{-2s_A} \sum_{0 \neq K \in \Lambda^*} \int_{\varepsilon^{-1}(\mathbb{B} - k^{(j)} + K)} (1 + |z|)^{2s_A} |\widehat{A}_j(z)|^2 \, \mathrm{d}z$$

$$\leqslant c\varepsilon^2 (1 + \varepsilon^{-1})^{-2s_A} \int_{\mathbb{R}^2 \setminus \varepsilon^{-1}(\mathbb{B} - k^{(j)})} (1 + |z|)^{2s_A} |\widehat{A}_j(z)|^2 \, \mathrm{d}z$$

$$\leqslant c\varepsilon^{2(1 + s_A)} \| \widehat{A}_j \|_{L_{s_A}^2}^2 . \tag{4.70}$$

By combining estimates (4.66)–(4.70) we arrive at

$$\| \tilde{u}_{\text{ans}} - \tilde{u} \|_{\mathcal{X}_2} \leqslant \frac{c}{\varepsilon} \left(\varepsilon^{1 + \min\{1, s_A(1-r)\}} \left(\varepsilon^2 + \varepsilon^{1 + s_A(1-r)} + \varepsilon^{1 + s_A} \right) \| \widehat{A}_j \|_{L_{s_A}^2} + \varepsilon^{3 - 2r} \right)$$

$$\leqslant c\varepsilon^{\min\{1, s_A(1-r)\}}$$

since $r \in \left\{ (0, \frac{1}{2} \right]$. Because $s_A > 1$, if we take $r \in \left(0, 1 - \frac{1}{s_A} \right]$ then $s_A(1 - r) \geqslant 1$ and hence

$$\| \tilde{u}_{\text{ans}} - \tilde{u} \|_{\mathcal{X}_2} \leqslant c\varepsilon.$$

The proof is thus complete recalling that the Bloch transform is an isomorphism between $\mathcal{X}_2$ and $H^2(\mathbb{R}^2)$, see (2.26).

Acknowledgments

Appendix A

In this last section we collect some auxiliary results needed throughout the paper. First, retracing the strategy of its standard proof (see e.g. [40, theorem 23.17]), we prove a Helmholtz decomposition adapted to our 'shifted' operator ∇'_k. This is employed in section 2.2 for the well-posedness of the eigenvalue problem in $L^2_{\#}(\mathcal{Q})$. Second, we address the regularity of the eigenfunctions $p_j(\cdot, k)$ and the Lipschitz continuity of the maps $k \mapsto \omega_j(k)$, $k \mapsto q_j(\cdot, k)$, and $k \mapsto p_j(\cdot, k)$, i.e. of the eigenvalue and eigenfunctions of the Bloch eigenvalue problems (2.8) and (2.10). The Lipschitz continuity is exploited in the nonlinear estimates of section 4.7.

A.1. Helmholtz decomposition

We note first that $H^1(\Omega, \mathbb{C}) = \{f \in L^2(\Omega, \mathbb{C}) \,|\, \nabla'_k f \in L^2(\Omega, \mathbb{C}^3)\}$ for any $k \in \mathbb{R}^2$ and any measurable $\Omega \subset \mathbb{R}^2$. We also define

$$H(\mathrm{curl}, \Omega) := \{v \in L^2(\Omega, \mathbb{C}^3) : \nabla' \times v \in L^2(\Omega, \mathbb{C}^3)\}.$$

Lemma A.1. *Let $\Omega \subset \mathbb{R}^2$ be a bounded domain, $\kappa \in \mathbb{R}\backslash\{0\}$ and $k \in \mathbb{R}^2$. Then*

$$H(\mathrm{curl}, \Omega) = W_k \oplus Z_k,$$

where

$$W_k := \left\{ w \in H(\mathrm{curl}, \Omega) \,\bigg|\, \int_\Omega w \cdot \overline{\nabla'_k f} = 0, \quad \forall\, f \in H^1(\Omega, \mathbb{C}) \right\}$$

and

$$Z_k := \{g \in L^2(\Omega, \mathbb{C}^3) \,|\, \exists\, \psi \in H^1(\Omega, \mathbb{C}) : g = \nabla'_k \psi\}$$

are closed subspaces.

Proof. Notice that W_k and Z_k are by definition orthogonal in $L^2(\Omega, \mathbb{C}^3)$ and $Z_k \subset H(\mathrm{curl}, \Omega)$ since $\nabla'_k \times (\nabla'_k \psi) = 0$ for all $\psi \in H^1(\Omega, \mathbb{C})$.
 Step 1: W_k *and* Z_k *are closed.*
 Let $(w_j)_j \subset W_k$, i.e. $\int_\Omega w_j \cdot \overline{\nabla'_k f} = 0$ for any $j \in \mathbb{N}$, and assume $w_j \to w$ in $H(\mathrm{curl}, \Omega)$. Then

$$\left| \int_\Omega w_j \cdot \overline{\nabla'_k f} - \int_\Omega w \cdot \overline{\nabla'_k f} \right| \leqslant \int_\Omega |w_j - w| |\nabla'_k f| \leqslant \|w_j - w\|_2 \|\nabla'_k f\|_2 \to 0,$$

therefore $w \in W_k$.
 Let now $(g_j)_j \subset Z_k$ be such that $g_j \to g$ in $H(\mathrm{curl}, \Omega)$. Then $g_j = \nabla'_k \psi_j$ with $\psi_j \in H^1(\Omega, \mathbb{C})$ for any $j \in \mathbb{N}$ and the sequence $(\nabla'_k \psi_j)_j$ is Cauchy in the L^2-norm. Noticing that

$$\|\nabla'_k \psi_j\|_2^2 = \|\partial_1 \psi_j + \mathrm{i} k_1 \psi_j\|_2^2 + \|\partial_2 \psi_j + \mathrm{i} k_2 \psi_j\|_2^2 + |\kappa|^2 \|\psi_j\|_2^2,$$

one immediately infers that also $(\psi_j)_j$ is Cauchy in $L^2(\Omega, \mathbb{C})$ and since $\|\psi_j\|_{H^1}^2 \leqslant \|\nabla_k' \psi_j\|_{L^2}^2 + c(k, \kappa)\|\psi_j\|_{L^2}^2$, the sequence $(\psi_j)_j$ is Cauchy also in $H^1(\Omega, \mathbb{C})$. Hence there exists $\psi \in H^1(\Omega, \mathbb{C})$ and thus $\psi_j \to \psi$ in $H^1(\Omega, \mathbb{C})$, such that $g_j = \nabla_k' \psi_j \to \nabla_k' \psi$ in $L^2(\Omega, \mathbb{C}^3)$. By the uniqueness of the limit we deduce $g = \nabla_k' \psi$, hence $g \in Z_k$.

Step 2: *decomposition.*

Let $v \in H(\mathrm{curl}, \Omega)$ and introduce $\mu_v \in (H^1(\Omega, \mathbb{C}))'$ via

$$\mu_v(\varphi) := \int_\Omega v \cdot \overline{\nabla_k' \varphi} \quad \text{for all } \varphi \in H^1(\Omega, \mathbb{C}),$$

as well as the sesquilinear form $S_k : H^1(\Omega, \mathbb{C}) \times H^1(\Omega, \mathbb{C}) \to \mathbb{C}$ via

$$\begin{aligned} S_k(\psi, \varphi) &:= \int_\Omega \nabla_k' \psi \cdot \overline{\nabla_k' \varphi} \\ &= \int_\Omega \nabla\psi \cdot \overline{\nabla\varphi} + (\kappa^2 + k_1^2 + k_2^2)\int_\Omega \psi\overline{\varphi} - \mathrm{i}\int_\Omega \nabla\psi \cdot k\overline{\varphi} + \mathrm{i}\int_\Omega \overline{\nabla\varphi} \cdot k\psi, \end{aligned}$$

which is clearly continuous in $H^1(\Omega, \mathbb{C})$. We now prove that S_k is also coercive in $H^1(\Omega, \mathbb{C})$:

$$\begin{aligned} S_k(\psi, \psi) &= \int_\Omega |\nabla\psi|^2 + (\kappa^2 + k_1^2 + k_2^2)\int_\Omega |\psi|^2 + 2\,\mathrm{Im}\left[\int_\Omega \nabla\psi \cdot k\overline{\psi}\right] \\ &\geqslant (1-\delta)\int_\Omega |\nabla\psi|^2 + \left(\kappa^2 + (1 - \frac{1}{\delta})(k_1^2 + k_2^2)\right)\int_\Omega |\psi|^2. \end{aligned} \quad (\text{A.1})$$

If we choose $\delta \in \left(\frac{|k|^2}{|\kappa|^2 + |k|^2}, 1\right)$, which is nonempty since $\kappa \neq 0$, both constants in (A.1) are positive and the sesquilinear form is coercive in $H^1(\Omega, \mathbb{C})$. By the theorem of Lax-Milgram we then find $\psi \in H^1(\Omega, \mathbb{C})$ such that

$$\int_\Omega \nabla_k' \psi \cdot \overline{\nabla_k' \varphi} = \int_\Omega v \cdot \overline{\nabla_k' \varphi} \quad \text{for all } \varphi \in H^1(\Omega, \mathbb{C}).$$

This means that $\nabla_k' \psi =: g \in L^2(\Omega, \mathbb{C}^3)$ and, being a gradient field, also $g \in H(\mathrm{curl}, \Omega)$. Hence, $w := v - g \in W_k$ since $w \in L^2(\Omega, \mathbb{C}^3)$ and

$$\int_\Omega w \cdot \overline{\nabla_k' f} = \int_\Omega v \cdot \overline{\nabla_k' f} - \int_\Omega g \cdot \overline{\nabla_k' f} = \int_\Omega v \cdot \overline{\nabla_k' f} - \int_\Omega \nabla_k' \psi \cdot \overline{\nabla_k' f} = 0.$$

$\square$

A.2. Regularity of $p_j(\cdot, k)$ and of the maps $k \mapsto \omega_j^2(k)$ and $k \mapsto p_j(\cdot, k)$

We prove here some regularity results for the Bloch eigenvalues $(\omega_j^2(k))_j$, $k \in \mathbb{B}$, and eigenfunctions $(q_j(\cdot, k))_j$ and $(p_j(\cdot, k))_j$ (for problems (2.4) and (2.10), respectively) described in sections 2.2 and 2.3. In particular, we show that the choice of our potential by (A1) and (A6), i.e. $0 < \epsilon \in W^{2,\infty}(\mathcal{Q})$, Λ-periodic and with $\epsilon^{-1} \in L^\infty(\mathcal{Q})$, is sufficient to have for each $j \in \mathbb{N}$

(a) $\sup_{k \in \mathbb{B}} \|p_j(\cdot, k)\|_{H^2(\mathcal{Q})} < \infty$ and $\sup_{k \in \mathbb{B}} \|p_j(\cdot, k)\|_{W^{2,\infty}(\mathcal{Q})} < \infty$;
(b) The map $\mathbb{K} \ni k \mapsto p_j(\cdot, k) \in H_\#^2(\mathcal{Q})$ is Lipschitz continuous, provided $\omega_j(k)$ is simple for all $k \in \mathbb{K} \subset \mathbb{B}$.

To this aim, several lemmas will be needed. In the whole section, in addition to the notation introduced in section 2, we denote $\| \cdot \|_{L^2_\#(Q)}$ by $\| \cdot \|_2$. Our method of proof is inspired by that in [7].

Lemma A.2. *The map $\mathbb{B} \ni k \mapsto \omega_j^2(k) \in [0, \infty)$ is Lipschitz continuous.*

Proof. Recall that the sesquilinear form $a_k(\cdot, \cdot)$ of the H-eigenvalue problem is defined as $a_k(\varphi, \psi) := \int_Q \frac{1}{\epsilon} \nabla'_k \times \varphi \cdot \overline{\nabla'_k \times \psi} \, \mathrm{d}x$ for $\varphi, \psi \in V_k$. Since for a fixed $k^0 \in \mathbb{B}$ it is $\nabla'_k = \nabla' + \mathrm{i}k = \nabla'_{k^0} + \mathrm{i}(k - k^0)$, one has

$$a_k(v, v) = a_{k^0}(v, v) + R(v, k, k^0),$$

with

$$
\begin{aligned}
|R(v, k, k^0)| &\leqslant \|\epsilon^{-1}\|_\infty \left(\|v\|_2^2 |k - k^0|^2 + 2\|\nabla'_{k^0} \times v\|_2 \|v\|_2 |k - k^0| \right) \\
&\leqslant c\, |k - k^0| \left(\|v\|_2^2 (|k| + |k^0|) + \|v\|_2 \left(\|\nabla' \times v\|_2 + |k^0|\|v\|_2 \right) \right) \\
&\leqslant c\, |k - k^0| \left(\|\nabla' \times v\|_2^2 + \|v\|_2^2 \right).
\end{aligned}
$$

Using the variational characterization of the eigenvalues

$$\omega_j^2(k) = \min_{\dim S = j} \max_{v \in S} \frac{a_k(v, v)}{\|v\|_2^2},$$

where S is an arbitrary subspace of V_k, we infer

$$
\begin{aligned}
\omega_j^2(k) - \omega_j^2(k^0) &\leqslant \min_{\dim S = j} \max_{v \in S} \frac{R(v, k, k^0)}{\|v\|_2^2} \\
&\leqslant c|k - k^0| \min_{\dim S = j} \max_{v \in S} \left(\frac{\|\nabla' \times v\|_2^2}{\|v\|_2^2} + 1 \right) \\
&= c\, |k - k^0|(\omega_j^2(0) + 1).
\end{aligned}
$$

Interchanging k and k^0, we finally get

$$|\omega_j^2(k) - \omega_j^2(k^0)| \leqslant c\, |k - k^0|.$$

$\square$

Notice that lemma A.2 evidently implies

$$\sup_{k \in \mathbb{B}} |\omega_j^2(k)| \leqslant C \tag{A.2}$$

for all $j \in \mathbb{N}$.

Lemma A.3. *For all $j \in \mathbb{N}$ and p_j defined in (2.9) one has*

$$\sup_{k \in \mathbb{B}} \|p_j(\cdot, k)\|_{H^2(Q)} < \infty. \tag{A.3}$$

Proof. First, by the choice of the normalization of the Bloch eigenfunctions in (2.11), one has

$$\|p_j(\cdot,k)\|_2^2 \leqslant \|\epsilon^{-1}\|_\infty \int_Q \epsilon(x)|p_j(x,k)|^2 \, \mathrm{d}x = \|\epsilon^{-1}\|_\infty.$$

Next, applying the divergence operator $\nabla'_k\cdot$ to (2.10), one finds

$$\nabla'_k \cdot p_j(x,k) = -\epsilon^{-1}(x)\nabla'_k\epsilon(x) \cdot p_j(x,k). \tag{A.4}$$

Noticing that

$$\|\nabla'_k\epsilon\|_\infty \leqslant \|\nabla\epsilon\|_\infty + \max_{k\in\mathbb{B}}(|k|+|\kappa|)\|\epsilon\|_\infty \leqslant c\,\|\epsilon\|_{W^{1,\infty}},$$

one infers

$$\sup_{k\in\mathbb{B}}\|\nabla'_k \cdot p_j(\cdot,k)\|_{L^2(Q)} \leqslant c\,\|\epsilon\|_{W^{1,\infty}}. \tag{A.5}$$

To have a bound on the H^1-norm, we need to estimate also $\nabla'_k \times p_j(\cdot,k)$. We exploit the definition (2.9) and equation (2.8) that $q_j(\cdot,k)$ satisfies in the L^2-sense to get

$$\nabla'_k \times p_j(\cdot,k) = \mathrm{i}\nabla'_k \times \left(\frac{1}{\epsilon}\nabla'_k \times q_j(\cdot,k)\right) = \mathrm{i}\omega_j^2(k)q_j(\cdot,k).$$

Therefore, from (A.2) and the normalization of the eigenfunctions $q_j(\cdot,k)$ we deduce

$$\sup_{k\in\mathbb{B}}\|\nabla'_k \times p_j(\cdot,k)\|_2 \leqslant \sup_{k\in\mathbb{B}}|\omega_j^2(k)|\|q_j(\cdot,k)\|_2 \leqslant c. \tag{A.6}$$

We can thus conclude by (A.5) and (A.6) that the same bound holds also in $H^1_\#(Q)$, i.e.

$$\sup_{k\in\mathbb{B}}\|p_j(\cdot,k)\|_{H^1(Q)} \leqslant c. \tag{A.7}$$

The H^2-norm is estimated similarly since

$$\|p_j(\cdot,k)\|_{H^2(Q)} \leqslant c\,\Big(\|p_j(\cdot,k)\|_{L^2(Q)} + \|\nabla'_k \cdot p_j(\cdot,k)\|_{H^1(Q)}$$
$$+ \|\nabla'_k \times p_j(\cdot,k)\|_{H^1(Q)}\Big). \tag{A.8}$$

First, by (2.10) and (A.2),

$$\|\nabla'_k \times p_j(\cdot,k)\|_{H^1(Q)} \leqslant c\,\left(\|\nabla'_k \times p_j(\cdot,k)\|_2 + \|\nabla'_k \times \nabla'_k \times p_j(\cdot,k)\|_2\right)$$
$$\leqslant c + \|\epsilon\|_\infty|\omega_j^2(k)|\|p_j(\cdot,k)\|_2 \leqslant c. \tag{A.9}$$

Next, from (A.4) we deduce

$$\nabla'_k\left(\nabla'_k \cdot p_j(\cdot,k)\right) = J'_k\left(\frac{\nabla'_k\epsilon}{\epsilon}\right)p_j(\cdot,k) + J'_k(p_j(\cdot,k))\frac{\nabla'_k\epsilon}{\epsilon},$$

where $J'_k(V)$ stands for the Jacobian of the vector field $V:\mathbb{R}^2 \to \mathbb{R}^2$ with the derivatives ∂_m replaced by the 'shifted' derivatives $\partial_m + \mathrm{i}k_m$ for $m \in \{1,2\}$. Hence from (A.7) we have

$$\left\| \nabla'_k \left(\nabla'_k \cdot p_j(\cdot, k) \right) \right\|_{L^2(\mathcal{Q})}$$

$$\leqslant \left\| J'_k \left(\frac{\nabla'_k \epsilon}{\epsilon} \right) \right\|_\infty \| p_j(\cdot, k) \|_2 + \| J'_k(p_j(\cdot, k)) \|_2 \left\| \frac{\nabla'_k \epsilon}{\epsilon} \right\|_\infty$$

$$\leqslant c \left(\| \epsilon \|_{W^{2,\infty}}, \| \epsilon^{-1} \|_\infty \right) \| p_j(\cdot, k) \|_{H^1(\mathcal{Q})} \leqslant c, \tag{A.10}$$

for all $k \in \mathbb{B}$. Combining (A.8) with (A.9) and (A.10), one infers (A.3) and the proof is concluded. $\qquad\square$

Lemma A.4. *For all $j \in \mathbb{N}$ and p_j defined in (2.9) one has*

$$\sup_{k \in \mathbb{B}} \| p_j(\cdot, k) \|_{W^{2,\infty}(\mathcal{Q})} < \infty.$$

Proof. By lemma A.3 and the embedding $H^2(\mathcal{Q}) \hookrightarrow L^\infty(\mathcal{Q})$ we infer $\sup_{k \in \mathbb{B}} \| p_j(\cdot, k) \|_\infty < \infty$. The upgrade to $W^{2,\infty}$-regularity is then accomplished by following the same steps as in the proof of lemma A.3. $\qquad\square$

Next, we aim to prove (b). Let $\mathbb{K}$ be a connected and contractible subset of $\mathbb{B}$ such that $\omega_j(k)$ is simple for all $k \in \mathbb{K}$. Notice that we meet such a condition if $\mathbb{K} = B_\delta(k^{(i)})$ with $j = n_*$, $k^{(i)} \in \{ k^{(1)}, \ldots, k^{(N)} \}$ and $0 < \delta \ll 1$ by assumption (A3). Indeed, the geometric simpleness of $\omega_{n_*}(k^{(i)})$ can be extended to $\omega_{n_*}(k)$ for k in a whole neighbourhood of $k^{(i)}$, see [26, theorem IV.3.16].

As a first step, we prove the following.

Lemma A.5. *The map $\mathbb{K} \ni k \mapsto q_j(\cdot, k) \in L^2_\#(\mathcal{Q})$ is of class C^2.*

In fact, Lipschitz continuity of the map in lemma A.5 will be enough for our purposes in section 4.7.

Proof. Define the operator $A_0(k) := \nabla'_k \times \left(\frac{1}{\epsilon} \nabla'_k \times \right) + a_0 I = L_k^{(H)} + a_0 I$, where a_0 is a positive constant. Since the spectrum of $L_k^{(H)}$ is contained in the non-negative half-line (see section 2.2), the operator $A_0(k)$ is invertible, so in particular $A_0(k)^{-1} : L^2_\#(\mathcal{Q}) \to V_k$, the latter space being the form domain of $L_k^{(H)}$ defined in (2.3). Hence

$$S_j(k) := A_0(k)^{-1} E - \nu_j(k) I, \quad k \in \mathbb{K},$$

where $\nu_j(k) := \left(a_0 + \omega_j^2(k) \right)^{-1}$ and $E : H^1_\#(\mathcal{Q}) \to L^2_\#(\mathcal{Q})$ is the identical embedding, is a well-defined Fredholm operator on $L^2_\#(\mathcal{Q})$ which depends on k in a C^2 fashion. Indeed, E is a compact embedding (see e.g. [1, theorems 3.5, 3.7]) and so $S_j(k)$ is a compact perturbation of (a multiple of) the identity. The C^2-regularity is a consequence of the same property that the map $k \mapsto \omega_j(k)$ enjoys, see assumption (A4). Moreover, it is easy to see that ker $S_j(k)$ coincides with the jth eigenspace of $L_k^{(H)}$ and so, by the geometric simpleness of $\omega_j(k)$, it is of dimension one for all $k \in \mathbb{K}$. This yields the structure of a vector bundle to ker $S_j := \bigcup_{k \in \mathbb{K}} \ker S_j(k)$ over $\mathbb{K}$, see [6, p 62]. Moreover, we claim that the map $k \mapsto \ker S_j(k)$ is C^2.

To this aim let $k^0 \in \mathbb{K}$. Since $S_j(k)$ is a self-adjoint Fredholm operator with a nontrivial kernel for all $k \in \mathbb{K}$, there exists an interval $[-\delta, \delta] \subset \mathbb{R}$, such that $\sigma(S_j(k)) \cap [-\delta, \delta] = \{0\}$.[2]

---

[2] If A is a self-adjoint operator on a Hilbert space and $\lambda \in \mathbb{C}$, then $A - \lambda$ is Fredholm if and only if λ is a discrete eigenvalue of finite multiplicity or lies in the resolvent of A. See also [22, chapter XVII theorem 2.1].

Since $k \mapsto S_j(k)$ is continuous, by spectral continuity [22, chapter II, theorem 4.2], such δ can be chosen independent of k for all $k \in B_r(k^0)$. Consider therefore the map

$$S_j(k) \mapsto P(k) := \frac{1}{2\pi \mathrm{i}} \oint_\Gamma \left(S_j(k) - \lambda\right)^{-1} d\lambda,$$

where Γ is a closed curve in $\mathbb{C}$ that isolates $\{0\}$ from the rest of the spectrum. Then $P(k)$ is a projection onto the eigenspace of the 0 eigenvalue for all $k \in \mathbb{K}$, i.e. $P(k) = \Pi_{\ker(S_j(k))}$, see [26, section 6.4] or [34, theorems XII.5–6]. It is clear then that $k \mapsto P(k)$ is C^2 relying on the same property of $k \mapsto S_j(k)$. Therefore the map $k \mapsto \mathrm{Im}\, P(k) = \ker(S_j(k))$ shares the same regularity too, and the claim is proved.

$\mathbb{K}$ being contractible, the vector bundle $\ker S_j$ is C^2-diffeomorphic to the trivial bundle $\mathbb{K} \times \mathbb{C}$, see e.g. [24, example 2 chapter 4.1], which clearly has a constant section $\tilde{s}_j : \mathbb{K} \to \mathbb{K} \times \mathbb{C}$ such that $\tilde{s}_j(k) = (k, 1)$. Then, calling such diffeomorphism φ_j, the map $s_j : \mathbb{K} \to \ker S_j$ defined as $s_j := \varphi_j^{-1} \circ \tilde{s}_j$ is a C^2 section over $\ker S_j$. This means that, up to a multiplication by a unitary complex function, it is possible to redefine the jth eigenfunction $q_j(\cdot, k)$ normalized as in (2.11) and such that the map $\mathbb{K} \ni k \mapsto q_j(\cdot, k) \in L^2_\#(\mathcal{Q})$ is C^2. $\qquad\square$

Before transferring such a property to the eigenfunctions $p_j(\cdot, k)$, we need a stronger result on $q_j(\cdot, k)$.

Lemma A.6. *The map $\mathbb{K} \ni k \mapsto q_j(\cdot, k) \in H^1_\#(\mathcal{Q})$ is Lipschitz continuous.*

Proof. In other words, we aim to prove that for an arbitrary $k^0 \in \mathbb{K}$ there exists a suitable constant $c(k^0) > 0$ such that

$$\|q_j(\cdot, k) - q_j(\cdot, k^0)\|_{H^1(\mathcal{Q})} \leqslant c\,|k - k^0|, \quad \text{for all } k \in \mathbb{K}. \tag{A.11}$$

Noticing that the Helmholtz decomposition of $H^1_\#(\mathcal{Q})$ in lemma A.1 holds with the operators $\nabla' \times$ and $\nabla' \cdot$, as the particular case when $k = 0$, we estimate separately $\|\nabla' \cdot q_j(\cdot, k) - \nabla' \cdot q_j(\cdot, k^0)\|_2$ and $\|\nabla' \times q_j(\cdot, k) - \nabla' \times q_j(\cdot, k^0)\|_2$. Since $\nabla'_k \cdot q_j(\cdot, k) = 0$ for all $k \in \mathbb{B}$, one has

$$\|\nabla' \cdot q_j(\cdot, k) - \nabla' \cdot q_j(\cdot, k^0)\|_2 = \|k \cdot q_j(\cdot, k) - k^0 \cdot q_j(\cdot, k^0)\|_2$$

$$\leqslant |k|\|q_j(\cdot, k) - q_j(\cdot, k^0)\|_2 - |k - k^0|\|q_j(\cdot, k^0)\|_2$$

$$\leqslant c\,|k - k^0|, \tag{A.12}$$

due to lemma A.5. The estimate for the difference of the curls' is more involved and is based on equation (2.4) which the eigenfunctions satisfy. First,

$$\|\nabla' \times q_j(\cdot, k) - \nabla' \times q_j(\cdot, k^0)\|_2$$

$$= \|\nabla'_k \times q_j(\cdot, k) - \nabla'_{k^0} \times q_j(\cdot, k^0)\|_2 + \|k \times q_j(\cdot, k) - k^0 \times q_j(\cdot, k^0)\|_2, \tag{A.13}$$

where the second term is estimated like above. Noticing that $\nabla'_k = \nabla'_{k^0} + \mathrm{i}(k - k^0)$, we write

$$\|\nabla_k' \times q_j(\cdot, k) - \nabla_{k^0}' \times q_j(\cdot, k^0)\|_2^2$$

$$= \int_{\mathcal{Q}} \nabla_k' \times q_j(\cdot, k) \cdot \overline{\left(\nabla_k' \times q_j(\cdot, k) - \left(\nabla_k' - \mathrm{i}(k - k^0) \right) \times q_j(\cdot, k^0) \right)}$$

$$- \int_{\mathcal{Q}} \nabla_{k^0}' \times q_j(\cdot, k^0) \cdot \overline{\left(\left(\nabla_{k^0}' + \mathrm{i}(k - k^0) \right) \times q_j(\cdot, k) - \nabla_{k^0}' \times q_j(\cdot, k^0) \right)}$$

$$= M_1 - \mathrm{i} M_2, \tag{A.14}$$

where

$$M_1 := \int_{\mathcal{Q}} \nabla_k' \times q_j(\cdot, k) \cdot \overline{\nabla_k' \times (q_j(\cdot, k) - q_j(\cdot, k^0))}$$

$$- \int_{\mathcal{Q}} \nabla_{k^0}' \times q_j(\cdot, k^0) \cdot \overline{\nabla_{k^0}' \times (q_j(\cdot, k) - q_j(\cdot, k^0))} \tag{A.15}$$

and

$$M_2 := \int_{\mathcal{Q}} \nabla_k' \times q_j(\cdot, k) \cdot \overline{(k - k^0) \times q_j(\cdot, k^0)} - \int_{\mathcal{Q}} \nabla_{k^0}' \times q_j(\cdot, k^0) \cdot \overline{(k - k^0) \times q_j(\cdot, k)}.$$

$$\tag{A.16}$$

We estimate M_1 and M_2 separately. First, using (2.4),

$$M_1 \leqslant \|\epsilon^{-1}\|_\infty \left| \int_{\mathcal{Q}} \frac{1}{\epsilon} \nabla_k' \times q_j(\cdot, k) \cdot \overline{\nabla_k' \times (q_j(\cdot, k) - q_j(\cdot, k^0))} \right.$$

$$\left. - \int_{\mathcal{Q}} \frac{1}{\epsilon} \nabla_{k^0}' \times q_j(\cdot, k^0) \cdot \overline{\nabla_{k^0}' \times (q_j(\cdot, k) - q_j(\cdot, k^0))} \right|$$

$$\leqslant c \left| \omega_j^2(k) \int_{\mathcal{Q}} q_j(\cdot, k) \cdot \overline{(q_j(\cdot, k) - q_j(\cdot, k^0))} - \omega_j^2(k^0) \int_{\mathcal{Q}} q_j(\cdot, k^0) \cdot \overline{(q_j(\cdot, k) - q_j(\cdot, k^0))} \right|$$

$$\leqslant c \left| \omega_j^2(k) - \omega_j^2(k^0) \right| \|q_j(\cdot, k^0)\|_2 \|q_j(\cdot, k) - q_j(\cdot, k^0)\|_2 + c \sup_{k \in \mathbb{B}} |\omega_j^2(k)| \|q_j(\cdot, k) - q_j(\cdot, k^0)\|_2^2$$

$$\leqslant c \left| k - k^0 \right|^2, \tag{A.17}$$

where in the last inequality we make use of (A.2) and lemmas A.2 and A.5. Similarly, we also get

$$M_2 \leqslant \|\nabla_k' \times q_j(\cdot, k)\|_2 |k - k^0| \|q_j(\cdot, k) - q_j(\cdot, k^0)\|_2$$

$$+ \|\nabla_k' \times q_j(\cdot, k) - \nabla_{k^0}' \times q_j(\cdot, k^0)\|_2 |k - k^0| \|q_j(\cdot, k^0)\|_2$$

$$\leqslant c \sup_{k \in \mathbb{B}} \|\nabla_k' \times q_j(\cdot, k)\|_2 |k - k^0|^2 + \delta \|\nabla_k' \times q_j(\cdot, k) - \nabla_{k^0}' \times q_j(\cdot, k^0)\|_2^2 + \frac{|k - k^0|^2}{4\delta}$$

$$\tag{A.18}$$

for a small $\delta > 0$. Therefore, combining equations (A.12)–(A.18), we finally infer (A.11). $\quad\square$

We are now in the position to prove (b).

Lemma A.7. *The map* $\mathbb{K} \ni k \mapsto p_j(\cdot, k) \in H_\#^2(\mathcal{Q})$ *is Lipschitz continuous.*

Proof. Notice that lemmas A.2 and A.6 and the definition $p_j(x, k) = \frac{\mathrm{i}}{\epsilon(x)\omega_j(k)}\nabla'_k \times q_j(x, k)$ already imply that the above map with values in $L^2_{\#}(\mathcal{Q})$ is Lipschitz continuous.

First we show that $\mathbb{K} \ni k \mapsto p_j(\cdot, k) \in H^1_{\#}(\mathcal{Q})$ is Lipschitz continuous. Once again we consider the ∇'-Helmholtz decomposition and estimate separately $\|\nabla' \cdot p_j(\cdot, k) - \nabla' \cdot p_j(\cdot, k^0)\|_2$ and $\|\nabla' \times p_j(\cdot, k) - \nabla' \times p_j(\cdot, k^0)\|_2$. Similarly to (A.13) we may confine ourselves to estimate $\|\nabla'_k \cdot p_j(\cdot, k) - \nabla'_{k^0} \cdot p_j(\cdot, k^0)\|_2$ and $\|\nabla'_k \times p_j(\cdot, k) - \nabla'_{k^0} \times p_j(\cdot, k^0)\|_2$.

First, by (A.4) we have

$$\|\nabla'_k \cdot p_j(\cdot, k) - \nabla'_{k^0} \cdot p_j(\cdot, k^0)\|_2 \leqslant \|\epsilon^{-1}\|_\infty \|\nabla'_k \epsilon \cdot p_j(\cdot, k) - \nabla'_{k^0} \epsilon \cdot p_j(\cdot, k^0)\|_2$$
$$\leqslant c|k - k^0|,$$

applying the triangular inequality and lemmas A.2 and A.6. Next, noticing that $\nabla'_k \times p_j(\cdot, k) = \mathrm{i}\omega_j(k) q_j(\cdot, k)$ by (2.8), we may similarly infer the estimate $\|\nabla'_k \times p_j(\cdot, k) - \nabla'_{k^0} \times p_j(\cdot, k^0)\|_2 \leqslant c|k - k^0|$ and, in turn, the Lipschitz continuity in the H^1-norm.

The upgrade to the H^2-norm can be deduced analogously, combining the estimates above with (A.8)–(A.10). □

ORCID iDs

Tomáš Dohnal ⓘ https://orcid.org/0000-0002-0379-5967
Giulio Romani ⓘ https://orcid.org/0000-0002-1563-3148

References

[1] Agmon S 1965 *Lectures on Elliptic Boundary Value Problems* (Princeton, NJ: Van Nostrand-Reinhold) Prepared for Publication by B Frank Jones, Jr With the Assistance of George W Batten, Jr Van Nostrand Mathematical Studies, No. 2

[2] Bao G, Cowsar L and Masters W (ed) 2001 *Mathematical Modeling in Optical Science* (*Frontiers in Applied Mathematics,* vol 22) (Philadelphia, PA: SIAM)

[3] Barton D, Lawrence M, Alaeian H, Baum B and Dionne J 2018 Parity-time symmetric plasmonics *Parity-time Symmetry and its Applications* ed D Christodoulides and J Yang (Berlin: Springer) pp 301–49

[4] Bartsch T, Dohnal T, Plum M and Reichel W 2016 Ground states of a nonlinear curl-curl problem in cylindrically symmetric media *Nonlinear Differ. Equ. Appl.* **23** 1–34

[5] Bender C M 2007 Making sense of non-Hermitian Hamiltonians *Rep. Prog. Phys.* **70** 947–1018

[6] Booss B and Bleecker D D 1985 The Atiyah-Singer index formula and gauge-theoretic physics *Topology and Analysis* (New York: Springer) Translated from the German by D D Bleecker and A Mader

[7] Conca C and Vanninathan M 1997 Homogenization of periodic structures via Bloch decomposition *SIAM J. Appl. Math.* **57** 1639–59

[8] Dauge M, Norton R A and Scheichl R 2015 Regularity for Maxwell eigenproblems in photonic crystal fibre modelling *Bit Numer. Math.* **55** 59–80

[9] Dohnal T and Pelinovsky D 2020 Bifurcation of nonlinear bound states in the periodic Gross–Pitaevskii equation with PT-symmetry *Proc. R. Soc. Edinburgh* A **150** 171–204

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

[11] Dohnal T and Romani G 2021 Eigenvalue bifurcation in doubly nonlinear problems with an application to surface plasmon polaritons *Nonlinear Differ. Equ. Appl.* **28** 30 Paper No. 9

[12] Dohnal T and Siegl P 2016 Bifurcation of eigenvalues in nonlinear problems with antilinear symmetry *J. Math. Phys.* **57** 093502

[13] Dohnal T and Uecker H 2009 Coupled mode equations and gap solitons for the 2D Gross–Pitaevskii equation with a non-separable periodic potential *Phys.* D **238** 860–79

[14] Dohnal T and Uecker H 2009 Erratum to 'coupled mode equations and gap solitons for the 2D Gross–Pitaevskii equation with a non-separable periodic potential' [Physica D 238 (2009) 860–879] *Phys.* D **240** 357–62

[15] Dohnal T and Uecker H 2016 Bifurcation of nonlinear Bloch waves from the spectrum in the Gross–Pitaevskii equation *J. Nonlinear Sci.* **26** 581–618

[16] Dohnal T and Wahlers L 2020 Coupled mode equations and gap solitons in higher dimensions *J. Differ. Equ.* **269** 2386–418

[17] Dohnal T and Wahlers L 2021 Bifurcation of gap solitons in coupled mode equations in d dimensions *J. Dynam. Differ. Equ.*

[18] Dörfler W, Lechleiter A, Plum M, Schneider G and Wieners C 2011 *On the Spectra of Periodic Differential Operators* (Berlin: Springer)

[19] El-Ganainy R, Makris K G, Khajavikhan M, Musslimani Z H, Rotter S and Christodoulides D N 2018 Non-Hermitian physics and PT symmetry *Nat. Phys.* **14** 11–9

[20] Fleischer J W, Segev M, Efremidis N K and Christodoulides D N 2003 Observation of two-dimensional discrete solitons in optically induced nonlinear photonic lattices *Nature* **422** 147–50

[21] Giannoulis J, Mielke A and Sparber C 2008 Interaction of modulated pulses in the nonlinear Schrödinger equation with periodic potential *J. Differ. Equ.* **245** 939–63

[22] Gohberg I, Goldberg S and Kaashoek M A 1990 *Classes of Linear Operators. Vol I* (*Operator Theory: Advances and Applications,* vol 49) (Basel: Birkhäuser)

[23] Goodman R H, Weinstein M I and Holmes P J 2001 Nonlinear propagation of light in one-dimensional periodic structures *J. Nonlinear Sci.* **11** 123–68

[24] Hirsch M W 1994 *Differential Topology* (*Graduate Texts in Mathematics,* vol 33) (New York: Springer) Corrected reprint of the 1976 original

[25] Kartashov Y V, Konotop V V and Zezyulin D A 2014 CPT-symmetric spin-orbit-coupled condensate *Europhys. Lett.* **107** 50002

[26] Kato T 1995 *Perturbation Theory for Linear Operators* (*Classics in Mathematics*) (Berlin: Springer) Reprint of the 1980 edition

[27] Kuchment P 1993 *Floquet Theory for Partial Differential Equations* (*Operator Theory*) (Berlin: Springer)

[28] Lin T-C and Wei J 2005 Ground state of N coupled nonlinear Schrödinger equations in $\mathbb{R}^n$, $n \leqslant 3$ *Commun. Math. Phys.* **255** 629–53

[29] Mandel R 2014 Minimal energy solutions for repulsive nonlinear Schrödinger systems *J. Differ. Equ.* **257** 450–68

[30] Mandel R 2015 Minimal energy solutions for cooperative nonlinear Schrödinger systems *Nonlinear Differ. Equ. Appl.* **22** 239–62

[31] Mederski J 2016 Nonlinear time-harmonic Maxwell equations in $\mathbb{R}^3$: recent results and open questions *Recent Advances in Nonlinear PDEs Theory* (*Lecture Notes of Seminario Interdisciplinare di Matematica,* vol 13) (Potenza: Semin. Interdiscip. Mat. (S.I.M.)) pp 47–57

[32] Oulton R F, Sorger V J, Zentgraf T, Ma R-M, Gladden C, Dai L, Bartal G and Zhang X 2009 Plasmon lasers at deep subwavelength scale *Nature* **461** 629–32

[33] Pankov A 2005 Periodic nonlinear schrödinger equation with application to photonic crystals *Milan J. Math.* **73** 259–87

[34] Reed M and Simon B 1972 *I. Functional Analysis* (*Methods of Modern Mathematical Physics*) (New York: Academic)

[35] Reed M and Simon B 1978 *IV. Analysis of Operators* (*Methods of Modern Mathematical Physics*) (New York: Academic)

[36] Rubinstein J, Sternberg P and Zumbrun K 2010 The resistive state in a superconducting wire: bifurcation from the normal state *Arch. Ration. Mech. Anal.* **195** 117–58

[37] Rüter C E, Makris K G, El-Ganainy R, Christodoulides D N, Segev M and Kip D 2010 Observation of parity-time symmetry in optics *Nat. Phys.* **6** 192–5

[38] Schneider G 1998 Nonlinear stability of Taylor vortices in infinite cylinders *Arch. Ration. Mech. Anal.* **144** 121–200

[39] Schneider G and Uecker H 2001 Nonlinear coupled mode dynamics in hyperbolic and parabolic periodically structured spatially extended systems *Asymptotic Anal.* **28** 163–80

[40] Schweizer B 2013 *Partielle Differentialgleichungen* (Berlin: Springer)

[41] Shen Y R 1984 *The Principles of Nonlinear Optics* (*Pure & Applied Optics Series: 1–349*) (New York: Wiley)

[42] Slusher R E and Eggleton B J 2003 *Nonlinear Photonic Crystals* (*Physics and Astronomy Online Library*) (Berlin: Springer)

[43] Dohnal T and Dörfler W 2013 Coupled mode equation modeling for out-of-plane gap solitons in 2D photonic crystals *Multiscale Model. Simul.* **11** 162–91

[44] Stuart C A 1993 Guidance properties of nonlinear planar waveguides *Arch. Ration. Mech. Anal.* **125** 145–200

[45] Sutherland R L, McLean D G and Kirkpatrick S 2003 *Handbook of Nonlinear Optics* (*Optical Engineering*) (New York: Marcel Dekker)