Bifurcation of Gap Solitons in Coupled Mode Equations in d Dimensions

Tomáš Dohnal[1] · **Lisa Wahlers[2]**

Abstract

We consider a system of first order coupled mode equations in $\mathbb{R}^d$ describing the envelopes of wavepackets in nonlinear periodic media. Under the assumptions of a spectral gap and a generic assumption on the dispersion relation at the spectral edge, we prove the bifurcation of standing gap solitons of the coupled mode equations from the zero solution. The proof is based on a Lyapunov–Schmidt decomposition in Fourier variables and a nested Banach fixed point argument. The reduced bifurcation equation is a perturbed stationary nonlinear Schrödinger equation. The existence of solitary waves follows in a symmetric subspace thanks to a spectral stability result. A numerical example of gap solitons in $\mathbb{R}^2$ is provided.

1 Introduction

First order coupled mode equations (CMEs) are used to describe a class of wavepackets in periodic structures [2,3,5,9–11,15]. They are modulation equations for the envelopes of asymptotically broad and small wavepackets. They were first studied in nonlinear optical fiber gratings, see e.g. [2,3]. A rigorous justification of such an approximation was performed in [15] and [11] for the one dimensional cubic nonlinear wave equation. In [9] the authors derived and justified CMEs as modulation equations for the d-dimensional periodic Gross–Pitaevskii equation. These CMEs have the form

$$\mathrm{i}(\partial_t A_j + v_g^{(j)} \cdot \nabla A_j) + \sum_{r=1}^{N} \kappa_{jr} A_r + N_j(\vec{A}) = 0, \ \ j = 1, \ldots, N, \tag{1.1}$$

✉ Tomáš Dohnal
    tomas.dohnal@mathematik.uni-halle.de

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

2 Fakultät für Mathematik, Technische Universität Dortmund, Vogelpothsweg 87, 44227 Dortmund,
   Germany

 ∅ Springer

where for $j, r \in \{1, \ldots, N\}$

$$N_j(\vec{A}) := \sum_{(m,n,o) \in \{1,\ldots,N\}^3} \gamma_j^{(m,n,o)} A_m \overline{A}_n A_o,$$

$$\gamma_j^{(m,n,o)} \in \mathbb{C}, \kappa_{jr} \in \mathbb{C}, v_g^{(j)} \in \mathbb{R}^d,$$

and where the matrix $\kappa = (\kappa_{jr})_{j,r=1}^N$ is Hermitian. This system (although only for the setting with $\kappa = 0$) was first derived in [10]. Like with all modulation equations, the application of CMEs is not limited to the Gross–Pitaevskii equation. It applies to wavepackets centered around N Bloch waves (N-wave mixing) with nonzero group velocities in models with nonlinearities that are cubic at lowest order.

The aim of this paper is to prove the existence of localized time harmonic solutions

$$\vec{A}(x,t) = e^{-i\omega t} \vec{B}(x), \ |\vec{B}(x)| \to 0 \text{ as } |x| \to \infty, \tag{1.2}$$

with ω in a gap of the linear spatial operator of (1.1). Such solutions are often called (standing) gap solitons. A necessary condition for the existence of a spectral gap is $\kappa \neq 0$. Hence, gap solitons cannot be obtained in the setting of [10]. We prove the existence of gap solitons in an asymptotic region near a spectral edge point ω_0. The result can be interpreted as a bifurcation from the zero solution at the spectral edge.

The equation for $\vec{B}$ is

$$\omega\vec{B} - L(\nabla)\vec{B} + \vec{N}(\vec{B}) = 0, \tag{1.3}$$

where

$$L(\nabla) = \begin{pmatrix} -iv_g^{(1)} \cdot \nabla & & \\ & \ddots & \\ & & -iv_g^{(N)} \cdot \nabla \end{pmatrix} - \kappa.$$

Our proof is constructive in that we use an asymptotic approximation of a solution $\vec{B}$ at $\omega = \omega_0 + O(\varepsilon^2)$, $\varepsilon \to 0$. The approximation is a modulation ansatz with a slowly varying envelope modulating the bounded linear solution at the spectral edge ω_0. The envelope is shown to satisfy a d−dimensional nonlinear Schrödinger equation (NLS) with constant coefficients. We prove that for sufficiently smooth PT symmetric (parity time symmetric) solutions of the NLS there are solutions $\vec{B}$ of (1.3) at $\omega = \omega_0 + O(\varepsilon^2)$ which are close to the asymptotic ansatz. The proof is carried out in Fourier variables in $L^1(\mathbb{R}^d)$. It is based on a decomposition of the solution in Fourier variables according to the eigenvectors of $L(ik) \in \mathbb{C}^{n \times n}$ and on a nested Banach fixed point argument. The reduced bifurcation equation is a perturbed stationary nonlinear Schrödinger equation (NLS). Solitary waves are then found via a persistence argument starting from solitary waves of the unperturbed NLS. The persistence holds in a symmetric subspace thanks to a spectral stability result of Kato.

The chosen approach is similar to that used in [6–8]. Unlike in these papers, where L^2-based spaces were used, we work here in L^1 in Fourier variables. This avoids the unfavorable scaling property of the L^2 norm of functions with an asymptotically slow dependence on x, namely $\|f(\varepsilon \cdot)\|_{L^2(\mathbb{R}^d)} = \varepsilon^{-d/2}\|f\|_{L^2(\mathbb{R}^d)}$. The L^1−approach was first used in this context for the bifurcation of time harmonic gap solitons in the one dimensional wave equation in [13].

The question of the existence of solitary waves of CMEs has previously been addressed only in one dimension in [2], where an explicit family of gap solitons was found for CMEs

describing the asymptotics of wavepackets in media with infinitesimally small contrast. These gap solitons are parametrized by the velocity $v \in (-1, 1)$ (after a rescaling). In [4] a numerical continuation was used to construct gap solitons also in one dimensional CMEs for finite contrast periodic structures. It was shown in [9] that for (1.1) in dimensions $d > 1$ a spectral gap of $L(\nabla)$ does not exist for N odd and for $N = 2$. Next, a gap was found in a special case of (1.1) with $d = 2$, $N = 4$ and standing gap solitons were computed numerically for this case. Here we assume the presence of a spectral gap and prove the existence of standing gap solitons of the form (1.2) for ω asymptotically close to the spectrum under the condition that the spectral edge is given by an isolated extremum of the dispersion relation.

The rest of the paper consists firstly of a formal derivation of the effective NLS equation for the modulation ansatz in Sect. 2. Next, in Sect. 3 we state and prove the main approximation result. Finally, Sect. 4 presents a numerical example of a solution $\vec{B}$ and a numerical verification of the convergence of the asymptotic error.

2 Formal Asymptotics of Gap Solitons

The formal asymptotics of localized solutions of (1.3) were performed already in [9]. We repeat here the calculation for readers' convenience.

The spectrum of $L(\nabla)$ can be determined using Fourier variables. We employ the Fourier transform

$$\widehat{f}(k) := (\mathcal{F}f)(k) := (2\pi)^{-d/2} \int_{\mathbb{R}^d} f(x)e^{-ik\cdot x}\,\mathrm{d}x$$

with the inverse formula $f(x) = (\mathcal{F}^{-1}\widehat{f})(x) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \widehat{f}(k)e^{ik\cdot x}\,\mathrm{d}k$. The spectrum of $L(\nabla)$ is

$$\sigma(L(\nabla)) = \cup_{j\in\{1,\dots,N\}}\lambda_j(\mathbb{R}^d),$$

where $\lambda_j(k)$ is the eigenvalue of $L(ik) \in \mathbb{C}^{N\times N}$ for each $k \in \mathbb{R}^d$, i.e.

$$L(ik)\vec{\eta}^{(j)}(k) = \lambda_j(k)\vec{\eta}^{(j)}(k)$$

for some $\vec{\eta}^{(j)}(k) \in \mathbb{C}^N \setminus \{0\}$. Because $L(ik)^* = L(ik)$ for all $k \in \mathbb{R}^d$, we have $\lambda_j : \mathbb{R}^d \to \mathbb{R}$. The mapping $k \mapsto (\lambda_1(k), \dots, \lambda_N(k))^T$ is the dispersion relation of (1.1).

The central assumptions of our analysis are

(A.1) The spectrum $\sigma(L(\nabla)) \subset \mathbb{R}$ has a gap, denoted by (α, β) with $\alpha < \beta$.
(A.2) $\omega_0 \in \{\alpha, \beta\}$ and for some $j_0 \in \mathbb{N}$, $k_0 \in \mathbb{B}$ we have

$$\omega_0 = \lambda_j(k) \quad \text{if and only if } (j, k) = (j_0, k_0).$$

As mentioned in the introduction, assumption (A.1) implies that if $d \geq 2$, then $N \geq 4$ and N even. Assumption (A.2) means that the spectral edge ω_0 is defined by one isolated extremum of the eigenvalue λ_{j_0} and that this is separated at $k = k_0$ from all other eigenvalues.

We make the following asymptotic ansatz for a gap soliton at $\omega = \omega_0 + \varepsilon^2\omega_1 \notin \sigma(L(\nabla))$, where $\varepsilon > 0$ is a small parameter and $\omega_1 = O(1)$ (as $\varepsilon \to 0$),

$$\vec{B}_{\mathrm{app}}(x) := \varepsilon C(\varepsilon x)e^{ik_0\cdot x}\vec{\eta}^{(j_0)}(k_0). \tag{2.1}$$

In Fourier variables this is

$$\widehat{\vec{B}}_{\mathrm{app}}(k) = \varepsilon^{1-d}\widehat{C}\left(\frac{k - k_0}{\varepsilon}\right)\vec{\eta}^{(j_0)}(k_0). \tag{2.2}$$

Substituting (2.2) and $\omega = \omega_0 + \varepsilon^2 \omega_1$ into the Fourier transform of the left hand side of (1.3), we get

$$(\omega_0 + \varepsilon^2 \omega_1)\widehat{\vec{B}}_{\mathrm{app}} - L(\mathrm{i}k)\widehat{\vec{B}}_{\mathrm{app}} + \widehat{\vec{N}}(\vec{B}_{\mathrm{app}})$$

$$= \varepsilon(\omega_0 + \varepsilon^2 \omega_1 - \lambda_{j_0}(k))\eta^{(j_0)}(k_0)\widehat{C}\left(\frac{k-k_0}{\varepsilon}\right)$$

$$+ \varepsilon^3 \vec{N}(\eta^{(j_0)}(k_0))(\widehat{C} * \widehat{\overline{C}} * \widehat{C})\left(\frac{k-k_0}{\varepsilon}\right)$$

$$= \varepsilon^3 \left[\left(\omega_1 - \left(\frac{k-k_0}{\varepsilon}\right)^T G_0 \frac{k-k_0}{\varepsilon}\right) \widehat{C}\left(\frac{k-k_0}{\varepsilon}\right) \right.$$

$$\left. + \Gamma(\widehat{C} * \widehat{\overline{C}} * \widehat{C})\left(\frac{k-k_0}{\varepsilon}\right) \right] \vec{\eta}^{(j_0)}(k_0) - \vec{\Theta}(k),$$

where

$$G_0 := \frac{1}{2} D^2 \lambda_{j_0}(k_0), \quad \Gamma := \vec{\eta}^{(j_0)}(k_0)^* \widehat{\vec{N}}(\vec{\eta}^{(j_0)}(k_0)),$$

(with v^* being the Hermitian transpose of a vector $v \in \mathbb{C}^N$) and where

$$\vec{\Theta}(k) := \varepsilon^3 \left(\lambda_{j_0}(k) - \omega_0 - \left(\frac{k-k_0}{\varepsilon}\right)^T G_0 \frac{k-k_0}{\varepsilon} \right) \vec{\eta}^{(j_0)}(k_0)\widehat{C}\left(\frac{k-k_0}{\varepsilon}\right)$$

$$+ \varepsilon^3 \left(\widehat{\vec{N}}(\vec{\eta}^{(j_0)}(k_0)) - \vec{\eta}^{(j_0)}(k_0)^* \widehat{\vec{N}}(\vec{\eta}^{(j_0)}(k_0))\vec{\eta}^{(j_0)}(k_0) \right) (\widehat{C} * \widehat{\overline{C}} * \widehat{C})\left(\frac{k-k_0}{\varepsilon}\right)$$

is small as implicitly shown in Sect. 3.

A necessary condition for the smallness of the residual corresponding to $\vec{B}_{\mathrm{app}}$ is the vanishing of the square brackets. This is equivalent to

$$\omega_1 C + \nabla^T (G_0 \nabla C) + \Gamma |C|^2 C = 0 \tag{2.3}$$

for $C : \mathbb{R}^d \to \mathbb{C}$. Equation (2.3) is the effective nonlinear Schrödinger equation (NLS) for the envelope C.

3 The Bifurcation and Approximation Result

Under assumptions (A.1-A.2) and the following assumption (A.3) we prove the bifurcation result below.

(A.3) The kernel of the Jacobian J corresponding to the NLS equation, as defined in (3.22), is $(n+1)$-dimensional (i.e. generated only by the continuous invariances of the NLS).

We define next the space $L_s^1(\mathbb{R}^d)$ for $s \geq 0$ as

$$L_s^1(\mathbb{R}^d) := \{f \in L^1(\mathbb{R}^d) : (1 + |\cdot|)^s f \in L^1(\mathbb{R}^d))\}.$$

For vector valued functions $f : \mathbb{R}^d \to \mathbb{C}^N$ we write $f \in L_s^1(\mathbb{R}^d)$ if $f_j \in L_s^1(\mathbb{R}^d)$ for each $j = 1, \ldots, N$.

The space of continuous functions $f : \mathbb{R}^d \to \mathbb{C}$ satisfying the asymptotics $f(x) \to 0$ as $|x| \to \infty$ is denoted by $C_0(\mathbb{R}^d)$. We equip the space with the supremum norm.

Theorem 1 *Choose ω_0 such that (A.1) and (A.2) are satisfied. Let $(\alpha, \beta) \subset \mathbb{R}$ be the spectral gap from (A.1) and let $\omega_1 \in \mathbb{R}$ be such that $sign(\omega_1) = 1$ if $\omega_0 = \alpha$ and $sign(\omega_1) = -1$ if $\omega_0 = \beta$. If C is a PT-symmetric (i.e. $C(-x) = \overline{C(x)}$) solution of (2.3) with $\widehat{C} \in L_4^1(\mathbb{R}^d)$ and such that (A.3) holds, then there are constants $c_1, c_2, \varepsilon_0 > 0$ such that for each $\varepsilon \in (0, \varepsilon_0)$ there is a solution $\vec{B}$ of Eq. (1.3) with $\omega = \omega_0 + \varepsilon^2 \omega_1$ which satisfies $\widehat{\vec{B}} \in L_2^1(\mathbb{R}^d)$ and*

$$\left\| \widehat{\vec{B}} - \varepsilon^{1-d} \widehat{C} \left(\frac{\cdot - k_0}{\varepsilon} \right) \vec{\eta}^{(j_0)}(k_0) \right\|_{L^1(\mathbb{R}^d)} \leq c_1 \varepsilon^2.$$

In particular,

$$\left\| \vec{B} - \varepsilon C(\varepsilon \cdot) \vec{\eta}^{(j_0)}(k_0) e^{ik_0 \cdot} \right\|_{C_0(\mathbb{R}^d)} \leq c_2 \varepsilon^2.$$

The constants c_1 and c_2 depend polynomially on $\|\widehat{C}\|_{L_4^1(\mathbb{R}^d)}$.

Clearly, due to $\widehat{\vec{B}} \in L^1(\mathbb{R}^d)$ the lemma of Riemann-Lebesgue implies the decay $\vec{B}(x) \to 0$ as $|x| \to \infty$.

The existence of a PT-symmetric solution C is satisfied, e.g., if G_0 is definite and $sign(\Gamma) = -sign(\omega_1)$. Due to the extremum of λ_{j_0} at $k = k_0$ we have then that Γ is positive/negative if $D^2 \lambda_{j_0}(k_0)$ is positive/negative definite respectively. Hence, the NLS is of focusing type and after a rescaling of the x variables it supports a real, positive, radially symmetric solution with exponential decay at infinity (Townes soliton).

Note that the condition on $sign(\omega_1)$ implies

$$\omega = \omega_0 + \varepsilon^2 \omega_1 \notin \sigma(L(\nabla)).$$

We proceed with the proof of Theorem 1. Like in Sect. 2 we work here in Fourier variables. We employ a Lyapunov-Schmidt-like decomposition. For each $k \in \mathbb{R}^d$ we split the solution $\widehat{\vec{B}}(k) \in \mathbb{C}^N$ into the component proportional to the eigenvector $\vec{\eta}^{(j_0)}(k)$ and the $l^2(\mathbb{C}^N)$−orthogonal complement. We define the projections

$$P_k : \mathbb{C}^N \to \text{span} \vec{\eta}^{(j_0)}(k), \ \vec{v} \mapsto (\vec{\eta}^{(j_0)}(k)^* \vec{v}) \vec{\eta}^{(j_0)}(k)$$

and

$$Q_k := I - P_k.$$

Then

$$\widehat{\vec{B}}(k) = \widehat{\vec{B}}_P(k) + \widehat{\vec{B}}_Q(k),$$

where

$$\widehat{\vec{B}}_P(k) = P_k \widehat{\vec{B}}(k) =: \psi(k) \vec{\eta}^{(j_0)}(k), \ \widehat{\vec{B}}_Q(k) = Q_k \widehat{\vec{B}}(k).$$

We aim to construct a solution $\widehat{\vec{B}}$ with ψ approximated by the envelope in our ansatz, i.e. by $\varepsilon^{1-d} \widehat{C} \left(\frac{\cdot - k_0}{\varepsilon} \right)$. We choose for ψ a decomposition according to the support

$$\psi(k) = \varepsilon^{1-d} \widehat{D} \left(\frac{k - k_0}{\varepsilon} \right) + \varepsilon^{1-d} \widehat{R} \left(\frac{k - k_0}{\varepsilon} \right), \tag{3.1}$$

where

$$\text{supp} \widehat{D} \subset B_{\varepsilon^{r-1}} := \left\{ k \in \mathbb{R}^d : |k| < \varepsilon^{r-1} \right\}, \ \text{supp} \widehat{R} \subset B_{\varepsilon^{r-1}}^c := \mathbb{R}^d \setminus B_{\varepsilon^{r-1}}$$

with $r \in (0, 1)$. At the moment r is a free parameter; it will be specified below. We also define

$$\widehat{\overline{B}}_D(k) := \varepsilon^{1-d} \widehat{D}\left(\tfrac{k-k_0}{\varepsilon}\right) \vec{\eta}^{(j_0)}(k), \quad \widehat{\overline{B}}_R(k) := \varepsilon^{1-d} \widehat{R}\left(\tfrac{k-k_0}{\varepsilon}\right) \vec{\eta}^{(j_0)}(k)$$

such that

$$\widehat{\overline{B}}_P = \widehat{\overline{B}}_D + \widehat{\overline{B}}_R.$$

In the ansatz in (3.1) we wish to find the component $\widehat{D}$ close to $\chi_{B_{\varepsilon^{r-1}}} \widehat{C}$ and the component $\widehat{R}$ small. The component $\widehat{\overline{B}}_D$ is then approximated by $\widehat{\overline{B}}_{app}$ in (2.2). If also $\widehat{\overline{B}}_Q$ is small, then the whole constructed solution $\widehat{\overline{B}}$ is close to $\widehat{\overline{B}}_{app}$.

The Fourier transform of (1.1) is

$$\omega \widehat{\overline{B}}(k) - L(ik)\widehat{\overline{B}}(k) + \widehat{N}(\widehat{\overline{B}})(k) = 0, \; k \in \mathbb{R}^d. \tag{3.2}$$

For the selected ansatz Eq. (3.2) becomes

$$\varepsilon^{1-d}(\omega_0 + \varepsilon^2 \omega_1 - \lambda_{j_0}(k)) \left(\widehat{D}\left(\tfrac{k-k_0}{\varepsilon}\right) + \widehat{R}\left(\tfrac{k-k_0}{\varepsilon}\right) \right) + \vec{\eta}^{(j_0)}(k)^* \widehat{N}(\widehat{\overline{B}})(k) = 0, \tag{3.3}$$

$$Q_k \left((\omega_0 + \varepsilon^2 \omega_1) I - L(ik) \right) Q_k \widehat{\overline{B}}_Q(k) + Q_k \widehat{N}(\widehat{\overline{B}})(k) = 0. \tag{3.4}$$

Because $\omega_0 \in \sigma(L(ik_0))$, the inverse of the matrix $(\omega_0 + \varepsilon^2 \omega_1)I - L(ik)$ is not bounded uniformly in ε. In a neighborhood of k_0 the norm of the inverse blows up as $\varepsilon \to 0$. However,

$$M_k := Q_k \left((\omega_0 + \varepsilon^2 \omega_1)I - L(ik) \right) Q_k$$

is invertible uniformly in ε due to assumption (A.2).

We separate the explicit part of (3.4) by writing

$$\widehat{N}(\widehat{\overline{B}})(k) = \widehat{N}(\widehat{\overline{B}}_P)(k) + (\widehat{N}(\widehat{\overline{B}})(k) - \widehat{N}(\widehat{\overline{B}}_P)(k))$$

and

$$\widehat{\overline{B}}_Q = \widehat{\overline{B}}_{Q,1} + \widehat{\overline{B}}_{Q,2},$$

where $\widehat{\overline{B}}_{Q,1}$ solves the explicit part, i.e.

$$\widehat{\overline{B}}_{Q,1}(k) = -M_k^{-1} Q_k \widehat{N}(\widehat{\overline{B}}_P)(k). \tag{3.5}$$

The system to solve is thus

$$\varepsilon^{1-d}(\omega_0 + \varepsilon^2 \omega_1 - \lambda_{j_0}(k)) \left(\widehat{D}\left(\tfrac{k-k_0}{\varepsilon}\right) + \widehat{R}\left(\tfrac{k-k_0}{\varepsilon}\right) \right) + \vec{\eta}^{(j_0)}(k)^* \widehat{N}(\widehat{\overline{B}})(k) = 0 \tag{3.6}$$

$$M_k \widehat{\overline{B}}_{Q,2}(k) + Q_k(\widehat{N}(\widehat{\overline{B}})(k) - \widehat{N}(\widehat{\overline{B}}_P)(k)) = 0. \tag{3.7}$$

Our procedure for constructing a solution can be sketched as follows.

(1) For any $\widehat{D}, \widehat{R} \in L^1(\mathbb{R}^d)$ Eq. (3.5) produces a small $\widehat{\overline{B}}_{Q,1}$ (because $\|\widehat{\overline{B}}_P\|_{L^1} = O(\varepsilon)$).
(2) For any $\widehat{D}, \widehat{R} \in L^1(\mathbb{R}^d)$ and $\widehat{\overline{B}}_{Q,1}$ from step 1 we solve (3.7) by a fixed point argument for a small $\widehat{\overline{B}}_{Q,2}$.
(3) For any $\widehat{D} \in L^1(\mathbb{R}^d)$ and for $\widehat{\overline{B}}_Q$ from steps 1 and 2 we solve (3.6) with $k \in B_{\varepsilon^r}(k_0)^c$ for a small $\widehat{R}$ by a fixed point argument.

(4) With the components obtained in the above steps we find a solution $\widehat{D} \in L_2^1(\mathbb{R}^d)$ of (3.6) with $k \in B_{\varepsilon^r}(k_0)$ close to a $\widehat{C} \in L_2^1(\mathbb{R}^d)$ (with C a solution of (2.3)) - provided such a C exists. In addition C needs to satisfy a certain symmetry, the PT-symmetry. Also here a fixed point argument is used - roughly speaking for the difference $\widehat{D} - \widehat{C}$.

(5) The error $\|\vec{B} - \vec{B}_{\text{app}}\|_{L^1(\mathbb{R}^d)}$ is $O(\varepsilon^2)$ if $\widehat{C}$ decays fast enough, namely if $\widehat{C} \in L_4^1(\mathbb{R}^d)$.

Lemma 1 *If $\widehat{D} \in L^1(\mathbb{R}^d)$, $\widehat{R} \in L_{s_R}^1(\mathbb{R}^d)$ for some $s_R \geq 0$, and $\operatorname{supp}\widehat{D} \subset B_{\varepsilon^{r-1}}$, $\operatorname{supp}\widehat{R} \subset B_{\varepsilon^{r-1}}^c$, then there are constant $c_1, c_2 > 0$ such that for all $\varepsilon > 0$ small enough*

$$\|\widehat{\vec{B}}_P\|_{L^1(\mathbb{R}^d)} \leq c_1 \varepsilon \left(\|\widehat{D}\|_{L^1(B_{\varepsilon^{r-1}})} + \varepsilon^{(1-r)s_R} \|\widehat{R}\|_{L_{s_R}^1(B_{\varepsilon^{r-1}}^c)} \right)$$

and

$$\|\widehat{\vec{N}}(\widehat{\vec{B}}_P)\|_{L^1(\mathbb{R}^d)} \leq c_2 \varepsilon^3 \left(\|\widehat{D}\|_{L^1(B_{\varepsilon^{r-1}})} + \varepsilon^{(1-r)s_R} \|\widehat{R}\|_{L_{s_R}^1(B_{\varepsilon^{r-1}}^c)} \right)^3.$$

Proof Because $\|\widehat{f}(\varepsilon^{-1}(\cdot - k_0))\|_{L^1(\mathbb{R}^d)} = \varepsilon^d \|\widehat{f}\|_{L^1(\mathbb{R}^d)}$, we have

$$\|\widehat{\vec{B}}_P\|_{L^1(\mathbb{R}^d)} \leq c\varepsilon \left(\|\widehat{D}\|_{L^1(\mathbb{R}^d)} + \int_{\mathbb{R}^d} (1 + |\kappa|)^{-s_R} (1 + |\kappa|)^{s_R} |\widehat{R}(\kappa)| \, d\kappa \right)$$

$$\leq c\varepsilon \left(\|\widehat{D}\|_{L^1(\mathbb{R}^d)} + \sup_{\kappa \in B_{\varepsilon^{r-1}}^c} (1 + |\kappa|)^{-s_R} \|\widehat{R}\|_{L_{s_R}^1(B_{\varepsilon^{r-1}}^c)} \right)$$

$$\leq c_1 \varepsilon \left(\|\widehat{D}\|_{L^1(B_{\varepsilon^{r-1}})} + \varepsilon^{(1-r)s_R} \|\widehat{R}\|_{L_{s_R}^1(B_{\varepsilon^{r-1}}^c)} \right).$$

The estimate for $\|\widehat{\vec{N}}(\widehat{\vec{B}}_P)\|_{L^1(\mathbb{R}^d)}$ follows by Young's inequality for convolutions. $\qquad \square$

(1) **Component $\widehat{\vec{B}}_{Q,1}$**

Because $\widehat{\vec{B}}_{Q,1} = -M_k^{-1} Q_k \widehat{\vec{N}}(\widehat{\vec{B}}_P)(k)$, we get from Lemma 1 the estimate

$$\|\widehat{\vec{B}}_{Q,1}\|_{L^1(\mathbb{R}^d)} \leq c\varepsilon^3 (\|\widehat{D}\|_{L^1(\mathbb{R}^d)} + \|\widehat{R}\|_{L^1(\mathbb{R}^d)})^3. \tag{3.8}$$

(2) **Component $\widehat{\vec{B}}_{Q,2}$**

With $\widehat{\vec{B}}_{Q,1}$ from above (and $\widehat{D}$, $\widehat{R}$ given) component $\widehat{\vec{B}}_{Q,2}$ satisfies

$$\widehat{\vec{B}}_{Q,2}(k) = M_k^{-1} Q_k \left(\widehat{\vec{N}}(\widehat{\vec{B}}_P)(k) - \widehat{\vec{N}}(\widehat{\vec{B}}_P + \widehat{\vec{B}}_{Q,1} + \widehat{\vec{B}}_{Q,2})(k) \right) =: \vec{G}(\widehat{\vec{B}}_{Q,2})(k).$$

Due to the cubic structure of $\vec{N}$ we have

$$\|\vec{G}(\widehat{\vec{B}}_{Q,2})\|_{L^1(\mathbb{R}^d)} \leq c\|\widehat{\vec{B}}_{Q,1} + \widehat{\vec{B}}_{Q,2}\|_{L^1(\mathbb{R}^d)} \big(\|\widehat{\vec{B}}_P\|_{L^1(\mathbb{R}^d)} + \|\widehat{\vec{B}}_{Q,1}\|_{L^1(\mathbb{R}^d)} + \|\widehat{\vec{B}}_{Q,2}\|_{L^1(\mathbb{R}^d)} \big)^2.$$

For $\widehat{\vec{B}}_{Q,2} \in B_{\varepsilon^\eta}$, $\eta > 0$ we have

$$\|\vec{G}(\widehat{\vec{B}}_{Q,2})\|_{L^1(\mathbb{R}^d)} \leq c(\|\widehat{D}\|_{L^1}, \|\widehat{R}\|_{L^1})(\varepsilon^5 + \varepsilon^{2+\eta} + \varepsilon^{1+2\eta} + \varepsilon^{3\eta}).$$

Hence, $\vec{G} : B_{c_0\varepsilon^5}^{(L^1)} \to B_{c_0\varepsilon^5}^{(L^1)}$ for some $c_0(\|\widehat{D}\|_{L^1}, \|\widehat{R}\|_{L^1}) > 0$, where $B_\alpha^{(L^1)} := \{\vec{v} \in L^1(\mathbb{R}^d) : \|\vec{v}\|_{L^1(\mathbb{R}^d)} \leq \alpha\}$. The constants c and c_0 depend polynomially on $\|\widehat{D}\|_{L^1}$ and $\|\widehat{R}\|_{L^1}$.

Similarly, we obtain the contraction (for $\varepsilon > 0$ small enough)

$$\|\vec{G}(\widehat{\vec{B}}_{Q,R}^{(1)}) - \vec{G}(\widehat{\vec{B}}_{Q,R}^{(2)})\|_{L^1(\mathbb{R}^d)} \le c\varepsilon^2 \|\widehat{\vec{B}}_{Q,R}^{(1)} - \widehat{\vec{B}}_{Q,R}^{(2)}\|_{L^1(\mathbb{R}^d)},$$

if $\widehat{\vec{B}}_{Q,R}^{(1)}, \widehat{\vec{B}}_{Q,R}^{(2)} \in B_{c_0\varepsilon^5}^{(L^1)}$. Hence, for $\varepsilon > 0$ small enough we have a unique solution $\widehat{\vec{B}}_{Q,2} \in B_{c_0\varepsilon^5}^{(L^1)}$ of $\widehat{\vec{B}}_{Q,2} = \vec{G}(\widehat{\vec{B}}_{Q,2})$, i.e.

$$\|\widehat{\vec{B}}_{Q,2}\|_{L^1(\mathbb{R}^d)} \le c_0(\|\widehat{D}\|_{L^1}, \|\widehat{R}\|_{L^1})\varepsilon^5. \tag{3.9}$$

(3) Component $\widehat{\vec{B}}_R$

For any $\widehat{D} \in L^1(\mathbb{R}^d)$ and with the above estimate on $\widehat{\vec{B}}_Q$ we look for a small $\widehat{R}$. The support of $\widehat{R}$ is $B_{\varepsilon^r-1}^c$, whence for $k \in \mathbb{R}^d \setminus B_{\varepsilon^r}(k_0)$ we can divide in (3.6) by $\omega_0 + \varepsilon^2\omega_1 - \lambda_{j_0}(k)$ and obtain

$$\widehat{R}\left(\frac{k - k_0}{\varepsilon}\right) = \varepsilon^{d-1}\nu(k)\vec{\eta}^{(j_0)}(k)^* \widehat{\vec{N}}(\widehat{\vec{B}})(k), \quad k \in \mathbb{R}^d \setminus B_{\varepsilon^r}(k_0)$$

with

$$\nu(k) := (\lambda_{j_0}(k) - \omega_0 - \varepsilon^2\omega_1)^{-1}.$$

Since $\nabla\lambda_{j_0}(k_0) = 0$, we have $|\lambda_{j_0}(k) - \omega_0| > c\varepsilon^{2r}$ for all $k \in \mathbb{R}^d \setminus B_{\varepsilon^r}(k_0)$ and hence $|\nu(k)| \le c\varepsilon^{-2r}$. In order to exploit the localized nature of $\widehat{\vec{B}}_D$ and the smallness of $\widehat{\vec{B}}_Q$, we write for $\kappa := \frac{k-k_0}{\varepsilon}$

$$\widehat{R}(\kappa) = H(\widehat{R})(\kappa),$$

where

$$H(\widehat{R})(\kappa) := \varepsilon^{d-1}\nu(k_0 + \varepsilon\kappa)\Big[h(k_0 + \varepsilon\kappa)^{-1}h(k_0 + \varepsilon\kappa)\vec{\eta}^{(j_0)}(k_0 + \varepsilon\kappa)^* \widehat{\vec{N}}(\widehat{\vec{B}}_D)(k_0 + \varepsilon\kappa)$$
$$+\vec{\eta}^{(j_0)}(k_0 + \varepsilon\kappa)^* \Big(\widehat{\vec{N}}(\widehat{\vec{B}})(k_0 + \varepsilon\kappa) - \widehat{\vec{N}}(\widehat{\vec{B}}_D)(k_0 + \varepsilon\kappa)\Big)\Big] \tag{3.10}$$

with $h(k) := \left(1 + \frac{|k-k_0|}{\varepsilon}\right)^{s_D}$. Using $\sup_{k \in \text{supp}\widehat{\vec{B}}_R} |h(k)^{-1}| \le c\varepsilon^{(1-r)s_D}$, the cubic form of $\widehat{\vec{N}}$ and the fact that $\left(\widehat{D}\left(\frac{\cdot - k_0}{\varepsilon}\right) * \widehat{\vec{D}}\left(\frac{\cdot + k_0}{\varepsilon}\right) * \widehat{D}\left(\frac{\cdot - k_0}{\varepsilon}\right)\right)(k_0 + \varepsilon\kappa) = \varepsilon^{2d}(\widehat{D} * \widehat{\vec{D}} * \widehat{D})(\kappa)$, we get

$$\|H(\widehat{R})\|_{L^1(\mathbb{R}^d)} \le c_1\varepsilon^{2-2r+(1-r)s_D}\|(1 + |\cdot|)^{s_D}(\widehat{D} * \widehat{\vec{D}} * \widehat{D})(\cdot)\|_{L^1(\mathbb{R}^d)}$$
$$+ c_2\varepsilon^{-2r-1}\left(\|\widehat{\vec{B}}_R\|_{L^1(\mathbb{R}^d)} + \|\widehat{\vec{B}}_Q\|_{L^1(\mathbb{R}^d)}\right)$$
$$\times \left(\|\widehat{\vec{B}}_D\|_{L^1(\mathbb{R}^d)} + \|\widehat{\vec{B}}_R\|_{L^1(\mathbb{R}^d)} + \|\widehat{\vec{B}}_Q\|_{L^1(\mathbb{R}^d)}\right)^2.$$

Since $\|\widehat{\vec{B}}_R\|_{L^1(\mathbb{R}^d)} \le c\varepsilon\|\widehat{R}\|_{L^1(\mathbb{R}^d)}$, $\|\widehat{\vec{B}}_D\|_{L^1(\mathbb{R}^d)} \le c\varepsilon\|\widehat{D}\|_{L^1(\mathbb{R}^d)}$, and $\|\widehat{\vec{B}}_Q\|_{L^1(\mathbb{R}^d)} \le c(\|\widehat{D}\|_{L^1}, \|\widehat{R}\|_{L^1})\varepsilon^3$ (with a polynomial c), we have

$$\|H(\widehat{R})\|_{L^1(\mathbb{R}^d)}$$
$$\le c\varepsilon^{2-2r}\left(\varepsilon^{(1-r)s_D}\|\widehat{D}\|_{L_{s_D}^1(\mathbb{R}^d)}^3 + \|\widehat{R}\|_{L^1(\mathbb{R}^d)}^3 + \|\widehat{R}\|_{L^1(\mathbb{R}^d)}^2 + \|\widehat{R}\|_{L^1(\mathbb{R}^d)} + \varepsilon^2\right)$$

with c depending polynomially on $\|\widehat{D}\|_{L^1}$ and $\|\widehat{R}\|_{L^1}$.

If $\widehat{D} \in L^1_{s_D}(\mathbb{R}^d)$, $s_D \geq 0$, then there is a constant c depending polynomially on $\|\widehat{D}\|_{L^1_{s_D}(\mathbb{R}^d)}$ such that

$$\|H(\widehat{R})\|_{L^1(\mathbb{R}^d)} \leq c(\|\widehat{D}\|_{L^1_{s_D}})\varepsilon^\alpha, \quad \alpha := \min\{4 - 2r, (2 + s_D)(1 - r)\} \qquad (3.11)$$

for all R with $\|\widehat{R}\|_{L^1(\mathbb{R}^d)} \leq c\varepsilon^\alpha$ and all $\varepsilon > 0$ small enough. Hence $H : B^{(L^1)}_{c\varepsilon^\alpha} \to B^{(L^1)}_{c\varepsilon^\alpha}$ for some $c > 0$ if $\varepsilon > 0$ is small enough and if $\widehat{D} \in L^1_{s_D}(\mathbb{R}^d)$.

Similarly, we get the contraction property of H on $B^{(L^1)}_{c\varepsilon^\alpha}$ for $\varepsilon > 0$ small enough. The constructed fixed point $\widehat{R} \in B^{(L^1)}_{c\varepsilon^\alpha}$ yields for any $s_D \geq 0$

$$\|\widehat{\vec{B}}_R\|_{L^1(\mathbb{R}^d)} \leq c(\|\widehat{D}\|_{L^1_{s_D}})\varepsilon^{\alpha+1}. \qquad (3.12)$$

(4) Component $\widehat{\vec{B}}_D$

Finally, we consider the component $\widehat{\vec{B}}_D$. For $k \in \mathrm{supp}(\widehat{\vec{B}}_D) = B_{\varepsilon^r}(k_0)$ we rewrite (3.6) as follows. We add and subtract $\widehat{N}(\vec{B}_D)$ like in (3.10), we Taylor expand $\lambda_{j_0}(k)$ at $k = k_0$, and we use the variable $\kappa = \varepsilon^{-1}(k - k_0)$. This leads to

$$(\omega_1 - \kappa^T G_0 \kappa)\widehat{D}(\kappa) + \chi_{B_{\varepsilon^{r-1}}}(\kappa)\Gamma(\widehat{D} * \overline{\widehat{D}} * \widehat{D})(\kappa) = \rho(\widehat{D})(\kappa), \qquad (3.13)$$

where

$$\rho(\widehat{D})(\kappa) = \chi_{B_{\varepsilon^{r-1}}}(\kappa)\left[\Gamma(\widehat{D} * \overline{\widehat{D}} * \widehat{D})(\kappa) - \varepsilon^{d-3}\vec{\eta}^{(j_0)}(k_0 + \varepsilon\kappa)^*\widehat{\vec{N}}(\vec{B}_D)(k_0 + \varepsilon\kappa)\right]$$
$$+ \varepsilon^{-2}\left[\delta(k_0 + \varepsilon\kappa)\widehat{D}(\kappa) + \varepsilon^{d-1}\chi_{B_{\varepsilon^{r-1}}}(\kappa)\vec{\eta}^{(j_0)}(k_0 + \varepsilon\kappa)^*\right.$$
$$\left.\left(\widehat{\vec{N}}(\vec{B})(k_0 + \varepsilon\kappa) - \widehat{\vec{N}}(\vec{B}_D)(k_0 + \varepsilon\kappa)\right)\right]$$

and $\delta(k) := \lambda_{j_0}(k) - \omega_0 - \frac{1}{2}(k - k_0)^T D^2\lambda_{j_0}(k_0)(k - k_0)$.

Next, we write $\rho = \rho_1 + \rho_2 + \rho_3$, where

$$\rho_1(\widehat{D})(\kappa) := \chi_{B_{\varepsilon^{r-1}}}(\kappa)\left[\Gamma(\widehat{D} * \overline{\widehat{D}} * \widehat{D})(\kappa) - \varepsilon^{d-3}\vec{\eta}^{(j_0)}(k_0 + \varepsilon\kappa)^*\widehat{\vec{N}}(\vec{B}_D)(k_0 + \varepsilon\kappa)\right],$$

$$\rho_2(\widehat{D})(\kappa) := \varepsilon^{-2}\delta(k_0 + \varepsilon\kappa)\widehat{D}(\kappa),$$

$$\rho_3(\widehat{D})(\kappa) := \varepsilon^{d-3}\chi_{B_{\varepsilon^{r-1}}}(\kappa)\vec{\eta}^{(j_0)}(k_0 + \varepsilon\kappa)^*\left(\widehat{\vec{N}}(\vec{B})(k_0 + \varepsilon\kappa) - \widehat{\vec{N}}(\vec{B}_D)(k_0 + \varepsilon\kappa)\right).$$

For ρ_1 we get

$$\rho_1(\widehat{D})(\kappa) = \varepsilon^{-2d}\sum_{\substack{m,n,o\in\{1,\dots,N\}\\j\in\{1,\dots,N\}}}\gamma_j^{(m,n,o)}\int_{B_{2\varepsilon^r}}\int_{B_{\varepsilon^r}(k_0)}\left(\overline{\eta}_j^{(j_0)}(k_0 + \varepsilon\kappa)\right.$$

$$\times \eta_m^{(j_0)}(k_0 + \varepsilon\kappa - s)\overline{\eta}_n^{(j_0)}(s - t)\eta_o^{(j_0)}(t) - \Gamma\Big)$$

$$\times \widehat{D}\left(\kappa - \frac{s}{\varepsilon}\right)\overline{\widehat{D}}\left(\frac{s - t + k_0}{\varepsilon}\right)\widehat{D}\left(\frac{t - k_0}{\varepsilon}\right)dt\,ds$$

$$= \sum_{\substack{m,n,o\in\{1,\dots,N\}\\j\in\{1,\dots,N\}}}\gamma_j^{(m,n,o)}\int_{B_{2\varepsilon^{r-1}}}\int_{B_{\varepsilon^{r-1}}}\left(\overline{\eta}_j^{(j_0)}(k_0 + \varepsilon\kappa)\right.$$

$$\times \eta_m^{(j_0)}(k_0 + \varepsilon(\kappa - \tilde{s}))\overline{\eta}_n^{(j_0)}(k_0 + \varepsilon(\tilde{s} - \tilde{t}))$$

$$\times \eta_o^{(j_0)}(k_0 + \varepsilon\tilde{t}) - \Gamma\Big)\widehat{D}(\kappa - \tilde{s})\overline{\widehat{D}}(\tilde{s} - \tilde{t})\widehat{D}(\tilde{t})\,d\tilde{t}\,d\tilde{s}.$$

Because $\Gamma = \sum_{m,n,o,j \in \{1,\dots,N\}} \gamma_j^{(m,n,o)} \overline{\eta}_j^{(j_0)}(k_0) \eta_m^{(j_0)}(k_0) \overline{\eta}_n^{(j_0)}(k_0) \eta_o^{(j_0)}(k_0)$, we get

$$\left| \sum_{m,n,o,j \in \{1,\dots,N\}} \gamma_j^{(m,n,o)} \overline{\eta}_j^{(j_0)}(k_0 + \varepsilon\kappa) \eta_m^{(j_0)}(k_0 + \varepsilon(\kappa - \tilde{s})) \overline{\eta}_n^{(j_0)}(k_0 + \varepsilon(\tilde{s} - \tilde{t})) \eta_o^{(j_0)}(k_0 + \varepsilon\tilde{t}) - \Gamma \right|$$
$$\leq c\varepsilon(|\kappa| + |\kappa - \tilde{s}| + |\tilde{s} - \tilde{t}| + |\tilde{t}|)$$

due to the Lipschitz continuity of $k \mapsto \eta^{(j_0)}(k)$. Hence, by Young's inequality for convolutions,

$$\|\rho_1(\widehat{D})\|_{L^1(\mathbb{R}^d)} \leq c\varepsilon \|\widehat{D}\|_{L_1^1(\mathbb{R}^d)}^3. \tag{3.14}$$

For ρ_2 we note that $|\delta(k_0 + \varepsilon\kappa)| \leq c\varepsilon^3 |\kappa|^3$ for $\kappa \in B_{\varepsilon^{r-1}}$ and $\varepsilon > 0$ small enough. We estimate

$$\|\rho_2(\widehat{D})\|_{L^1(\mathbb{R}^d)} \leq c\varepsilon \int_{B_{\varepsilon^{r-1}}} |\kappa|^3 |\widehat{D}(\kappa)| \, d\kappa \leq c\varepsilon \sup_{\kappa \in B_{\varepsilon^{r-1}}} |\kappa|^\beta \int_{\mathbb{R}^d} |\kappa|^{3-\beta} |\widehat{D}(\kappa)| \, d\kappa$$
$$\leq c\varepsilon^{1-\beta(1-r)} \|\widehat{D}\|_{L_{3-\beta}^1(\mathbb{R}^d)} \tag{3.15}$$

for any $\beta \in [0, 3)$.

Finally, we estimate ρ_3. Note that $\widehat{\widehat{N}}(\widehat{\widehat{B}})(k_0 + \varepsilon\cdot) - \widehat{\widehat{N}}(\widehat{\widehat{B}}_D)(k_0 + \varepsilon\cdot)$ appears also in (3.10). We have

$$\|\widehat{\widehat{N}}(\widehat{\widehat{B}})(k_0 + \varepsilon\cdot) - \widehat{\widehat{N}}(\widehat{\widehat{B}}_D)(k_0 + \varepsilon\cdot)\|_{L^1(\mathbb{R}^d)}$$
$$\leq \varepsilon^{-d} \left(\|\widehat{\widehat{B}}_R\|_{L^1(\mathbb{R}^d)} + \|\widehat{\widehat{B}}_Q\|_{L^1(\mathbb{R}^d)} \right) \left(\|\widehat{\widehat{B}}_D\|_{L^1(\mathbb{R}^d)} + \|\widehat{\widehat{B}}_R\|_{L^1(\mathbb{R}^d)} + \|\widehat{\widehat{B}}_Q\|_{L^1(\mathbb{R}^d)} \right)^2$$
$$\leq c_1(\|\widehat{D}\|_{L_{s_D}^1}) \varepsilon^{-d} \left((\varepsilon^{\alpha+3} + \varepsilon^5) \|\widehat{D}\|_{L^1(\mathbb{R}^d)}^2 + (\varepsilon^{2\alpha+3} + \varepsilon^7) \|\widehat{D}\|_{L^1(\mathbb{R}^d)} + \varepsilon^{3\alpha+3} + \varepsilon^9 \right)$$
$$\leq c_2(\|\widehat{D}\|_{L_{s_D}^1})(\varepsilon^{\alpha+3-d} + \varepsilon^{5-d})$$

for $\varepsilon > 0$ small enough, where we have made use of (3.8), (3.9), (3.12), the fact that $\|\widehat{\widehat{B}}_D\|_{L^1} \leq \varepsilon \|\widehat{D}\|_{L^1}$ and the estimate $\|\widehat{R}\|_{L^1} \leq c(\|\widehat{D}\|_{L_{s_D}^1}) \varepsilon^\alpha$. The dependence of c_1 and c_2 on $\|\widehat{D}\|_{L_{s_D}^1}$ is polynomial. As a result

$$\|\rho_3(\widehat{D})\|_{L^1(\mathbb{R}^d)} \leq c(\|\widehat{D}\|_{L_{s_D}^1})(\varepsilon^\alpha + \varepsilon^2). \tag{3.16}$$

The whole right hand side of (3.13) is thus estimated as

$$\|\rho(\widehat{D})\|_{L^1(\mathbb{R}^d)} \leq c(\|\widehat{D}\|_{L_{s_D}^1}, \|\widehat{D}\|_{L_{3-\beta}^1}) \left(\varepsilon^{1-\beta(1-r)} + \varepsilon^\alpha \right)$$

for any $\beta \in [0, 3)$. Once again, the constant c depends polynomially on its arguments.

In order to solve (3.13) for $\widehat{D}$ below (using a fixed point argument), we need to consider ρ as an in-homogeneity. The linearized operator to be inverted in the iteration is of second order such that in Fourier space it acts from $L_2^1(\mathbb{R}^d)$ to $L^1(\mathbb{R}^d)$. For that reason we need to choose $s_D = 2$ and $\beta \geq 1$ above. The choice $s_D = 2$, $\beta = 1$, $r = 1/2$ leads to $\alpha = 2$ and $1 - \beta(1 - r) = 1/2$, i.e. $\|\rho(\widehat{D})\|_{L^1(\mathbb{R}^d)} \leq c\varepsilon^{1/2}$ provided $\widehat{D} \in L_2^1(\mathbb{R}^d)$. For $s_D = 2$, $\beta = 1$ the largest value of $\min\{1 - \beta(1 - r), \alpha\}$ is $4/5$ attained at $r = 4/5$. Hence, with $r = 4/5$ we get the best possible estimate

$$\|\rho(\widehat{D})\|_{L^1(\mathbb{R}^d)} \leq c(\|\widehat{D}\|_{L_2^1}) \varepsilon^{4/5}.$$

This order determines the accuracy of the approximation and turns out to be insufficient. It leads to $\|\widehat{\widehat{B}} - \widehat{\widehat{B}}_{\text{app}}\|_{L^1(\mathbb{R}^d)} \leq c\varepsilon^{9/5}$ instead of $c\varepsilon^2$.

Clearly, the leading order term in the residual ρ is caused by the error from the Taylor expansion of λ_{j_0}. We introduce a refined ansatz for $\widehat{D}$ in order to make this error of higher order. Note that Eq. (3.13) is a perturbation of the NLS (2.3). Writing

$$f_{\text{NLS}}(C) := \omega_1 C + \nabla^T (G_0 \nabla C) + \Gamma |C|^2 C,$$

Eq. (3.13) is

$$\chi_{B_{\varepsilon^{r-1}}} \widehat{f_{\mathrm{NLS}}(D)} = \rho(\widehat{D}). \qquad (3.17)$$

We search for D close to a solution C of the NLS, i.e. of $f_{\mathrm{NLS}}(C) = 0$. For that we need to look for $\widehat{D}$ in a vicinity of

$$\widehat{C}^{(\varepsilon)} := \chi_{B_{\varepsilon^{r-1}}} \widehat{C}.$$

We choose the following ansatz

$$\widehat{D}(\kappa) = \left(1 + \varepsilon \frac{D^3 \lambda_{j_0}(k_0)(\kappa, \kappa, \kappa)}{6(\omega_1 - \kappa^T G_0 \kappa)}\right) \widehat{C}^{(\varepsilon)}(\kappa) + \widehat{d}(\kappa), \qquad (3.18)$$

where $D^3 \lambda_{j_0}(k_0)(\kappa, \kappa, \kappa)$ is the third order term in the Taylor expansion of $\varepsilon^{-3} \lambda_{j_0}(k_0 + \varepsilon \kappa)$ and where $\mathrm{supp}(\widehat{d}) \subset B_{\varepsilon^{r-1}}$. We look for a solution with a small $\widehat{d}$.

We also define

$$\widehat{D}_0 := \widehat{C}^{(\varepsilon)} + \widehat{d}, \quad \nu(\kappa) := \frac{D^3 \lambda_{j_0}(k_0)(\kappa, \kappa, \kappa)}{6(\omega_1 - \kappa^T G_0 \kappa)},$$

such that $\widehat{D} = \widehat{D}_0 + \varepsilon \nu \widehat{C}^{(\varepsilon)}$. With this notation Eq. (3.17) reads

$$\chi_{B_{\varepsilon^{r-1}}} \widehat{f_{\mathrm{NLS}}(D_0)} = \rho(\widehat{D}) + \chi_{B_{\varepsilon^{r-1}}} \left(\widehat{f_{\mathrm{NLS}}(D_0)} - \widehat{f_{\mathrm{NLS}}(D)}\right) =: \tilde{\rho}(\widehat{D}). \qquad (3.19)$$

Compared to ρ the right hand side $\tilde{\rho}$ is smaller as we show next. It is

$$\tilde{\rho}(\widehat{D}) = \rho_1(\widehat{D}) + \rho_3(\widehat{D}) + \tilde{\rho}_2(\widehat{D}),$$

where

$$\tilde{\rho}_2(\widehat{D})(\kappa) = \varepsilon^{-2} \left(\lambda_{j_0}(k_0 + \varepsilon \kappa) - \omega_0 - \varepsilon^2 \kappa^T G_0 \kappa\right) \widehat{D}(\kappa) - \varepsilon(\omega_1 - \kappa^T G_0 \kappa) \nu(\kappa) \widehat{C}^{(\varepsilon)}(\kappa)$$

$$+ \chi_{B_{\varepsilon^{r-1}}}(\kappa) \Gamma \left[\widehat{D}_0 * \overline{\widehat{D}_0} * \widehat{D}_0 - \widehat{D} * \overline{\widehat{D}} * \widehat{D}\right](\kappa)$$

$$= \varepsilon^{-2} \left(\lambda_{j_0}(k_0 + \varepsilon \kappa) - \omega_0 - \varepsilon^2 \kappa^T G_0 \kappa - \frac{\varepsilon^3}{6} D^3 \lambda_{j_0}(k_0)(\kappa, \kappa, \kappa)\right) \widehat{C}^{(\varepsilon)}(\kappa)$$

$$+ \varepsilon^{-2} \left(\lambda_{j_0}(k_0 + \varepsilon \kappa) - \omega_0 - \varepsilon^2 \kappa^T G_0 \kappa\right) \left(\varepsilon \nu(\kappa) \widehat{C}^{(\varepsilon)}(\kappa) + \widehat{d}(\kappa)\right)$$

$$+ \chi_{B_{\varepsilon^{r-1}}}(\kappa) \Gamma \left[\widehat{D}_0 * \overline{\widehat{D}_0} * \widehat{D}_0 - \widehat{D} * \overline{\widehat{D}} * \widehat{D}\right](\kappa).$$

To estimate $\tilde{\rho}_2$ note that the first Taylor expansion error term (i.e. the first line) can be estimated in $L^1(\mathbb{R}^d)$ by $c \varepsilon^2 \|\widehat{C}^{(\varepsilon)}\|_{L_4^1(\mathbb{R}^d)} \leq c \varepsilon^2 \|\widehat{C}\|_{L_4^1(\mathbb{R}^d)}$. For the second Taylor expansion error note first that $|\nu(\kappa)| \leq c(1 + |\kappa|)$ for all $\kappa \in \mathbb{R}^d$ such that

$$\|\varepsilon^{-2} \left(\lambda_{j_0}(k_0 + \varepsilon \cdot) - \omega_0 - \varepsilon^2 (\cdot)^T G_0(\cdot)\right)(\varepsilon \nu \widehat{C}^{(\varepsilon)} + \widehat{d})\|_{L^1(\mathbb{R}^d)}$$

$$\leq c \left(\varepsilon^2 \|\widehat{C}\|_{L_4^1(\mathbb{R}^d)} + \varepsilon \int_{\mathbb{R}^d} |\kappa|^3 |\widehat{d}(\kappa)| \, d\kappa\right).$$

Just like in (3.15) we get

$$\varepsilon \int_{\mathbb{R}^d} |\kappa|^3 |\widehat{d}(\kappa)| \, d\kappa \leq c \varepsilon^{1 - \beta(1 - r)} \|\widehat{d}\|_{L_{3-\beta}^1(\mathbb{R}^d)}$$

for any $\beta \in [0, 3)$. Once again, we need to choose $\beta \geq 1$ in order for $\tilde{\rho}_2 : \widehat{d} \mapsto \tilde{\rho}_2(\widehat{d})$ to be a mapping from $L_2^1(\mathbb{R}^d)$ to $L^1(\mathbb{R}^d)$. With $\beta = 1$ and $r = 1/2$ we get $1 - \beta(1 - r) = 1/2$.

The last term in $\tilde{\rho}_2$ is

$$\chi_{B_{\varepsilon^{r-1}}} \Gamma \left[\widehat{D}_0 * \overline{\widehat{D}_0} * \widehat{D}_0 - \widehat{D} * \overline{\widehat{D}} * \widehat{D} \right]$$

$$= \chi_{B_{\varepsilon^{r-1}}} \Gamma \left[(\widehat{D} - \varepsilon v \widehat{C}^{(\varepsilon)}) * (\overline{\widehat{D} - \varepsilon v \widehat{C}}^{(\varepsilon)}) * (\widehat{D} - \varepsilon v \widehat{C}^{(\varepsilon)}) - \widehat{D} * \overline{\widehat{D}} * \widehat{D} \right].$$

In the L^1-norm this can be estimated using Young's inequality by

$$c \left(\varepsilon \|\widehat{C}\|_{L^1_1(\mathbb{R}^d)} \|\widehat{D}\|^2_{L^1(\mathbb{R}^d)} + \varepsilon^2 \|\widehat{C}\|^2_{L^1_1(\mathbb{R}^d)} \|\widehat{D}\|_{L^1(\mathbb{R}^d)} + \varepsilon^3 \|\widehat{C}\|^3_{L^1_1(\mathbb{R}^d)} \right).$$

In summary, with the above choice of β and r we get

$$\|\tilde{\rho}_2\|_{L^1(\mathbb{R}^d)} \leq c \left(\varepsilon^{1/2} \|\widehat{d}\|_{L^1_2} + \varepsilon^2 \|\widehat{C}\|_{L^1_4(\mathbb{R}^d)} \right.$$

$$\left. + \varepsilon \|\widehat{C}\|_{L^1_1(\mathbb{R}^d)} \|\widehat{D}\|^2_{L^1(\mathbb{R}^d)} + \varepsilon^2 \|\widehat{C}\|^2_{L^1_1(\mathbb{R}^d)} \|\widehat{D}\|_{L^1(\mathbb{R}^d)} + \varepsilon^3 \|\widehat{C}\|^3_{L^1_1(\mathbb{R}^d)} \right). \tag{3.20}$$

The terms ρ_1 and ρ_3 are estimated in (3.14) and (3.16). As discussed above, we seek $\widehat{D}$ in $L^1_2(\mathbb{R}^d)$ and hence we set $s_D = 2$ in (3.11). This yields $\alpha = 2$ and

$$\|\rho_3(\widehat{D})\|_{L^1(\mathbb{R}^d)} \leq c(\|\widehat{D}\|_{L^1_2})\varepsilon^2.$$

In summary, for any $\widehat{C} \in L^1_4(\mathbb{R}^d)$ fixed (with C being a solution of the NLS), the right hand side of (3.19) is estimated as

$$\|\tilde{\rho}(\widehat{D})\|_{L^1(\mathbb{R}^d)} \leq c \left(\varepsilon + \|\widehat{d}\|_{L^1_2} \varepsilon^{1/2} + (\|\widehat{d}\|^2_{L^1_2} + \|\widehat{d}\|^3_{L^1_2})\varepsilon \right). \tag{3.21}$$

We proceed with a fixed point argument for the correction d. In order to obtain a differentiable function (to use the Jacobian of the NLS), we write f_{NLS} in real variables. Writing $D = D_R + iD_I$, $C = C_R + iC_I$, $d = d_R + id_I$, and $\tilde{\rho} = \tilde{\rho}_R + i\tilde{\rho}_I$, we define

$$F_{\text{NLS}}(D_R, D_I) := \begin{pmatrix} \text{Re}(f_{\text{NLS}}(D_R + iD_I)) \\ \text{Im}(f_{\text{NLS}}(D_R + iD_I)) \end{pmatrix},$$

the Jacobian

$$J := DF_{\text{NLS}}(C_R, C_I) \tag{3.22}$$

as well as the Fourier-truncation of the Jacobian

$$\widehat{J}_\varepsilon := \chi_{B_{\varepsilon^{r-1}}} \mathcal{F} \left(DF_{\text{NLS}}(C_R^{(\varepsilon)}, C_I^{(\varepsilon)}) \right),$$

where $\widehat{C_R}^{(\varepsilon)} := \chi_{B_{\varepsilon^{r-1}}} \widehat{C_R}$, $C_R^{(\varepsilon)} := \mathcal{F}^{-1}(\widehat{C_R}^{(\varepsilon)})$ and $\widehat{C_I}^{(\varepsilon)} := \chi_{B_{\varepsilon^{r-1}}} \widehat{C_I}$, $C_I^{(\varepsilon)} := \mathcal{F}^{-1}(\widehat{C_I}^{(\varepsilon)})$. $\widehat{J}_\varepsilon$ has the form

$$\widehat{J}_\varepsilon = \chi_{B_{\varepsilon^{r-1}}} (\omega_1 + \kappa^T G_0 \kappa) \text{Id}_{2\times 2}$$

$$+ \chi_{B_{\varepsilon^{r-1}}} \Gamma \begin{pmatrix} 3\widehat{C_R}^{(\varepsilon)} * \widehat{C_R}^{(\varepsilon)} * + \widehat{C_I}^{(\varepsilon)} * \widehat{C_I}^{(\varepsilon)} * & 2\widehat{C_R}^{(\varepsilon)} * \widehat{C_I}^{(\varepsilon)} * \\ 2\widehat{C_R}^{(\varepsilon)} * \widehat{C_I}^{(\varepsilon)} * & 3\widehat{C_I}^{(\varepsilon)} * \widehat{C_I}^{(\varepsilon)} * + \widehat{C_R}^{(\varepsilon)} * \widehat{C_R}^{(\varepsilon)} * \end{pmatrix}.$$

With this notation (3.19) reads

$$\widehat{J}_\varepsilon \vec{\widehat{d}} = W(\vec{\widehat{d}}),$$

where

$$\vec{\widehat{d}} := (\widehat{d_R}, \widehat{d_I})^T \text{ and } W(\vec{\widehat{d}}) := \chi_{B_{\varepsilon^{r-1}}} \vec{\tilde{\rho}}((1 + \varepsilon v)\vec{\widehat{C}}^{(\varepsilon)} + \vec{\widehat{d}}) - \chi_{B_{\varepsilon^{r-1}}} \left(F_{\text{NLS}}(\vec{\widehat{C}}^{(\varepsilon)} + \vec{\widehat{d}}) - \widehat{J}_\varepsilon \vec{\widehat{d}} \right)$$

with $\vec{\widehat{C}}^{(\varepsilon)} := (\widehat{C_R}^{(\varepsilon)}, \widehat{C_I}^{(\varepsilon)})^T$ and $\vec{\tilde{\rho}} := (\tilde{\rho}_R, \tilde{\rho}_I)^T$.

The aim is to construct a small fixed point $\widehat{\widetilde{d}} \in L^1_{s_D}(\mathbb{R}^d)$ of $\widehat{J}_\varepsilon^{-1} W$. The difficulty is that the inverse of $\widehat{J}_\varepsilon$ is not bounded uniformly in ε. This is due to the presence of the $d+1$ zero eigenvalues of J caused by the d spatial shift invariances and the phase invariance of the NLS, see assumption (A.3). The distance of the essential spectrum of $\widehat{J}_\varepsilon$ from zero is $|\omega_1|$ due to the choice of $\mathrm{sign}(\omega_1)$. To eliminate the zero eigenvalues, we work in a symmetric subspace of $L^1_{s_D}(\mathbb{R}^d)$ in which the invariances do not hold. A natural symmetry is the PT-symmetry. Hence, we consider the fixed point problem

$$\widehat{\widetilde{d}} = \widehat{J}_\varepsilon^{-1} W(\widehat{\widetilde{d}})$$

in the space

$$X^{\mathrm{sym}}_{s_D} := \{\widehat{\widetilde{d}} \in L^1_{s_D}(\mathbb{R}^d) : \mathrm{supp}(\widehat{\widetilde{d}}) \subset B_{\varepsilon^{r-1}},\, d(-x) = \overline{d(x)}\ \forall x \in \mathbb{R}^d\}$$

$$= \{\widehat{\widetilde{d}} \in L^1_{s_D}(\mathbb{R}^d) : \mathrm{supp}(\widehat{\widetilde{d}}) \subset B_{\varepsilon^{r-1}},\, \mathrm{Im}(\widehat{d_R}) = -\mathrm{Re}(\widehat{d_I})\}.$$

Note that $\widehat{J}_0, \widehat{J}_\varepsilon : L^1_q(\mathbb{R}^d) \to L^1_{q-2}(\mathbb{R}^d)$ for any $q \geq 2$. Because of assumption (A.3) $\widehat{J}_0^{-1}$ is bounded in $X^{\mathrm{sym}}_{s_D}$ for any $s_D \geq 2$. Since $\widehat{J}_\varepsilon$ is a perturbation of $\widehat{J}_0$, we still need to ensure that 0 is not an eigenvalue of $\widehat{J}_\varepsilon$. For that we use a spectral stability result of Kato, see [12, Theorem IV.3.17]. Applied to our problem in the Banach space $X^{\mathrm{sym}}_{s_D-2}$, $s_D \geq 2$ with the domain of $\widehat{J}_0$ being $D(\widehat{J}_0) = X^{\mathrm{sym}}_{s_D}$, it reads:

Assume $\widehat{J}_\varepsilon - \widehat{J}_0$ is $\widehat{J}_0$-bounded, i.e. $\mathrm{domain}(\widehat{J}_0) \subset \mathrm{domain}(\widehat{J}_\varepsilon - \widehat{J}_0)$ and for some $a, b \geq 0$ is

$$\|(\widehat{J}_\varepsilon - \widehat{J}_0)\widehat{\widetilde{d}}\|_{L^1_{s_D-2}} \leq a\|\widehat{\widetilde{d}}\|_{L^1_{s_D-2}} + b\|\widehat{J}_0\widehat{\widetilde{d}}\|_{L^1_{s_D-2}} \quad \text{for all } \widehat{\widetilde{d}} \in X^{\mathrm{sym}}_{s_D}. \tag{3.23}$$

If for some $\zeta \in \rho(\widehat{J}_0)$

$$a\|(\widehat{J}_0 - \zeta)^{-1}\|_{L^1_{s_D-2} \to L^1_{s_D}} + b\|\widehat{J}_0(\widehat{J}_0 - \zeta)^{-1}\|_{L^1_{s_D-2} \to L^1_{s_D-2}} < 1, \tag{3.24}$$

then $\zeta \in \rho(\widehat{J}_\varepsilon)$ and

$$\|(\widehat{J}_\varepsilon - \zeta)^{-1}\|_{L^1_{s_D-2} \to L^1_{s_D}} \leq \|(\widehat{J}_0 - \zeta)^{-1}\|_{L^1_{s_D-2} \to L^1_{s_D}} \left(1 - a\|(\widehat{J}_0 - \zeta)^{-1}\|_{L^1_{s_D-2} \to L^1_{s_D}}\right.$$
$$\left. - b\|\widehat{J}_0(\widehat{J}_0 - \zeta)^{-1}\|_{L^1_{s_D-2} \to L^1_{s_D-2}}\right)^{-1}.$$

We check now (3.23) and (3.24) for $\zeta = 0$. For $\widehat{\widetilde{d}} \in X^{\mathrm{sym}}_{s_D}$ one has

$$\|(\widehat{J}_\varepsilon - \widehat{J}_0)\widehat{\widetilde{d}}\|_{L^1_{s_D-2}} = |\Gamma| \left\| \left(\widehat{C_R}^{(\varepsilon)} * \widehat{C_R}^{(\varepsilon)} - \widehat{C_R} * \widehat{C_R}\right) * (3\widehat{d_R} + \widehat{d_I})\right.$$
$$+ \left(\widehat{C_I}^{(\varepsilon)} * \widehat{C_I}^{(\varepsilon)} - \widehat{C_I} * \widehat{C_I}\right) * (3\widehat{d_I} + \widehat{d_R})$$
$$\left. + 2\left(\widehat{C_R}^{(\varepsilon)} * \widehat{C_I}^{(\varepsilon)} - \widehat{C_R} * \widehat{C_I}\right) * (\widehat{d_R} + \widehat{d_I})\right\|_{L^1_{s_D-2}}.$$

Writing $\widehat{C_{R,I}}^{(\varepsilon)} =: \widehat{C_{R,I}} + \widehat{\gamma_{R,I}}^{(\varepsilon)}$, it is $\mathrm{supp}(\widehat{\gamma_{R,I}}^{(\varepsilon)}) \subset B^c_{\varepsilon^{r-1}}$. The difference $(\widehat{J}_\varepsilon - \widehat{J}_0)\widehat{\widetilde{d}}$ consists of terms that are linear or quadratic in $\widehat{\gamma_{R,I}}^{(\varepsilon)}$; for instance terms like $\widehat{\gamma_R}^{(\varepsilon)} * \widehat{C_R}^{(\varepsilon)} * \widehat{d_R}$ or $\widehat{\gamma_I}^{(\varepsilon)} * \widehat{C_R}^{(\varepsilon)} * \widehat{d_I}$. Because

$$\|\widehat{\gamma_{R,I}}^{(\varepsilon)}\|_{L^1_{s_D-2}} = \int_{B^c_{\varepsilon^{r-1}}} (1 + |\kappa|)^{s_D-2} |\widehat{\gamma_{R,I}}(\kappa)|\, \mathrm{d}\kappa$$

$$\leq \sup_{|\kappa| > \varepsilon^{r-1}} \frac{1}{(1 + |\kappa|)^2} \|\widehat{\gamma_{R,I}}^{(\varepsilon)}\|_{L^1_{s_D}} \leq c\varepsilon^{2(1-r)} \|\widehat{\gamma_{R,I}}^{(\varepsilon)}\|_{L^1_{s_D}},$$

Young's inequality for convolutions yields

$$\|(\widehat{J}_\varepsilon - \widehat{J}_0)\widehat{\widetilde{d}}\|_{L^1_{s_D-2}} \leq c(\|\widehat{C}\|_{L^1_{s_D}})\varepsilon^{2(1-r)} \|\widehat{\widetilde{d}}\|_{L^1_{s_D-2}}$$

with c depending polynomially on $\|\widehat{C}\|_{L^1_{s_D}}$. Conditions (3.23) and (3.24) for $\zeta = 0$ are thus satisfied with $a = c\varepsilon^{2(1-r)}$ and $b = 0$ if $\varepsilon > 0$ is small enough and if $\widehat{C} \in L^1_{s_D}(\mathbb{R}^d)$.

For the fixed point problem we use $s_D = 2$ and $r = 1/2$ and show firstly that if $\widehat{C} \in X^{\text{sym}}_4$, then there is some $c > 0$ such that for all $\varepsilon > 0$ small enough

$$\widehat{J}^{-1}_\varepsilon W : B^{2,\,\text{sym}}_{c\varepsilon} \to B^{2,\,\text{sym}}_{c\varepsilon}, \tag{3.25}$$

where $B^{2,\,\text{sym}}_{c\varepsilon}$ is the $c\varepsilon$-ball in X^{sym}_2, i.e.

$$B^{2,\,\text{sym}}_{c\varepsilon} := \{\varphi \in X^{\text{sym}}_2 : \|\varphi\|_{L^1_2} \le c\varepsilon\}.$$

Note that the requirement $\widehat{C} \in L^1_4$ is dictated by (3.20).

Secondly, we prove that $\widehat{J}^{-1}_\varepsilon W$ is contractive provided $\widehat{C} \in \{\widehat{f} \in L^1_4(\mathbb{R}^d) : \operatorname{Im}(\widehat{f_R}) = -\operatorname{Re}(\widehat{f_I})\}$. We start by showing $\widehat{J}_\varepsilon : X^{\text{sym}}_2 \to X^{\text{sym}}_0$. The loss of 2 in the weight is due to the second order nature of the operator J, i.e. due to the factor $\kappa^T G_0 \kappa$ in Fourier variables. The entries $\omega_1 + \kappa^T G_0 \kappa$ clearly preserve the PT-symmetry. For the convolution terms we have, for instance

$$\operatorname{Im}(\widehat{C_R}) = -\operatorname{Re}(\widehat{C_I}) \;\Rightarrow\; \operatorname{Im}(\widehat{C_R}^{(\varepsilon)}) = -\operatorname{Re}(\widehat{C_I}^{(\varepsilon)}) \;\Rightarrow\; C^{(\varepsilon)}(-x) = \overline{C^{(\varepsilon)}(x)} \;\forall x.$$

Hence $C^{(\varepsilon)^2}_R d_R$, $C^{(\varepsilon)^2}_I d_R$, and $C^{(\varepsilon)}_R C^{(\varepsilon)}_I d_I$ are even and $C^{(\varepsilon)^2}_I d_I$, $C^{(\varepsilon)^2}_R d_I$, and $C^{(\varepsilon)}_R C^{(\varepsilon)}_I d_R$ are odd such that the $PT-$symmetry is preserved also by the convolution terms. In end effect, $\operatorname{Im}((\widehat{J}_\varepsilon \widehat{d})_1) = -\operatorname{Re}((\widehat{J}_\varepsilon \widehat{d})_2)$. Hence, $\widehat{J}_\varepsilon : X^{\text{sym}}_2 \to X^{\text{sym}}_0$ and for $\widehat{J}^{-1}_\varepsilon$ we get $\widehat{J}^{-1}_\varepsilon : X^{\text{sym}}_0 \to X^{\text{sym}}_2$.

Next, we show that $W : B^{2,\,\text{sym}}_{c\varepsilon} \to B^{0,\text{sym}}_{c\varepsilon}$ if $\widehat{C} \in X^{\text{sym}}_4$. The term ρ is estimated in (3.21) and dictates the order ε^1.

The difference $F_{\text{NLS}}(\widehat{C}^{(\varepsilon)} + \widehat{d}) - \widehat{J}_\varepsilon \widehat{d}$ consists of terms quadratic in $\widehat{d}$ and hence is bounded in $L^1(\mathbb{R}^d)$ by $c_1(\|\widehat{d}\|^2_{L^1(\mathbb{R}^d)} + \|\widehat{d}\|^3_{L^1(\mathbb{R}^d)})$. In summary,

$$\|W(\widehat{d})\|_{L^1(\mathbb{R}^d)} \le c_2(\varepsilon + \varepsilon^{1/2}\|\widehat{d}\|_{L^1_2(\mathbb{R}^d)} + \|\widehat{d}\|^2_{L^1_2(\mathbb{R}^d)} + \|\widehat{d}\|^3_{L^1_2(\mathbb{R}^d)}) \le c\varepsilon$$

if $\|\widehat{d}\|_{L^1_2(\mathbb{R}^d)} \le c\varepsilon$ and $\varepsilon > 0$ is small enough. Due to the boundedness of $\widehat{J}^{-1}_\varepsilon : X^{\text{sym}}_0 \to X^{\text{sym}}_2$ we thus have (3.25).

The contractive property of $\widehat{J}^{-1}_\varepsilon W$ in $B^{2,\,\text{sym}}_{c\varepsilon}$ is now clear due to the quadratic nature of $F_{\text{NLS}}(\widehat{C}^{(\varepsilon)} + \widehat{d}) - \widehat{J}_\varepsilon \widehat{d}$.

We conclude that if the solution C of the NLS (2.3) satisfies $\widehat{C} \in \{\widehat{f} \in L^1_4(\mathbb{R}^d) : \operatorname{Im}(\widehat{C_R}) = -\operatorname{Re}(\widehat{C_I})\}$, then there is $c > 0$ such that for all $\varepsilon > 0$ small enough the constructed solution D of (3.13) satisfies

$$\widehat{D} \in \{\widehat{f} \in L^1_2(\mathbb{R}^d) : \operatorname{Im}(\widehat{f_R}) = -\operatorname{Re}(\widehat{f_I})\}$$

and due to (3.18)

$$\|\widehat{D} - \chi_{B_{\varepsilon^{-1/2}}}\widehat{C}\|_{L^1_2(\mathbb{R}^d)} \le c(\|\widehat{C}\|_{L^1_4(\mathbb{R}^d)})\varepsilon. \tag{3.26}$$

Here we have also used $\|\varepsilon\nu\chi_{B_{\varepsilon^{-1/2}}}\widehat{C}\|_{L^1_2} \le c\varepsilon\|\widehat{C}\|_{L^1_3}$.

This allows us to estimate $\widehat{\widehat{B}}_D - \widehat{\widehat{B}}_{\text{app}}$. We have

$$\|\widehat{\widehat{B}}_D - \widehat{\widehat{B}}_{\text{app}}\|_{L^1(\mathbb{R}^d)} \le \varepsilon^{1-d}\left\{\left\|\left(\widehat{D}\left(\frac{\cdot - k_0}{\varepsilon}\right) - \widehat{C}\left(\frac{\cdot - k_0}{\varepsilon}\right)\right)\vec{\eta}^{(j_0)}(k_0)\right\|_{L^1(B_{\varepsilon^r}(k_0))}\right.$$

$$+ \left\|\widehat{D}\left(\frac{\cdot - k_0}{\varepsilon}\right)\left(\vec{\eta}^{(j_0)}(\cdot) - \vec{\eta}^{(j_0)}(k_0)\right)\right\|_{L^1(B_{\varepsilon^r}(k_0))}$$

$$\left. + \left\|\widehat{C}\left(\frac{\cdot - k_0}{\varepsilon}\right)\vec{\eta}^{(j_0)}(k_0)\right\|_{L^1(B^c_{\varepsilon^r}(k_0))}\right\}$$

Next we use (3.26), the Lipschitz continuity of $\vec{\eta}^{(j_0)}$, and the estimate $\|\widehat{C}\|_{L^1(B^c_{\varepsilon^r-1})} \leq \varepsilon^{(1-r)s_C} \|\widehat{C}\|_{L^1_{s_C}(\mathbb{R}^d)}$ for all $s_C \geq 0$. This produces at $r = 1/2$

$$\|\widehat{\overline{B}}_D - \widehat{\overline{B}}_{\text{app}}\|_{L^1(\mathbb{R}^d)} \leq c\left(\varepsilon^2 + \varepsilon^2\|\widehat{D}\|_{L^1_1(\mathbb{R}^d)} + \varepsilon^{1+s_C/2}\|\widehat{C}\|_{L^1_{s_C}(\mathbb{R}^d)}\right)$$

$$\leq c\left(1 + \|\widehat{D}\|_{L^1_1(\mathbb{R}^d)} + \|\widehat{C}\|_{L^1_4(\mathbb{R}^d)}\right)\varepsilon^2$$

if $s_C = 4$.

We can now summarize the error estimate

$$\|\widehat{\overline{B}} - \widehat{\overline{B}}_{\text{app}}\|_{L^1(\mathbb{R}^d)} \leq \|\widehat{\overline{B}}_D - \widehat{\overline{B}}_{\text{app}}\|_{L^1(\mathbb{R}^d)} + \|\widehat{\overline{B}} - \widehat{\overline{B}}_D\|_{L^1(\mathbb{R}^d)}$$

$$\leq \|\widehat{\overline{B}}_D - \widehat{\overline{B}}_{\text{app}}\|_{L^1(\mathbb{R}^d)} + \|\widehat{\overline{B}}_Q\|_{L^1(\mathbb{R}^d)} + \|\widehat{\overline{B}}_R\|_{L^1(\mathbb{R}^d)}.$$

The components $\widehat{\overline{B}}_Q$ and $\widehat{\overline{B}}_R$ are estimated in (3.8), (3.9), and (3.12). Having now estimated $\|\widehat{R}\|_{L^1}$ in terms of $\|\widehat{C}\|_{L^1}$ and $\|\widehat{D}\|_{L^1}$ in terms of $\|\widehat{C}\|_{L^1_2}$, we get for $r = 1/2$ and $s_D = 2$

$$\|\widehat{\overline{B}}_Q\|_{L^1(\mathbb{R}^d)} \leq c_1(\|\widehat{C}\|_{L^1_2(\mathbb{R}^d)})\varepsilon^3, \qquad \|\widehat{\overline{B}}_R\|_{L^1(\mathbb{R}^d)} \leq c_2(\|\widehat{C}\|_{L^1_2(\mathbb{R}^d)})\varepsilon^3,$$

where c_1 and c_2 depend polynomially on $\|\widehat{C}\|_{L^1_2(\mathbb{R}^d)}$. Hence, the estimate in Theorem 1 is proved.

4 Numerical Example of Bifurcating Gap Solitons for $d = 2$

In [9] it is shown that assumption (A.1), i.e. the existence of a spectral gap is satisfied for $N = 4$ in the symmetric case

$$\begin{aligned}
&v_g^{(1)} = -v_g^{(2)} =: v, v_g^{(3)} = -v_g^{(4)} =: w, \\
&\kappa_{12} = \kappa_{34} =: \alpha_1, \\
&\kappa_{14} = \kappa_{32} =: \alpha_2, \\
&\kappa_{13} = \kappa_{42} =: \alpha_3, \\
&\kappa_{jj} = 0, j = 1, \ldots, 4
\end{aligned} \tag{4.1}$$

provided $|\alpha_1|^2 > 2(|\alpha_2|^2 + |\alpha_3|^2)$. In the following example we choose $v = (0, 1)^T$, $w = (1, 0)^T$, $\alpha_1 = 2$, and $\alpha_2 = \alpha_3 = 1$. The dispersion relation $\omega_j : \mathbb{R}^2 \to \mathbb{R}$, $j = 1, \ldots, 4$ of (1.1) is plotted in Fig. 1. The gap appears even though the sufficient condition $|\alpha_1|^2 > 2(|\alpha_2|^2 + |\alpha_3|^2)$ is not satisfied.

We see that the second eigenvalue λ_2 has an isolated maximum at $k = k_0 := 0$. The corresponding frequency is $\omega_0 := \lambda_2(0) = 0$. The eigenvector corresponding to $\lambda_2(0)$ is $\vec{\eta}_{j_0}(0) = \frac{1}{\sqrt{2}}(1, 1, -1, -1)^T$.

We use the following special case of the coefficients $\gamma_j^{(m,n,o)}$ in (1.3)

$$\begin{aligned}
1 = \gamma_j^{(j,j,j)} &= \gamma_j^{(j,i,i)} = \gamma_j^{(i,i,j)}, \quad i, j = 1, \ldots, 4, \\
&= \gamma_1^{(3,2,4)} = \gamma_1^{(4,2,3)} = \gamma_2^{(3,1,4)} = \gamma_2^{(4,1,3)} \\
&= \gamma_3^{(1,4,2)} = \gamma_3^{(2,4,1)} = \gamma_4^{(1,3,2)} = \gamma_4^{(2,3,1)}, \\
\gamma_j^{(m,n,o)} &= 0 \text{ otherwise.}
\end{aligned} \tag{4.2}$$

Clearly, coefficients (4.1) and (4.2) allow symmetric solutions with $B_2 = \overline{B_1}$ and $B_4 = \overline{B_3}$. We do not make a direct use of this symmetry in our computations. We construct the approximation $\vec{B}_{\text{app}}$ of a solution of (1.3) at $\omega = \omega_0 + \varepsilon^2\omega_1$ for six values of ε: 0.2, 0.1, 0.05, 0.025, 0.0125, and 0.00625. The coefficients of the effective NLS (2.3) are $G_0 = -0.25 \, I_{2x2}$ and $\Gamma = 2.25$ and we choose $\omega_1 = 1$. A
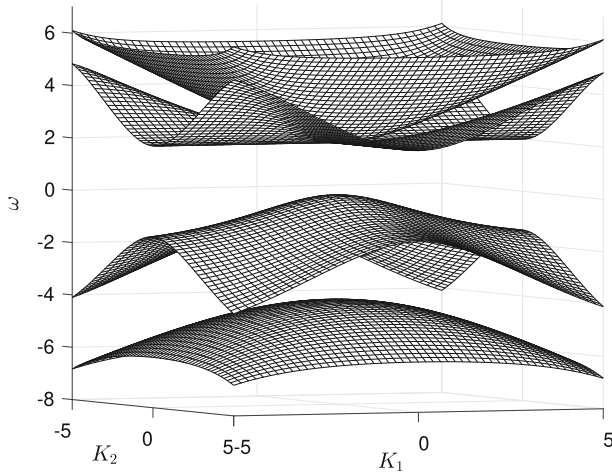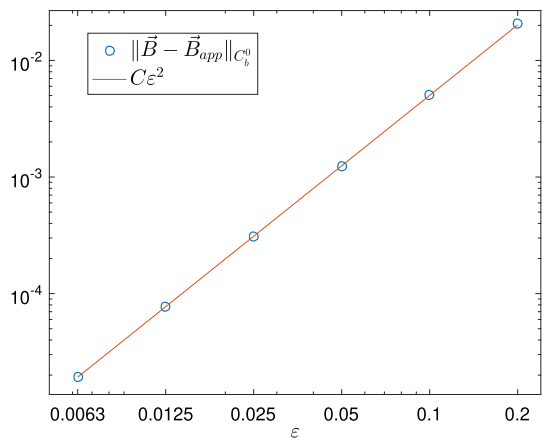
Fig. 1 Dispersion relation of (1.1) for the example in Sect. 4
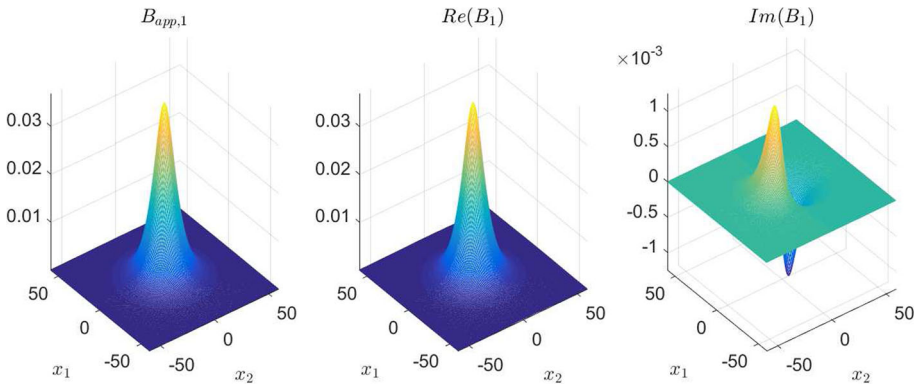
Fig. 2 Asymptotic approximation $B_{\mathrm{app},1}$ and the numerical solution B_1 (real and imaginary part) at $\varepsilon = 0.05$

Fig. 3 Convergence of the asymptotic error in ε

real C radially symmetric was chosen in this example. It was computed using the shooting method for the NLS in polar variables.

Using the numerical Petviashvili iteration [1,14], we also produce a numerical approximation of a solution $\vec{B}$ at $\omega = \omega_0 + \varepsilon^2$. The Petviashvili iteration is a fixed point iteration in Fourier variables with a stabilizing normalization factor. The initial guess of the iteration was chosen as $\vec{B}_{\mathrm{app}}$. Note that although $\vec{B}_{\mathrm{app}}$ can be real (if a real solution C of the NLS is chosen), Eq. (1.3) does not allow real solutions $\vec{B}$ due to the term $i\nabla\vec{B}$ and due to the realness of α_1, α_2, and α_3 and of $\gamma_j^{(m,n,o)}$. Nevertheless, if $\vec{B}_{\mathrm{app}}$ is real, there must be a solution $\vec{B}$ with $\mathrm{Im}(\vec{B}) = O(\varepsilon^2)$. Figure 2 shows $\vec{B}_{\mathrm{app}}$ and $\vec{B}$ for $\varepsilon = 0.05$.

The numerical parameters for the Petviashvili iteration were selected as follows: we compute on the domain $x \in [-3/\varepsilon, 3/\varepsilon]^2$ with the discretization given by 160x160 grid points, i.e. $dx_1 = dx_2 = 3/(80\varepsilon)$. Note that because $k_0 = 0$, the relatively coarse discretization for small values of ε does not matter (there are no oscillations to be resolved).

For each ε we evaluate the asymptotic error $E := \|\vec{B} - \vec{B}_{\mathrm{app}}\|_{C_0}$. Figure 3 shows the convergence of the error in ε. Clearly, $E(\varepsilon) \sim c\varepsilon^2$, which confirms the convergence rate proved in our theorem.

References

1. Ablowitz, M.J., Musslimani, Z.H.: Spectral renormalization method for computing self-localized solutions to nonlinear systems. Opt. Lett. **30**(16), 2140–2142 (2005)
2. Aceves, A.B., Wabnitz, S.: Self induced transparency solitons in nonlinear refractive media. Phys. Lett. A **141**, 37–42 (1989)
3. Agueev, D., Pelinovsky, D.: Modeling of wave resonances in low-contrast photonic crystals. SIAM J. Appl. Math. **65**(4), 1101–1129 (2005)
4. Dohnal, T.: Traveling solitary waves in the periodic nonlinear Schrödinger equation with finite band potentials. SIAM J. Appl. Math. **74**, 306–321 (2014)
5. Dohnal, T., Helfmeier, L.: Justification of the coupled mode asymptotics for localized wavepackets in the periodic nonlinear Schrödinger equation. J. Math. Anal. Appl. **450**, 691–726 (2017)
6. Dohnal, T., Pelinovsky, D.: Bifurcation of nonlinear bound states in the periodic Gross-Pitaevskii equation with PT-symmetry. Proc. R. Soc. Edinb. Sect. A Math. **150**(1), 171–204 (2020)
7. Dohnal, T., Pelinovsky, D., Schneider, G.: Coupled-mode equations and gap solitons in a two-dimensional nonlinear elliptic problem with a separable periodic potential. J. Nonlinear Sci. **19**(2), 95–131 (2009)
8. Dohnal, T., Uecker, H.: Coupled mode equations and gap solitons for the 2D Gross-Pitaevskii equation with a non-separable periodic potential. Phys. D **238**(9–10), 860–879 (2009)
9. Dohnal, T., Wahlers, L.: Coupled mode equations and gap solitons in higher dimensions. J. Differ. Equ. **269**(3), 2386–2418 (2020)
10. Giannoulis, J., Mielke, A., Sparber, C.: Interaction of modulated pulses in the nonlinear Schrödinger equation with periodic potential. J. Differ. Equ. **245**(4), 939–963 (2008)
11. Goodman, R.H., Weinstein, M.I., Holmes, P.J.: Nonlinear propagation of light in one-dimensional periodic structures. J. Nonlinear Sci. **11**(2), 123–168 (2001)

12. Katō, T.: Perturbation Theory for Linear Operators. Grundlehren der mathematischen Wissenschaften. Springer, Berlin (1995)
13. Pelinovsky, D., Schneider, G.: Justification of the coupled-mode approximation for a nonlinear elliptic problem with a periodic potential. Appl. Anal. **86**(8), 1017–1036 (2007)
14. Petviashvili, V.I.: Equation of an extraordinary soliton. Plasma Phys. **2**, 469 (1976)
15. Schneider, G., Uecker, H.: Nonlinear coupled mode dynamics in hyperbolic and parabolic periodically structured spatially extended systems. Asymptot. Anal. **28**(2), 163–180 (2001)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.