Waves in Random and Complex Media
Vol. 15, No. 2, May 2005, 145–228

Nonlinear photonic crystals: IV. Nonlinear Schrödinger equation regime

A. BABIN and A. FIGOTIN*

Department of Mathematics, University of California at Irvine, CA 92697, USA

We study here the nonlinear Schrödinger (NLS) equation as the first term in a sequence of approximations for an electromagnetic (EM) wave propagating according to the nonlinear Maxwell (NLMs) equations. The dielectric medium is assumed to be periodic, with a cubic nonlinearity, and with its linear background possessing inversion symmetric dispersion relations. The medium is excited by a current **J** producing an EM wave. The wave nonlinear evolution is analysed based on the modal decomposition and an expansion of the exact solution to the NLM into an asymptotic series with respect to three small parameters α, β and ϱ. These parameters are introduced through the excitation current **J** to scale, respectively (i) its amplitude and consequently the magnitude of the nonlinearity; (ii) the range of wavevectors involved in its modal composition, with β^{-1} scaling its spatial extension; (iii) its frequency bandwidth, with ϱ^{-1} scaling its time extension. We develop a consistent theory of approximations of increasing accuracy for the NLM with its first term governed by the NLS. We show that such NLS regime is the medium response to an almost monochromatic excitation current **J**. The developed approach not only provides rigorous estimates of the approximation accuracy of the NLM with the NLS in terms of powers of α, β and ϱ, but it also produces a new extended NLS (ENLS) providing better approximations. Remarkably, quantitative estimates show that properly tailored ENLS can significantly improve the approximation accuracy of the NLM compared with the classical NLS equation.

Notations and abbreviations

almost single-mode excitation—*see* (55)
bidirectional quadruplet—*see* (341)
directly excited modes—*see* (219), (58)
doublet—*see* (339)
ENLS extended nonlinear Schrödinger equation—*see* (100), (101) or (114) and section 1.4
Floquet–Bloch modal decomposition—*see* (146)
FM—frequency matching condition (223), frequency matched, *see* (227)
FNLR first nonlinear response—*see* (14), (176), (167), (174)
Fourier transform—*see* (425)
GVM group velocity matching condition—*see* (73), (222)
indirectly excited modes—*see* (220), (68), (69)
interaction quadruplet—*see* (72)
linear response—*see* (13)
NLM—nonlinear Maxwell equation, *see* (3)
NLS—nonlinear Schrödinger equation, *see* (36), (37) and (309)

*Corresponding author. E-mail: afigotin@uci.edu

Waves in Random and Complex Media
ISSN: 1745-5030 (print), 1745-5049 (online) © 2005 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/17455030500196929

NLS regime—a situation when the evolution of an electromagnetic (EM) wave is governed by the NLM equations and it can be approximated by an NLS or, may be, by a slightly more general extended NLS

non-FM—non-frequency-matched

rectifying coordinates—*see* (281), (282)

susceptibility $\chi_D^{(3)}$—*see* (156)

susceptibility $\chi^{(3)}$—*see* (157)

unidirectional **excitation**—*see* (188)

$\alpha_\pi = 3\alpha(2\pi)^{2d}$

$\gamma_{(\nu)}(\boldsymbol{\eta})$—the Taylor polynomial of $\omega_{\bar{n}}\ (\mathbf{k}_* + \boldsymbol{\eta})$ of order ν (200)

$\gamma_{(2)}(\boldsymbol{\eta})$—the second-order Taylor polynomial of $\omega_{\bar{n}}\ (\mathbf{k}_* + \boldsymbol{\eta})$ (201)

$\delta^{\pm}_{\times,\zeta}$—coefficients defined by (349), (345), (346)

$\zeta = \pm 1$ or $\zeta = \pm$—band binary number, when used in indices is abbreviated to $\zeta = \pm$, namely $V_\zeta = V_+$ if $\zeta = +1$, $V_\zeta = V_-$ if $\zeta = -1$.

$\vec{\zeta}_{0,\times}$—vectors defined by (345)

$\theta = \frac{\varrho}{\beta^2}$—inverse dispersion parameter, *see* (24)

$\xi = (\xi_1, \ldots, \xi_d)$—Fourier wavevector variable

$\omega_{\bar{n}}(\mathbf{k}) = \zeta\omega_n(\mathbf{k})$—dispersion relation of the band (ζ, n), *see* (138)

$\omega'_{n_0}(\mathbf{k}) = \nabla_{\mathbf{k}}\omega_{n_0}(\mathbf{k})$—group velocity vector

$\omega''_{n_0}(\mathbf{k}_*) = \nabla_{\mathbf{k}}\nabla_{\mathbf{k}}\omega_{n_0}(\mathbf{k}_*)$—Hessian matrix of $\omega_{n_0}(\mathbf{k})$ at $\mathbf{k}_*$

$\tilde{\mathbf{G}}_{\bar{n}}(\mathbf{r}, \mathbf{k})$—the eigenfunction (eigenmode) corresponding to band index $\bar{n}$ and quasimomentum $\mathbf{k}$, *see* (137), (141)

$h_\zeta(\mathbf{r}),\ \zeta = \pm$—initial data for the NLS (36), (37), (301)

$\mathcal{F}^{(1)}_{\mathrm{NL}}$—*see* (162)

$\hat{h}_\zeta(\frac{1}{\beta}\boldsymbol{\xi}),\ \zeta = \pm$—Fourier transform of the initial data $h_\zeta(\beta\mathbf{r})$ for the NLS (189), (425)

$I_{\bar{n},\zeta',\zeta'',\zeta'''}(\mathbf{k}, \tau)$—interaction integral (211)

$\mathbf{J}$—excitation current, *see* (35)

$\mathbf{k} = (k_1, \ldots, k_d)$—quasimomentum (wavevector) variable

$\mathbf{k}_* = (k_{*1}, \ldots, k_{*d})$—centre of the wavepacket, directly excited mode

$\vec{k} = (\mathbf{k}, \mathbf{k}', \mathbf{k}'', \mathbf{k}''')$, $\vec{q} = (\mathbf{q}, \mathbf{q}', \mathbf{q}'', \mathbf{q}''')$—four-wave interaction wavevector, *see* (177), (232)

$\vec{k}_{*,\times,\pm}$—vectors defined by (346)

$\bar{n} = (\zeta, n)$—band index, *see* (138)

n—band number

n_0—band number of directly exited band

$\vec{n} = (\overline{n}, \overline{n}', \overline{n}'', \overline{n}''')$—four-wave interaction band index

$\vec{n}_0 = ((\zeta, n), (\zeta, n_0), (\zeta, n_0), (-\zeta, n_0))$—*see* (228)

$\uparrow n_0, \mathbf{k}_* \downarrow = \{(1, n_0, \mathbf{k}_*), (-1, n_0, -\mathbf{k}_*)\}$—modal doublet, doublet, *see* (339)

$\nabla_{\mathbf{r}} = (\partial/\partial r_1, \partial/\partial r_2, \cdots, \partial/\partial r_d)$

$O(\mu)$—any quantity having the property that $O(\mu)/\mu$ is bounded as $\mu \to 0$

$O(|\mathbf{U}^{(1)}|)$—magnitude of the FNLR, estimated by (290)

$\breve{Q}_{\vec{n}}(\vec{k})$—modal susceptibility defined by (179)

$\breve{Q}_{\vec{n},\bar{l}}(\vec{k})$—a component of the modal susceptibility, *see* (393)

$\vec{q} = (\mathbf{q}, \mathbf{q}', \mathbf{q}'', \mathbf{q}''')$—*see* (177), (232), (230)

$\vec{q}^{\,\flat} = (\mathbf{q}, \mathbf{q}, \mathbf{q}, -\mathbf{q})$—*see* (289)

$\mathbf{q}'''(\vec{q}) = \mathbf{q} - \mathbf{q}' - \mathbf{q}''$—*see* (255)

$\vec{q}^{\,0} = (\mathbf{q}, \mathbf{q}', \mathbf{q}'', \mathbf{q} - \mathbf{q}' - \mathbf{q}'')$—*see* (255)

$\mathbf{r} = (r_1, \ldots, r_d)$—spatial variable

$\vec{q} = (\mathbf{q}, \mathbf{q}', \mathbf{q}'', \mathbf{q}''')$—*see* (232)

$\vec{q}^{\,\star} = (\mathbf{q}', \mathbf{q}'', \mathbf{q}''')$—*see* (261)
$\vec{q}^{\,\sharp} = (\mathbf{q}' + \mathbf{q}'' + \mathbf{q}''', \mathbf{q}', \mathbf{q}'', \mathbf{q}''')$—*see* (261)
$\mathbf{U}$—solution of the NLM, *see* (4)
$\tilde{\mathbf{U}}(\mathbf{k}, \mathbf{r}, t)$—Floquet–Bloch transform of $\mathbf{U}(\mathbf{r}, t)$, *see* (52)
$\tilde{\mathbf{U}}_{\bar{n}}^{(0)}$—modal component of the linear response, *see* (165)
$\tilde{\mathbf{U}}_{\bar{n}}^{(1)}$—modal component of the first nonlinear response, *see* (165)
$\tilde{U}_{\bar{n}}(\mathbf{k}, \tau) = \tilde{u}_{\bar{n}}\,(\mathbf{k}, \tau)\mathrm{e}^{-\mathrm{i}\omega_{\bar{n}}(\mathbf{k})t}$—modal amplitudes, *see* (147), (150)
$u^{(m)}, u^{(m_1,m_2)}$—coefficients of the power series expansions of $u = u(\alpha, \beta, \varrho)$, *see* (160), (174), (175)
$\tilde{u}_{\bar{n}}^{(1)}(\mathbf{k}, \tau)$—modal amplitude depending on the slow time τ of the first nonlinear response in written causal form, *see* (383), (394)
$\tilde{u}_{\bar{n}}^{(1,0)}(\mathbf{k}, \tau)$—modal amplitude depending on the slow time τ of the first nonlinear response in the time-harmonic approximation, *see* (176), (394)
$\hat{V}(\mathbf{q})$—Fourier transform of $V(\mathbf{r})$, *see* (425)
$\hat{V}_{\zeta}(\boldsymbol{\eta}, t) = \hat{v}_{\zeta}(\boldsymbol{\eta}, \tau)\,\mathrm{e}^{-\mathrm{i}\zeta\gamma_{(\nu)}(\zeta\boldsymbol{\eta})t}$—amplitudes, *see* (332)
$\psi_0(\tau)$—slowly time cut off function, *see* (191)
$\psi(\tau) = \int_0^{\tau} \psi_0(\tau)\,\mathrm{d}\tau_1$—auxiliary function with $\psi_0(\tau)$ satisfying (191), *see also* (311)
Ψ—cutoff function in quasimomentum domain, *see* (190)
$\phi_{\bar{n}}(\vec{k}) = \zeta\omega_n(\mathbf{k}) - \zeta'\omega_{n'}(\mathbf{k}') - \zeta''\omega_{n''}(\mathbf{k}'') - \zeta'''\omega_{n'''}(\mathbf{k}''')$—four-wave interaction phase function, *see* (178)
$\Phi^{(\nu)}(\vec{\zeta}_0, \beta\vec{q})$—polynomial phase function (254)
$\zeta = \pm$ binary index
$\vec{\zeta} = (\zeta, \zeta', \zeta'', \zeta''')$—four-wave interaction binary band index
$\vec{\zeta}_0 = (\zeta, \zeta, \zeta, -\zeta)$—*see* (234)
$Z_{\pm}$—solution of the NLS or ENLS
Z^{*}—complex conjugate to Z

1. Introduction

The subject of this work is the accuracy of approximation of solutions to the nonlinear Maxwell (NLM) equations for periodic dielectric media, i.e. photonic crystals (see [35]) by solutions to the nonlinear Schrödinger (NLS) equations or, more broadly, by similar to the NLS equations. The theory of nonlinear photonic crystals is treated in many books and papers, see for references [3-4-14-30-32-59]. Both NLMs and NLSs are widely used in nonlinear optics, and there are many derivations of the NLS in different situations and of different levels of rigour in the physical literature. There is extensive literature devoted to studies of solutions to the NLS (see [1–9] and references therein). If the evolution of an electromagnetic (EM) wave is governed by the NLM and it can be approximated by an NLS or, may be, by a slightly more general extended NLS (ENLS), we refer to it as *NLS regime of propagation or just NLS regime*.

The NLS describes a universal wave propagation regime occurring in a dispersive medium with a dispersive relation $\omega(\mathbf{k})$ for its linear background. A derivation of the NLS equations emphasizing its universal nature can be obtained by introducing an amplitude-dependent dispersion relation (see [7, pp. 4–5; 10, pp. 50–51; 11]) of the form $\omega(\mathbf{k}) + \delta|Z|^2$ formally implying the following NLS evolution:

$$\partial_t Z = -\mathrm{i}[\omega(-\mathrm{i}\vec{\nabla}_{\mathbf{r}})Z + \delta|Z|^2 Z] \tag{1}$$

More elaborate derivations of the NLS based on the NLM evolution make use of Fourier expansions of the involved fields in infinite space [7, pp. 6, 7; 10, pp. 67–71, 83–104]. A similar analysis of the NLM equation for periodic dielectric media, i.e. photonic crystals, was carried out based on Bloch expansions in [12]. In a number of mathematical studies the NLS are derived based on equations other than the nonlinear Maxwell equations (see [13–16]).

Looking at different NLS derivations one can see that they are based on the following fundamental assumptions.

- The nonlinear component of the wave is relatively small (the nonlinearity is weak).
- The wave is defined as a real-valued function.
- The wavevectors (quasimomenta) $\mathbf{k}$ involved in the wave composition are close to a certain $\mathbf{k}_*$.
- The time evolution of the wave envelope is slow compared to the typical carrier wave frequency.
- The dispersion relation $\omega(\mathbf{k}) = \omega(\mathbf{k}_* + \boldsymbol{\eta})$ in the vicinity of $\mathbf{k}_*$ is approximated by its second-order Taylor polynomial $\gamma_{(2)}(\boldsymbol{\eta})$.
- The non-frequency-matched wave interactions (in particular the third harmonic generation) are neglected.
- The frequency dependence of the susceptibility tensor is neglected and its value at $\mathbf{k}_*$ is used.

All the above factors are presented and to some degree are refined in our quantitative approach to the approximation of solutions to the NLM by the NLS. The approach is based on the framework described in [17–19], and its outline is as follows. A wave propagating in the nonlinear media is generated by an excitation current $\mathbf{J}$ which is turned on at time $t = 0$ and is turned off at a later time $t = t_0$. Hence for $t > t_0$ there are no external currents and the wave dynamics is determined entirely by the medium.

Suppose that the excitation current $\mathbf{J}$ has the form of a wavepacket with the carrier frequency $\omega = \omega_{n_0}(\mathbf{k}_*)$ where $\omega_n(\mathbf{k})$ is the dispersion relation of the underlying linear medium with the band number n and the wave numbers (quasimomenta) $\mathbf{k}$, and n_0 and $\mathbf{k}_*$ are chosen. The envelope amplitude of the excitation current $\mathbf{J}$ is supposed to vary slowly in space and time. The current $\mathbf{J}$ and the resulting wave evolution are determined by three dimensionless small parameters α, β and ϱ. The first small parameter α scales the relative magnitude of the wave nonlinear component and is related to the amplitude of the excitation current. The second parameter β scales the range of the wavevectors $\mathbf{k}$ in a vicinity $\mathbf{k}_*$ involved in the modal composition $\mathbf{J}$, and, consequently, β^{-1} scales the spatial extension of $\mathbf{J}$. Finally, the parameter ϱ scales the frequency bandwidth of $\mathbf{J}$, and, consequently, ϱ^{-1} scales the time extension of $\mathbf{J}$. It turns out [17] that, in particular, ϱ determines the slow time $\tau = \varrho t$ related to the nonlinear evolution.

Supposing that there is an excitation current $\mathbf{J} = \mathbf{J}(\alpha, \varrho, \beta)$ as described above, we consider the resulting wave $\mathbf{U} = \mathbf{U}(\alpha, \varrho, \beta)$ which is a solution to the NLM. The NLM is a rather complicated nonlinear evolution equation for electromagnetic vector fields varying in time and space, and, naturally, we are interested in simpler scalar equations approximating the NLM. It is well known that the NLS is one of such approximations and we are interested in finding how the exact solutions $\mathbf{U}(\alpha, \varrho, \beta)$ of the NLM for small α, β and ϱ for $t > t_0$ are approximated by solutions to an NLS. In our analysis *we take into account all the modes and all possible interactions* as functions of the parameters α, β and ϱ. Using relevant series expansions rigorously justified in [20] we study the exact solution of the NLM for small but still finite values of all three parameters α, β and ϱ, and relate this solution to a solution of a properly tailored NLS. In particular, we show that the scalar amplitudes of the Bloch modes in the modal composition of the solution $\mathbf{U}$ can be approximated by amplitudes of the Fourier modes in the Fourier composition of the solution Z of the relevant NLS with high precision, providing also error estimates. Having a good control over all the steps of the approximation, we identify all additional terms which should be added to the classical NLS to improve the approximation accuracy. Those more accurate equations are referred to as extended NLS (ENLSs). We provide explicit expressions for those additional terms in the ENLS which represent the dominant discrepancy between the exact NLM and its classical NLS approximation. Consequently, the derived ENLSs are intimately related to the NLMs. We provide here some analysis of the ENLS, for more information on the subject see [67] and references therein.

One of interesting results of our quantitative analysis of the NLS regimes for the NLM is their remarkable accuracy for small α, β and ϱ. Namely, quantitative estimates of nonlinear wave interactions show that a properly tailored ENLS can be far more accurate than the classical NLS. *In particular, for the classical NLS characterized by scaling $\alpha \sim \varrho \sim \beta^2$ its approximation accuracy of the NLM is proportional to β whereas properly tailored ENLSs of third and fourth order have the approximation accuracy proportional, respectively, to β^2 and β^3*. An explanation of this phenomenon is based on an analysis of nonlinear wave interactions [17–19]. Namely, we show in following sections that under the condition $\alpha \sim \varrho \sim \beta^2$ the nonlinear wave interactions that lead to the NLS-like regimes and are described by different ENLSs essentially exhaust all significant interactions up to the order β^4 whereas other nonlinear interactions under same conditions are of the order not greater than β^5. *In other words, just by using ENLSs, which are only a little more complex than the classical NLS, we can improve the total approximation accuracy of the NLM on time intervals of order β^{-2} from β to β^3.*

Complete analysis of the accuracy of the approximation of the NLM with the NLS is laborious, and it is helpful to keep in mind the following key elements of that analysis.

- The *dispersion relations $\omega_n(\mathbf{k})$* of the underlying linear periodic medium, with n and $\mathbf{k}$ being respectively the band number and the quasimomentum, *are inversion symmetric*, i.e.

$$\omega_n(-\mathbf{k}) = \omega_n(\mathbf{k}), n = 1, 2, \ldots \tag{2}$$

  The inversion symmetry condition (2) is an important factor for NLS regimes in dielectric media with cubic nonlinearities.
- We use modal decompositions of all involved fields with respect to the related Bloch modes of the underlying linear medium. We consider only weakly nonlinear regimes for which, as it turns out, the modal decomposition is instrumental to the analysis of the wave propagation. The physical and mathematical significance of the spectral decomposition with respect to the Bloch modes lies in the fact that they don't exchange energy under the linear evolution.
- *The NLS regime as a phenomenon of nonlinear wave interactions is characterized by a lack of significant nonlinear interactions and energy exchanges between different spectral bands and different quasimomenta*. More exactly, if the wave is initially composed of eigenmodes characterized by a single band number n_0 and chosen quasimomentum $\pm\mathbf{k}_*$ then under the NLS regime its modal composition remains confined to this band, and its quasimomenta remain close to $\pm\mathbf{k}_*$ for long times with the nonlinear interactions essentially occurring only between this narrow group of quasimomenta, whereas nonlinear interactions with all other bands and quasimomenta are negligibly small.
- The NLS describes approximately the evolution of the modal coefficient of the solution of the NLM generated by a real-valued almost time-harmonic excitation current composed of eigenmodes with a single band number n_0 and the quasimomentum $\mathbf{k}$ from a small vicinity of a chosen point $\mathbf{k}_*$. *The NLS regime is a dielectric medium response to almost time-harmonic excitations.*
- The linear part of the NLS is determined by the second order (or higher order for the ENLS) Taylor polynomial $\gamma_{(2)}(\boldsymbol{\eta})$ of $\omega_{n_0}(\mathbf{k}_* + \boldsymbol{\eta})$ at $\mathbf{k}_*$.

- To relate the NLM and the NLS we introduce spatial and time scales through the excitation currents in the NLM, and then study exact solutions and their asymptotic expansions with respect to the parameters α, β and ϱ assuming that they are small. After that we tailor the parameters of the NLS or an ENLS so that their solutions have the same asymptotic expansions up to a prescribed accuracy. The solutions comparison is carried out after excitation currents are turned off. *We do not make any* a priori *assumptions on the form of solutions to the NLM*, and our analysis of the solutions is not based on any specific ansatz. This allows us not to impose strict functional dependence between the parameters α, β and ϱ, and, consequently, the significance of different terms in the NLS and ENLS and their relation with the NLM can be studied for different ranges of parameters.
- The analysis of involved fields and equations is based on asymptotic series expansions of interaction integrals with respect to small α, ϱ and β and the fourth small parameter which equals either ϱ/β^2 or β^3/ϱ. In other words, we consider two cases: ϱ/β^2 is small or β^3/ϱ is small. *The asymptotic expansions involving β and ϱ stem from oscillatory interaction integrals and they are not Taylor series expansions.*

Following [17–19] we recast the classical NLMs in the following non-dimensional operator form

$$\partial_t \mathbf{U}(\mathbf{r},t) = -\mathrm{i}\mathbf{M}\mathbf{U}(\mathbf{r},t) + \alpha\mathcal{F}_{\mathrm{NL}}(\mathbf{U}(\mathbf{r},t)) - \mathbf{J};\ \mathbf{U}(\mathbf{r},t) = \mathbf{J}(\mathbf{r},t) = 0 \text{ for } t \leq 0 \tag{3}$$

$$\mathbf{U}(\mathbf{r},t) = \begin{bmatrix} \mathbf{D}(\mathbf{r},t) \\ \mathbf{B}(\mathbf{r},t) \end{bmatrix},\ \mathbf{M}\mathbf{U}(\mathbf{r},t) = \mathrm{i}\begin{bmatrix} \nabla\times\mathbf{B}(\mathbf{r},t) \\ -\nabla\times(\varepsilon^{-1}(\mathbf{r})\mathbf{D}(\mathbf{r},t)) \end{bmatrix} \tag{4}$$

$$\mathbf{J}(\mathbf{r},t) = 4\pi\begin{bmatrix} \mathbf{J}_D(\mathbf{r},t) \\ \mathbf{J}_B(\mathbf{r},t) \end{bmatrix}$$

where $\mathcal{F}_{\mathrm{NL}}$ is a nonlinearity with a cubic principal part which may have a general tensorial form, and $\varepsilon(\mathbf{r})$ is the electric permittivity tensor depending on the three-dimensional spatial variable $\mathbf{r} = (r_1, r_2, r_3)$. We consider in this article the case of a lossles medium, i.e. $\varepsilon(\mathbf{r})$ is a Hermitian matrix satisfying

$$\varepsilon(\mathbf{r}) = [\varepsilon(\mathbf{r})]^*, \quad \mathbf{r} = (r_1, r_2, r_3) \tag{5}$$

and our special interest is in the case when the permittivity tensor $\varepsilon(\mathbf{r})$ is also a real symmetric matrix, i.e.

$$\varepsilon(\mathbf{r}) = \{\varepsilon_{jm}(\mathbf{r})\}_{j,m=1}^{3} \text{ where all } \varepsilon_{jm}(\mathbf{r}) = \varepsilon_{mj}(\mathbf{r}) \text{ are real-valued} \tag{6}$$

Notice that the condition (6) implies the inversion symmetry property (2) as well as the complex conjugation property of the eigenmodes (see equation (145)). Though almost all our constructions assume only the inversion symmetry property (2), the dielectric media for which the condition (6) is satisfied get our special attention since they can support real-valued waves described very accurately by the classical NLS. Without the condition (6) but still under the inversion symmetry condition (2), we obtain instead complex-valued waves described by a system of two coupled NLSs (see section 1.4.5 below).

The cases when $\varepsilon(\mathbf{r})$, $\mathbf{J}(\mathbf{r})$ and $\mathbf{U}(\mathbf{r})$ depend only on r_1 or on r_1, r_2 are called, respectively, one-dimensional and two-dimensional. All the fields $\mathbf{D}$, $\mathbf{B}$, $\mathbf{J}_D$ and $\mathbf{J}_B$ are assumed to be divergence free. The dielectric permittivity $\varepsilon(\mathbf{r})$ and the nonlinear polarization $\mathbf{P}_{\mathrm{NL}}(\mathbf{r})$ involved in $\mathcal{F}_{\mathrm{NL}}$ are assumed to be periodic with respect to every r_i, $i = 1, 2, 3$ with the period one for simplicity. The nonlinearity $\mathcal{F}_{\mathrm{NL}}$ originates from the nonlinear polarization which can be written in the following canonical form (see [21]):

$$\mathbf{P}_{\mathrm{NL}}(\mathbf{r},t;\mathbf{E}(\cdot)) = \mathbf{P}^{(3)}(\mathbf{r},t;\mathbf{E}(\cdot)) + \mathbf{P}^{(5)}(\mathbf{r},t;\mathbf{E}(\cdot)) + \cdots \tag{7}$$

where $\mathbf{P}^{(m)}$ is an m-homogeneous operator of the form

$$\mathbf{P}^{(m)}(\mathbf{r}, t; \mathbf{E}(\cdot)) = \int_{-\infty}^{t} \ldots \int_{-\infty}^{t} \mathbf{R}^{(m)}(\mathbf{r}; t - t_1, \ldots, t - t_m) \vdots \prod_{j=1}^{m} \mathbf{E}(\mathbf{r}, t_j) \mathrm{d}t_j \tag{8}$$

with $\mathbf{R}^{(m)}$, $m = 3, 5, \ldots$, describing the medium response. The convergence of the series and the reduction of the nonlinear Maxwell equations to the operator form (3) are discussed in detail in [20]. We consider here the case when the series (7) has only odd order terms that is typical for a medium with central symmetry allowing though the dependence on $\mathbf{r}$ which may be not central symmetric. The parameter α in (3) evidently determines the relative magnitude of the nonlinearity.

We assume that $\alpha \ll 1$, consequently considering weakly nonlinear phenomena. Note that if we rescale $\mathbf{U}$ and $\mathbf{J}$ in (3) by replacing $\mathbf{U}$ by $\xi\mathbf{U}$ and $\mathbf{J}$ by $\xi\mathbf{J}$ with a scaling parameter ξ, we obtain the same equation (3) with α replaced by $\xi^2\alpha$.

Hence, *taking small values for α is equivalent to taking small amplitudes for the excitation current* $\mathbf{J}$ and all three small parameters α, β, and ϱ are ultimately introduced into the NLM through the choice of the excitation current $\mathbf{J}$.

As already mentioned, we assume that the excitation current $\mathbf{J}(\mathbf{r}, t)$ is nonzero only in a finite time interval $[0, \tau_0/\varrho]$, i.e.

$$\mathbf{J}(\mathbf{r}, t) = 0 \text{ if } t \leq 0 \text{ or } t \geq \frac{\tau_0}{\varrho}, \quad \text{where } \tau_0 > 0 \text{ is a small constant} \tag{9}$$

and consider the NLS regime after the current is switched off, i.e. for $t \geq \tau_0/\varrho$. We also assume that the time dependence of the modal coefficients of the currents $\mathbf{J}(\mathbf{r}, t)$ is described by *almost time-harmonic* functions, that is functions of the form:

$$a(t) = a_\varrho(t) = \mathrm{e}^{-\mathrm{i}\omega_0 t}\psi(\varrho t) \text{ where } \psi(\tau) = 0 \text{ for } \tau \leq 0 \text{ and } \tau \geq \tau_0 \tag{10}$$

This type of time dependence corresponds to the well-known slowly-varying-amplitude approximation [22]. Note that since we prescribe this form to the excitation currents that are at our disposal and not to the solutions, no approximation is made yet at this state. It turns out that such currents in both linear and weakly nonlinear regimes generate waves that also have almost time-harmonic amplitudes.

It was shown in [17] and [20] that the exact solution of (3) can be written in the form

$$\mathbf{U}(\mathbf{r}, t) = \mathbf{U}^{(0)}(\mathbf{r}, t) + \alpha\mathbf{U}^{(1)}(\mathbf{r}, t) + O(\alpha^2), \; 0 \leq t \leq \frac{\tau_*}{\varrho}, \; \tau_* \gg \tau_0 \text{ is a constant} \tag{11}$$

We recall that for any quantity ξ the notation $O(\xi)$ stands for any quantity such that

$$|O(\xi)| \leq C|\xi|, \text{ where } C \text{ is a constant} \tag{12}$$

In (11) the term $\mathbf{U}^{(0)}(t)$ is the solution to the linear equation

$$\partial_t \mathbf{U}^{(0)} = -\mathrm{i}\mathbf{M}\mathbf{U}^{(0)} - \mathbf{J}^{(0)}; \; \mathbf{U}^{(0)}(t) = 0 \text{ for } t \leq 0 \tag{13}$$

obtained from (3) by setting there $\alpha = 0$, and we refer to this term as the medium *linear response*. The component $\mathbf{U}^{(1)}(t)$ in (11), called the medium *first nonlinear response (FNLR)*, is a solution of the linear equation obtained by substitution of (11) into (3) with consequent collection of terms proportional to α, namely

$$\partial_t \mathbf{U}^{(1)} = -\mathrm{i}\mathbf{M}\mathbf{U}^{(1)} + \mathcal{F}_{\mathrm{NL}}(\mathbf{U}^{(0)}) - \mathbf{J}^{(1)}; \; \mathbf{U}^{(1)}(t) = 0 \text{ for } t \leq 0 \tag{14}$$

For the reader's convenience we list basic quantities essential for the analysis of the NLM in table 1.

Table 1. The basic quantities needed for NLM analysis.

Basic quantities related to the NLM	
EM wave, a solution to the NLM: six-component vector field	$\mathbf{U}(\mathbf{r}, t)$
Excitation currents	$\mathbf{J}(\mathbf{r}, t)$
Linear part of the NLM	
First-order Hermitian differential operator with 1-periodic coefficients	M
EM wave, a solution to the linear part	$\mathbf{U}^{(0)}(\mathbf{r}, t)$
Dispersion relations of $\mathbf{M}$, generic 2π-periodic functions	$\omega_n(\boldsymbol{k})$
Floquet-Bloch eigenmodes of $\mathbf{M}$	$\tilde{\mathbf{G}}_{\zeta,n}(\mathbf{r}, \mathbf{k})$
Modal coefficients of $\mathbf{U}(\mathbf{r}, t)$ with respect to $\tilde{\mathbf{G}}_{\zeta,n}(\mathbf{r}, \mathbf{k})$	$\tilde{U}_{\zeta,n}(\mathbf{k}, t)$
Phase of the linear wave $\mathbf{U}^{(0)}(\mathbf{r}, t)$	$\zeta\omega_n(\mathbf{k})\frac{\tau}{\varrho}$
Nonlinear part of the NLM	
Tensorial nonlinearity	$\alpha\mathcal{F}_{\mathrm{NL}}(\mathbf{U})$
Cubic susceptibility tensor $\chi^{(3)}(\mathbf{r}; \zeta'\omega_{n'}(\mathbf{k}'), \zeta''\omega_{n''}(\mathbf{k}''), \zeta'''\omega_{n'''}(\mathbf{k}'''))$	
Phase of cubic nonlinear interactions $[\zeta\omega_n(\mathbf{k}) - \zeta'\omega_{n'}(\mathbf{k}') - \zeta''\omega_{n''}(\mathbf{k}'') - \zeta'''\omega_{n'''}(\mathbf{k}''')]\frac{\tau}{\varrho}$	

1.1 *Sketch of nonlinear evolution essentials*

1.1.1 Magnitude, space and time scales. We study NLM–NLS approximations for the following time range

$$\frac{\tau_0}{\varrho} \le t \le \frac{\tau_*}{\varrho}, \ \frac{\tau_*}{\varrho} \le \frac{\alpha_0}{\alpha}, \text{ where } \alpha_0, \ \tau_0, \ \tau_* \text{ are constants} \tag{15}$$

The constant α_0 is related to the convergence of the series (11), and it is independent of the small parameters ϱ and β. Observe that the relations (15) imply that $\alpha \le \varrho\frac{\alpha_0}{\tau_*}$, that is $\alpha = O(\varrho)$ and in the case of a power dependence

$$\alpha \sim \varrho^{\varkappa_0}, \ \varkappa_0 \ge 1 \tag{16}$$

Our primary focus is on an important particular case of (16) when

$$\alpha \sim \varrho, \ \varkappa_0 = 1 \tag{17}$$

and in section 7 we discuss the wave evolution for longer time intervals.

The nonlinear evolution governed by the NLM naturally involves two time scales related to t and $\tau = \varrho t$ (for details see section 5.2). The time (*fast time*) t is just the "real" time, whereas the *slow time* $\tau = \varrho t$ describes a typical time scale for a noticeable nonlinear evolution as in the rescaled NLS (22) below. In other words, $1/\varrho$ is the time for which a noticeable nonlinear evolution can occur. Recasting (15) in terms of the slow time τ we obtain

$$\tau_0 \le \tau \le \tau_*, \text{ where } \tau_0 < \tau_* \text{ are constants} \tag{18}$$

The lesser times $t \le \tau_0/\varrho$, corresponding to transient regimes (see (9)), are beyond the scope of our studies.

Now we give a preliminary sketch of the NLS which approximates the NLM in the one-dimensional case $d = 1$. The NLS has the form

$$\partial_t Z = -\mathrm{i}\gamma_0 Z - \gamma_1\partial_x Z + \mathrm{i}\gamma_2\partial_x^2 Z + \mathrm{i}\alpha q|Z|^2 Z, \ Z(x,t)|_{t=0} = h(\beta x) \tag{19}$$

where

$$\gamma_0 = \omega_{n_0}(\mathbf{k}_*), \ \gamma_1 = \omega'_{n_0}(\mathbf{k}_*), \ \gamma_2 = \frac{1}{2}\omega''_{n_0}(\mathbf{k}_*) \tag{20}$$

Note that the *spatial scale* $1/\beta$ is explicitly introduced in the initial condition for Z in (19). In the rescaled variables

$$\tau = \varrho t, \ \ y = \beta x, \ \ Z(x,t) = z(y,\tau) \tag{21}$$

equation (19) turns into

$$\partial_\tau z = -\mathrm{i}\frac{\gamma_0}{\varrho} z - \frac{\beta\gamma_1}{\varrho}\partial_y z + \mathrm{i}\gamma_2\frac{\beta^2}{\varrho}\partial_y^2 z + \mathrm{i}\frac{\alpha}{\varrho}q|z|^2 z, \ \ z(y,t)|_{\tau=0} = h(y) \tag{22}$$

Evidently, the coefficients of the NLS (22) explicitly depend on the small parameters α, ϱ and β whereas the initial condition does not depend on them. The terms $(\beta\gamma_1/\varrho)\partial_y z$ and $\mathrm{i}\gamma_0/\varrho z$, describing respectively the propagation of the wavepacket with the group velocity $\beta\gamma_1/\varrho$ and time oscillations at the frequency γ_0/ϱ, can be eliminated by a standard change of variables yielding the following reduced classical NLS

$$\partial_\tau z = \mathrm{i}\gamma_2\frac{\beta^2}{\varrho}\partial_y^2 z + \mathrm{i}\frac{\alpha}{\varrho}q|z|^2 z \tag{23}$$

Let us look now at the term $\mathrm{i}\gamma_2(\beta^2/\varrho)\partial_y^2 z$ in (23), describing the linear dispersive effect, and introduce the following parameter

$$\theta = \frac{\varrho}{\beta^2} \tag{24}$$

which we refer to as the *inverse dispersion parameter* since it determines the magnitude of the linear dispersion effects. *It is well known that the ultimate magnitude of nonlinear effects is essentially determined by an interplay between nonlinearity caused by sufficiently large wave amplitudes and the linear wave dispersion causing a reduction of the wave amplitude.* In particular,

$$\begin{aligned} &\text{if } \gamma_2\theta^{-1} \ll 1 \text{ the dispersive effects are weaker} \\ &\text{if } \gamma_2\theta^{-1} \gg 1 \text{ the dispersive effects are stronger} \end{aligned} \tag{25}$$

The significance of the inverse dispersion parameter θ is also supported by our analysis of the error of the NLM–NLS approximation. The dispersive effects already show themselves when θ is fixed and bounded uniformly in β and ϱ. Indeed, in the linear case $\alpha = 0$, the dispersion causes a reduction of the wave amplitude approximately at the rate $(1+\gamma_2\tau/\theta)^{-1/2}$ as the slow time τ increases. In contrast, in the nonlinear case $\alpha \neq 0$ the wave amplitude does not fall with time as in the linear case under assumption that ϱ/α is bounded, indicating a significant nonlinear effect on the wave evolution. In particular, if

$$\alpha \sim \varrho \sim \beta^2 \tag{26}$$

the NLS (23) has soliton solutions with amplitudes that do not fall as τ increases. A qualitative comparative picture of the wave amplitude evolutions for a linear medium versus a nonlinear one is shown in figure 1, which indicates, in particular, that for *for time ranges as in (18) and under conditions (26) the wave evolution shows significant nonlinear effects.*

Note that the effect of the nonlinearity is already significant when the fraction $\tau_0/\tau_* < 1$ in (15) is a fixed number and it does not have to be infinitesimally small.

A closer look at the classical NLS equation (23) shows *that if the small parameters vary so that*

$$\frac{\beta^2}{\varrho} = \theta^{-1} = \text{constant}, \ \ \frac{\alpha}{\varrho} = \text{constant} \tag{27}$$

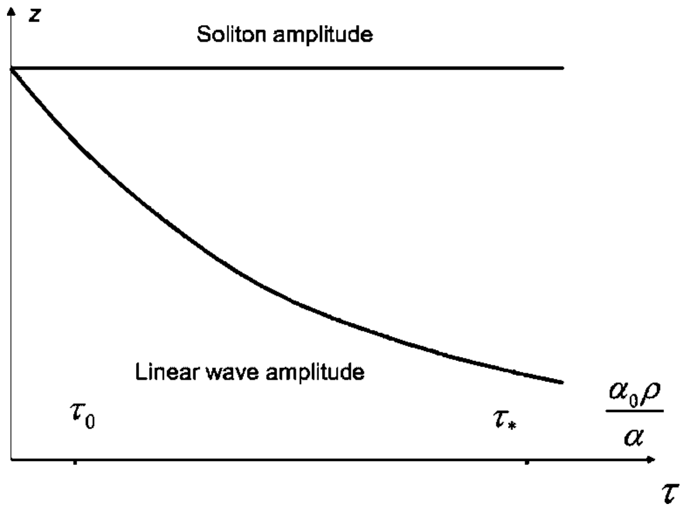

Figure 1. A qualitative comparative picture of the wave amplitude evolution for a linear medium versus a medium with cubic nonlinearity. If $\varrho \sim \alpha \sim \beta^2$, the linear dispersion is exactly balanced by the nonlinearity in a relevant soliton wave.

its form is essentially preserved. We refer to the relations (27) as *classical NLS scaling*. Notice that the condition (26) is an equivalent form of the classical NLS scaling. In particular, from the linear wave dispersion point of view the classical scaling is the marginal case when the inverse dispersion parameter θ is neither infinitesimally small nor large but rather it is finite. *Existence of the solitons manifests the balance between dispersion and nonlinearity* reached at the classical NLS scaling (26).

As to further analysis of the interplay of the linear dispersion with the nonlinearity we consider two cases: (i) $\theta \to 0$; (ii) $\theta \geq \theta_0 > 0$. The first case as $\theta \to 0$, corresponds to stronger dispersion effects and it can be characterized more accurately by the inequality

$$\|\omega''_{n_0}(\mathbf{k}_*)^{-1}\|\theta = \|\omega''_{n_0}(\mathbf{k}_*)^{-1}\|\frac{\varrho}{\beta^2} \ll 1 \tag{28}$$

where $\omega''_{n_0}(\mathbf{k}_*)$ in the multidimensional case $d > 1$ is the matrix of the second differential of $\omega_{n_0}(\mathbf{k})$ at $\mathbf{k}_*$ and the symbol $\|\cdot\|$ stands for the matrix norm. Notice that in the case $d = 1$ the expression $\omega''_{n_0}(\mathbf{k}_*)$ is just the second derivative and $\|\omega''_{n_0}(\mathbf{k}_*)\|$ is its absolute value, implying $\|\omega''_{n_0}(\mathbf{k}_*)^{-1}\| = \|\omega''_{n_0}(\mathbf{k}_*)\|^{-1}$. In view of (25), we refer to the case described by (28) as *dispersive case*. In particular, the dispersive case takes place if

$$\varrho \sim \beta^{\varkappa_1}, \ \varkappa_1 > 2 \tag{29}$$

The other case, $\theta \geq \theta_0 > 0$, occurs if

$$\frac{\|\omega''_{n_0}(\mathbf{k}_*)\|}{\theta} = \frac{\|\omega''_{n_0}(\mathbf{k}_*)\|\beta^2}{\varrho} \ll 1 \quad \text{or} \quad \frac{\|\omega''_{n_0}(\mathbf{k}_*)\|\beta^2}{\varrho} \sim 1 \tag{30}$$

Again, in view of (25), we refer to the case described by (30) as *weakly dispersive*. In particular, the weakly dispersive case takes place if

$$\varrho \sim \beta^{\varkappa_1}, \ 2 \geq \varkappa_1 > 0 \tag{31}$$

Notice that the classical NLS scaling is covered by the second alternative condition of the weakly dispersive case (30) and (31), namely $\varkappa_1 = 2$ and

$$\frac{|\omega''_{n_0}(\mathbf{k}_*)|}{\beta^2} \sim \frac{1}{\varrho} \tag{32}$$

Formula (32) determines the time scale on which the dispersive effects are already significant, as one can see from figure 1. We will consider the weakly dispersive case in more detail since it exhibits stronger nonlinear effects.

For both the dispersive and weakly dispersive cases we get the same set of NLSs or ENLSs, but the properties of the equations are different in different ranges of the parameters. Mathematical techniques used to study them are different as well. Namely, in the dispersive case of (28) we apply the stationary phase method. The weakly dispersive case (30) is technically simpler and is studied based on the Taylor expansion of relevant oscillatory phases. Remarkably, in both cases the dynamics of the directly excited modes is explicitly expressed in terms of a solution of the same NLS.

In addition to the above conditions, we assume that ϱ is small enough to provide the condition

$$\frac{\omega_{n_0}(\mathbf{k}_*)}{\varrho} \gg 1 \tag{33}$$

which signifies the relevance of the frequency matching condition. We also assume the following condition

$$\frac{|\omega'_{n_0}(\mathbf{k}_*)|\beta}{\varrho} \gg 1 \tag{34}$$

which allows use of the group velocity for the analysis of wave interactions. Note though that the conditions (33) and (34) are not always necessary.

1.1.2 Relation between the NLM and the NLS. Observe that the excitation current $\mathbf{J}$ determines uniquely the solution of the NLM (3) whereas the initial data h determines the solution of the NLS (19). Consequently, if we want to select regimes of the NLM that are well approximated by solutions of the NLS we have to (i) construct the NLS, in other words to determine its coefficients, based on the NLM and (ii) describe the correspondence between $\mathbf{J}$ and h. It turns out that the current $\mathbf{J}$ should be of the form

$$\mathbf{J}(\mathbf{r},t) = \mathbf{J}^{(0)}(\mathbf{r},t) + \alpha\mathbf{J}^{(1)}(\mathbf{r},t),\ \mathbf{J}^{(j)}(\mathbf{r},t) = 0 \text{ if } t \le 0 \quad \text{or} \quad t \ge \frac{\tau_0}{\varrho},\ j = 0,1 \tag{35}$$

where the principal part $\mathbf{J}^{(0)}$ and the corrective part $\mathbf{J}^{(1)}$ of the current $\mathbf{J}$ are properly selected to produce a NLS-type regime (see sections 2.1 and 5.2 for details). Notice that the current $\mathbf{J}$ substitutes for the initial data for the NLM and is based on the initial data of the NLS. As was explained in [17], the introduction of the excitation current $\mathbf{J}$ is both mathematically and physically a more suitable option for the NLM having nonlinear polarization of the form (8), since a prescription of instantaneous initial data for $t = 0$ is inconsistent with the form of the nonlinearity (8) which requires one to know the fields at all times. The standard classical NLS though assumes the prescription of an instantaneous initial data at $t = 0$. We overcome this difference by setting a proper form for the current $\mathbf{J}(\mathbf{r},t)$ including, in particular, its composition of the form (35). To produce NLS-type regimes the current $\mathbf{J}$ has to possess two properties. Firstly, it should be almost time-harmonic, as in (10) with a deviation from time-harmonicity measured by a small parameter ϱ which consequently determines the ratio of the slow time and the fast time scales. Secondly, following the framework described in [17–20] we choose the excitation current composed of Bloch modes from a single spectral band, described

Table 2. The basic quantities needed for NLS analysis.

Basic quantities related to the NLS	
Wave, a solution to the NLS: scalar function	$Z_\zeta(\mathbf{r}, t)$
Initial data	$h(\beta r)$
Linear part of the NLS	
Second-order Hermitian differential operator	$\gamma_{(2)}(-\mathrm{i}\partial_r)$
Wave, a solution to the linear part	$Z_\zeta^{(0)}(\mathbf{r}, t)$
Dispersion relations of $\gamma_{(2)}(-\mathrm{i}\vec{\nabla}_{\mathrm{r}})$, a polynomial	$\gamma_{(2)}(\boldsymbol{\eta})$
Fourier eigenmodes of $\gamma_{(2)}(-\mathrm{i}\vec{\nabla}_{\mathrm{r}})$: exponentials	$e^{i\mathbf{r}\cdot\eta}$
Modal coefficients of $Z_\zeta(\mathbf{r}, t)$ with respect to $e^{i\mathbf{r}\cdot\eta}$	$\hat{Z}_\zeta(\boldsymbol{\eta}, t)$
Phase of the linear wave $Z_\zeta^{(0)}(\mathbf{r}, t)$	$\gamma_{(2)}(\boldsymbol{\eta})\frac{\tau}{\varrho}$
Nonlinear part of the NLS	
Scalar nonlinearity	$\alpha Q_\zeta \lvert Z\rvert^2 Z$
Phase of nonlinear interactions	
$[\gamma_{(2)}(\zeta\beta\boldsymbol{\eta}) - \gamma_{(2)}(\zeta\beta\boldsymbol{\eta}') - \gamma_{(2)}(\zeta\beta\boldsymbol{\eta}'') + \gamma_{(2)}(-\zeta\beta\boldsymbol{\eta}''')]\frac{\tau}{\varrho}$	

by an index $n = n_0$, and with the quasimomenta $\mathbf{k}$ from a small β-vicinity $|\mathbf{k} \pm \mathbf{k}_*| = O(\beta)$ of a fixed quasimomentum $\pm\mathbf{k}_*$ in the Brillouin zone. The reason for having two quasimomenta $\pm\mathbf{k}_*$ (a doublet) rather than just one $\mathbf{k}_*$ is that it is the minimal set of quasimomenta producing a real-valued $\mathbf{U}(\mathbf{r}, t)$. In the case of the NLS the parameter β is introduced through its initial data $h_\pm(\beta\mathbf{r})$ at $t = 0$. Then we provide an explicit construction of the excitation current $\mathbf{J}(\mathbf{r}, t)$ based on the prescribed initial data $h_\pm(\beta\mathbf{r})$ for the NLS. Notice that β^{-1} determines the length scale for $\mathbf{J}(\mathbf{r}, t)$ too. Thus, both parameters ϱ and β are introduced into the NLM via the excitation current $\mathbf{J}(\mathbf{r}, t)$. The two relevant modal coefficients near $\pm\mathbf{k}_*$ in the composition of the solution $\mathbf{U}(\mathbf{r}, t)$ to the NLM are approximated by Fourier modes of two scalar functions $Z_\pm(\mathbf{r}, t)$. These two intimately related scalar functions $Z_\pm(\mathbf{r}, t)$ satisfy two related NLSs. We write now these equations in the simpler one-dimensional case $d = 1$, i.e. when the medium coefficients and solutions of (3) depend only on the coordinate $r_1 = x$ and do not depend on the remaining coordinates r_2 and r_3. The equations have the form

$$\begin{aligned} \partial_t Z_+ &= -\mathrm{i}\gamma_0 Z_+ - \gamma_1 \partial_x Z_+ + \mathrm{i}\gamma_2 \partial_x^2 Z_+ + \alpha_\pi Q_+ Z_- Z_+^2 \\ Z_+(x, t)|_{t=0} &= h_+(\beta x),\ \alpha_\pi = 3\alpha(2\pi)^2, \end{aligned} \tag{36}$$

$$\begin{aligned} \partial_t Z_- &= \mathrm{i}\gamma_0 Z_- - \gamma_1 \partial_x Z_- - \mathrm{i}\gamma_2 \partial_x^2 Z_- + \alpha_\pi Q_- Z_-^2 Z_+ \\ Z_-(x, t)|_{t=0} &= h_-(\beta x),\ h_-(\beta x) = h_+^*(\beta x) \end{aligned} \tag{37}$$

with the asterisk denoting the complex conjugation. In (36) and (37) the coefficients $\gamma_0, \gamma_1, \gamma_2$ satisfy (20) and $h_+(x)$ is a smooth function decaying sufficiently fast as $x \to \infty$. The function $h_+(x)$ can be chosen as we please. The coefficients $Q_\pm$ in (36) and (37) are certain complex-valued numbers related to the the third-order susceptibility tensor associated with the cubic nonlinearity $\mathcal{F}_{\mathrm{NL}}(\mathbf{U})$. The relevant properties of the NLS are given in table: which can be compound with table 1 for the NLM. We do not impose any conditions on the structure of the cubic tensor in $\mathcal{F}_{\mathrm{NL}}(\mathbf{U})$, which affects only the values of the coefficients $Q_\pm$ in (36) and (37). With no structural conditions imposed on the tensors related to the nonlinearity, the complex coefficients $Q_\pm$ may be such that $Q_+ \neq Q_-^*$. In the latter case $Z_-(x, t)$ might be different from $Z_+^*(x, t)$. *Though in the case when the nonlinearity maps real-valued fields into real-valued and (6) holds we always have*

$$Z_-(x, t) = Z_+^*(x, t),\ Z_- Z_+^2 = |Z_+|^2 Z_+ \tag{38}$$

and the system (36) and (37) is effectively reduced to a single scalar equation (36):

$$\partial_t Z = -\mathrm{i}\gamma_0 Z - \gamma_1 \partial_x Z + \mathrm{i}\gamma_2 \partial_x^2 Z + \alpha_\pi Q_+ |Z|^2 Z,\ Z(x, t)|_{t=0} = h_+(\beta x) \tag{39}$$

The two functions $Z_{\pm}(x,t)$ satisfying the NLSs (36) and (37) yield an approximation $\mathbf{U}_Z(\mathbf{r},t)$ to the exact solution $\mathbf{U}(\mathbf{r},t)$ of the NLM. An analysis of the approximate solution $\mathbf{U}_Z(\mathbf{r},t)$ leads to a natural partition of modes involved in its composition into two groups: 'directly' and 'indirectly' excited modes, and it suggests splitting of the approximate solution into two parts

$$\mathbf{U}_Z(\mathbf{r},t) = \mathbf{U}_Z^{\mathrm{dir}}(\mathbf{r},t) + \mathbf{U}_Z^{\mathrm{ind}}(\mathbf{r},t) \tag{40}$$

The *directly excited modes* that contribute to $\mathbf{U}_Z^{\mathrm{dir}}$ are the ones presented in the excitation current $\mathbf{J}$ and excited through the linear mechanism, i.e. when $\alpha = 0$, whereas *indirectly excited modes*, which form $\mathbf{U}_Z^{\mathrm{ind}}$, are excited only through the nonlinearity and, consequently, for $\alpha = 0$ their amplitudes are zero (see section 3 for details). The modal coefficients of the indirectly excited modes are much smaller than those related to the directly excited modes, therefore $\mathbf{U}_Z^{\mathrm{ind}}$ is much smaller than $\mathbf{U}_Z^{\mathrm{dir}}$, i.e.

$$\left|\mathbf{U}_Z^{\mathrm{ind}}\right| = O(\alpha) O\left(|\mathbf{U}_Z^{\mathrm{dir}}|\right) \tag{41}$$

It turns out that high-precision approximations for the modal coefficients of the indirectly excited modes are based only on the FNLR, and consequently are expressed in terms of the excitation currents and do need the NLS. *In contrast, approximations of the same accuracy for directly excited modes are ultimately reduced to relevant NLSs which account for nonlinear self-interactions of these modes.*

The directly excited part of the approximate solution $\mathbf{U}_Z$ has the following form in the space domain:

$$\mathbf{U}_Z^{\mathrm{dir}}(\mathbf{r},t) = \tilde{\mathbf{G}}_{+,n_0}(\mathbf{r},\mathbf{k}_*) Z_+(\mathbf{r},t) + \tilde{\mathbf{G}}_{-,n_0}(\mathbf{r},\mathbf{k}_*) Z_-(\mathbf{r},t) + \beta \mathbf{U}_{Z,n_0}^1(\mathbf{r},t) + O(\beta^2) \tag{42}$$

with $\tilde{\mathbf{G}}_{\zeta,n_0}(\mathbf{r},\mathbf{k}_*)$, $\zeta = \pm$, being Bloch eigenmodes of the linear Maxwell operator $\mathbf{M}$. We recall that

$$\tilde{\mathbf{G}}_{\zeta,n_0}(\mathbf{r},\mathbf{k}_*) = \mathrm{e}^{\mathrm{i}\mathbf{k}_*\cdot\mathbf{r}} \hat{\boldsymbol{G}}_{\zeta,n_0}(\mathbf{r},\mathbf{k}_*) \quad \text{where } \hat{\boldsymbol{G}}_{\zeta,n_0}(\mathbf{r},\mathbf{k}_*) \text{ is periodic in } \mathbf{r} \tag{43}$$

The next order correction $\mathbf{U}_Z^1(\mathbf{r},t)$ in this representation in the one-dimensional case when $\mathbf{r} = x$ is given by the following formula (see section 5.5 for the general case of the space dimensions 2 and 3)

$$\mathbf{U}_Z^1(\mathbf{r},t) = \mathbf{U}_{Z_+}^1(x,t) + \mathbf{U}_{Z_-}^1(x,t) \tag{44}$$

$$\mathbf{U}_{Z_\zeta}^1(x,t) = -\zeta \mathrm{i}\mathrm{e}^{\mathrm{i}\zeta k_*\cdot x} \beta^{-1} \partial_x Z_\zeta(x,t) \partial_k \hat{\boldsymbol{G}}_{\zeta,n_0}(x,k_*), \ \zeta = \pm \tag{45}$$

This correction reflects finer effects of the periodicity of the medium and it is present even in the linear case when $\alpha = 0$. The terms $\beta^{-1}\partial_x Z_{\pm}(x,t)$ in (45) are bounded for small β because, after the rescaling (21), $\beta^{-1}\partial_x Z_{\pm}(x,t)$ equals $\partial_y z_{\pm}(y,t)$ where $z_{\pm}(y,t)$ solve equations of the form (22). Note also that the approximate expression (42) is not an ansatz, it is a consequence of the exact formula (58) written in section 1.2 in terms of the Floquet–Bloch transform. Note that, according to (42) and (43), the quasimomentum $\mathbf{k}_*$ describes the phase shift of the carrier wave over the period cell.

The difference between the approximate solution $\mathbf{U}_Z(\mathbf{r},t)$, based on the NLS equations (36), (37) and the exact solution $\mathbf{U}(\mathbf{r},t)$ is called the *approximation error*. Using the modal decomposition and analytic methods developed in [17–20] we proved the following estimate for the approximation error:

$$\mathbf{U}(\mathbf{r},t) - \mathbf{U}_Z(\mathbf{r},t) = O(\beta) + O(\varrho) \tag{46}$$

with the symbol $O(\eta)$ defined by (12).

The approximation error $\mathbf{U}(\mathbf{r}, t) - \mathbf{U}_Z(\mathbf{r}, t)$ can be reduced by adding certain corrective terms to the NLSs (36) and (37). We call such equations with added corrective terms *extended NLSs (ENLSs)* (see section 1.4 for details). The simplest ENLSs have corrective terms of the form $\partial_x^3 Z_+$, $Z_- \partial_x Z_+^2$ and $Z_+^2 \partial_x Z_-$ with calculable coefficients $Q_{1,\pm}$, and they are as follows

$$\begin{aligned} \partial_t Z_\zeta &= -\mathrm{i}\zeta\gamma_0 Z_\zeta - \gamma_1 \partial_x Z_\zeta + \mathrm{i}\zeta\gamma_2 \partial_x^2 Z_\zeta + \gamma_3 \partial_x^3 Z_\zeta + \alpha_\pi [Q_\zeta Z_\zeta^2 Z_{-\zeta} \\ &\quad + Q_{1,\zeta} Z_\zeta Z_{-\zeta} \partial_x Z_\zeta + Q_{1,*,\zeta} Z_\zeta^2 \partial_x Z_{-\zeta}] \\ Z_\zeta(x,t)|_{t=0} &= h_\zeta(\beta x), \ \zeta = \pm \end{aligned} \tag{47}$$

The coefficients $Q_{1,\pm}$, $Q_{1,*,\pm}$ in (47) take into account the dependence of the susceptibility and Bloch eigenfunctions on the Bloch spectral variable $\mathbf{k}$ (the quasimomentum) which are neglected in the standard NLSs (36) and (37). In the real-valued case we use (38) to reduce the two equations (47) to one equation, which after the change of variables $\beta x = y$ takes a form similar to (22), i.e.

$$\begin{aligned} \partial_\tau z &= -\mathrm{i}\frac{\gamma_0}{\varrho} z - \frac{\beta\gamma_1}{\varrho}\partial_y z + \frac{\beta^2}{\varrho}\left[\mathrm{i}\gamma_2 \partial_y^2 z + \beta\gamma_3 \partial_y^3 z\right] \\ &\quad + \frac{\alpha}{\varrho}[\mathrm{i}q_0|z|^2 z + \beta q_1 |z|^2 \partial_y z + \beta q_{1,*,+} z^2 \partial_y z^*], \quad z(y,t)|_{\tau=0} = h(y) \end{aligned} \tag{48}$$

Table 3 lists additional terms of the order of β showing their relations to the NLM. If $Z_\pm$ are solutions to the ENLS (47) then the approximate solution of the NLM given by (42) with $\mathbf{U}_Z$ based on (47) provides a better approximation of $\mathbf{U}$ than with $\mathbf{U}_Z$ based on the standard NLSs (36) and (37), i.e.

$$\mathbf{U}(\mathbf{r}, t) - \mathbf{U}_Z(\mathbf{r}, t) = O(\beta^2) + O(\varrho) \tag{49}$$

Estimates (49) and (46) show that the introduction of corrective terms into the ENLS improves the accuracy of the approximation, namely the error term $O(\beta)$ is replaced by a smaller $O(\beta^2)$. Evidently that is a significant improvement when $\varrho \sim \beta^2$. Such a refinement of the approximation is possible due to the specific form of matching between solutions of the NLS and the NLM which is described in the next section 1.2, see (58). In section 1.4 we consider an ENLS having more corrective terms and yielding even better approximations.

Another way to construct approximate solutions of the NLM is by using not the differential equations of the form of NLS or ENLS, but rather finite-difference lattice equations (see section 9 for details). In the one-dimensional case the lattice counterpart of equation (39) is

$$\begin{aligned} \partial_t Z_+(m) &= -\mathrm{i}(\gamma_0 + \gamma_2) Z_+(m) - \gamma_1 \left(\frac{1}{2}[Z_+(m+1) - Z_+(m-1)]\right) \\ &\quad + \mathrm{i}\frac{\gamma_2}{2}[Z_+(m+1) + Z_+(m-1)] + \alpha_\pi Q_+ |Z_+(m)|^2 Z_+(m) \\ Z_+(m,t)|_{t=0} &= h_+(\beta m), \ \alpha_\pi = 3\alpha(2\pi)^2, \ m = \cdots - 1, 0, 1, 2, \ldots \end{aligned} \tag{50}$$

Table 3. Additional terms in the third-order ENLS which improve the accuracy of the NLM–NLS approximation.

Source in the NLM	Term in the ENLS
Dispersion relation	$\beta\gamma_3\partial_y^3 z$
Susceptibility	$\beta q_1 |z|^2 \partial_y z + \beta q_{1,*,+} z^2 \partial_y z^*$

Table 4. The origin of terms in the classical second-order NLS as an approximation of the NLM.

NLM characteristics	Mechanism of correspondence	NLS characteristics
Dispersion relation $\omega_{n_0}(\mathbf{k})$	$\gamma_{(2)}(\boldsymbol{\eta})$ is Taylor polynomial of $\omega_{n_0}(\mathbf{k})$ for $\mathbf{k} = \mathbf{k}_* + \boldsymbol{\eta}$	Dispersion relation $\gamma_{(2)}(\boldsymbol{\eta})$
Nonlinearity $\mathcal{F}_{\mathrm{NL}}(\mathbf{U})$ and the susceptibility $\chi^{(3)}$	The susceptibility $\chi^{(3)}$ and modes $\tilde{\mathbf{G}}_{\zeta,n_0}(\mathbf{r},\mathbf{k})$ at $\mathbf{k} = \mathbf{k}_*$ determine Q_+	Coefficient Q_+ at the nonlinearity $\vert Z_+\vert^2 Z_+$

The lattice equation (50) is obtained by a direct approximation of the NLM, and it is not a finite-difference approximation of the NLS (36). Technically, approximations of dispersion relations by algebraic polynomials yield differential operators whereas approximations by trigonometric polynomials yield finite-difference lattice operators. Instead of (42) a similar formula holds with the same leading term (see equation (477–480) for details). *The accuracy of the approximation of the NLM in terms of the lattice NLS is the same*, it is given by (46).

Summarizing, we single out the following factors essential for forming NLS-type regimes of the NLM and for determining the coefficients of the relevant NLS or ENLS:

- dispersion relations $\omega_n(\mathbf{k})$;
- band number n_0, quasimomentum $\mathbf{k}_*$ and the dispersion relation $\omega_{n_0}(\mathbf{k})$ which determine, in particular, the wave carrier frequency $\omega_{n_0}(\mathbf{k}_*)$;
- the susceptibility tensor $\chi^{(3)}$;
- the Bloch mode $\tilde{\mathbf{G}}_{\zeta,n_0}(\mathbf{r},\mathbf{k}_*)$ corresponding to the band n_0 and quasimomentum $\mathbf{k}_*$;
- the chosen order ν of the NLS which often equals 2.

Table 4 shows elements of the construction of the classical second-order NLS, for the order $\nu = 2$. After the value of ν is chosen and the NLS is constructed we move to the construction of NLS-type solutions for the NLM based on the initial data $h(\beta\mathbf{r})$. Such NLS-type solutions are constructed by setting a proper expression for the excitation currents $\mathbf{J}$ in terms of the initial data $h(\beta\mathbf{r})$.

1.2 *Basics of the modal analysis*

Following [17–20] we study the NLMs in periodic media based on the Floquet–Bloch modal decomposition. The importance and even necessity of such a decomposition is based on the absence of the energy transfer between Bloch modes in the linear approximation, which is instrumental for the construction of the perturbation theory of the nonlinear evolution. As long as the amplitude of the wave component due to the nonlinearity does not exceed the amplitude of its linear component, the Floquet–Bloch modal expansions continue to be an excellent framework capturing well the nonlinear evolution. The Floquet-Bloch expansion of the exact solution of (3) has the form

$$\mathbf{U}(\mathbf{r},t) = \frac{1}{(2\pi)^d} \sum_{\bar{n}} \int_{[-\pi,\pi]^d} \tilde{U}_{\bar{n}}(\mathbf{k},t)\tilde{\mathbf{G}}_{\bar{n}}(\mathbf{r},\mathbf{k})\,\mathrm{d}\mathbf{k} \tag{51}$$

where $\tilde{\mathbf{G}}_{\bar{n}}(\mathbf{r},\mathbf{k})$ are the Bloch eigenfunctions corresponding to the eigenvalues $\omega_{\bar{n}}(\mathbf{k})$ of the Maxwell operator $\mathbf{M}$ and $\mathbf{k}$ is the quasimomentum with values in the Brillouin zone $[-\pi,\pi]^d$. The scalar functions $\tilde{U}_{\bar{n}}(\mathbf{k},t)$ in (51) are the *modal coefficients* of $\mathbf{U}(\mathbf{r},t)$ corresponding to the mode $(\bar{n},\mathbf{k})$. In the combined index $\bar{n} = (\zeta, n)$, the integer index $n = 1, 2, \ldots$ is the band number and the binary index $\zeta = \pm 1$ labels two conjugate eigenfunctions of the Maxwell operator $\mathbf{M}$ with opposite eigenvalues $\omega_{\bar{n}}(\mathbf{k}) = \omega_{\zeta,n}(\mathbf{k}) = \zeta\omega_n(\mathbf{k})$. The field

$$\tilde{\mathbf{U}}(\mathbf{k},\mathbf{r},t) = \sum_{\bar{n}} \tilde{U}_{\bar{n}}(\mathbf{k},t)\tilde{\mathbf{G}}_{\bar{n}}(\mathbf{r},\mathbf{k}) = \sum_{\zeta=\pm1}\sum_{n=1}^{\infty} \tilde{U}_{\zeta,n}(\mathbf{k},t)\tilde{\mathbf{G}}_{\zeta,n}(\mathbf{r},\mathbf{k}) \tag{52}$$

which is the integrand of the integral in the right-hand side of (51), is called the Floquet-Bloch transform of $\mathbf{U}(\mathbf{r},t)$, see [17] for details. By setting $\alpha = 0$, $\mathbf{J} = \mathbf{0}$ in (3) we obtain the linear homogeneous Maxwell equation:

$$\partial_t \mathbf{U}(\mathbf{r},t) = -\mathrm{i}\mathbf{M}\mathbf{U}(\mathbf{r},t) \tag{53}$$

Its general solution has the following Floquet–Bloch transform

$$\tilde{\mathbf{U}}(\mathbf{k},\mathbf{r},t) = \sum_{\bar{n}} \tilde{u}_{\bar{n}}(\mathbf{k})\mathrm{e}^{-\mathrm{i}\omega_{\bar{n}}(\mathbf{k})t}\tilde{\mathbf{G}}_{\bar{n}}(\mathbf{r},\mathbf{k}) \tag{54}$$

If we ask now what kind of current $\mathbf{J}(\mathbf{r},t)$ can produce a wave that evolves essentially according to an NLS the answer is as follows. We set $\mathbf{J}(\mathbf{r},\mathbf{t})$, firstly, to be of the form (35) and composed of eigenmodes with a single band number n_0, and, secondly, we set the modal form of its principal part $\mathbf{J}^{(0)}(\mathbf{r},\mathbf{t})$ to be as follows

$$\begin{aligned}
&\tilde{\mathbf{J}}^{(0)}_{n_0}(\mathbf{r},\mathbf{k},t) = \tilde{j}^{(0)}_{+,n_0}(\mathbf{k},\tau)\tilde{\mathbf{G}}_{+,n_0}(\mathbf{r},\mathbf{k})\mathrm{e}^{-\mathrm{i}\omega_{n_0}(\mathbf{k})t} + \tilde{j}^{(0)}_{-,n_0}(\mathbf{k},\tau)\tilde{\mathbf{G}}_{-,n_0}(\mathbf{r},\mathbf{k})\mathrm{e}^{\mathrm{i}\omega_{n_0}(\mathbf{k})t}\\
&\tilde{j}^{(0)}_{\zeta,n_0}(\mathbf{k},\tau) = -\varrho\beta^{-d}\psi_0(\tau)\Psi(\mathbf{k}-\zeta\mathbf{k}_*)\hat{h}_\zeta\left(\frac{1}{\beta}(\mathbf{k}-\mathbf{k}_*)\right),\ \tau = \varrho t,\ \zeta = \pm\\
&\tilde{\mathbf{J}}^{(0)}_{n}(\mathbf{r},\mathbf{k},t) = 0,\ n \neq n_0
\end{aligned} \tag{55}$$

We call such an excitation current *almost single-mode excitation*. Evidently, the current $\mathbf{J}^{(0)}(\mathbf{r},\mathbf{k},t)$ defined by (55) is an almost time-harmonic function of the time t as in (10) for every $\mathbf{k}$. Observe also that in (55) $\mathbf{k}_*$ is a chosen quasimomentum in the Brillouin zone $[-\pi,\pi]^d$. The currents of the above form are determined by the choice of the function $\beta^{-d}\hat{h}_\pm(\boldsymbol{\eta}/\beta)$, which is the Fourier transform of the function $h_\pm(\beta\mathbf{r})$, which, in turn, corresponds to the initial data of the NLS. Therefore, $h_\pm(\beta\mathbf{r})\tilde{\mathbf{G}}_{\pm,n_0}(\mathbf{r},\mathbf{k}_*)$ is a proper substitute for the initial data for the NLM. Note that for small β the spread of the function $h(\beta\mathbf{r})$ is large and proportional to $1/\beta$, whereas the spread of its Fourier transform $\beta^{-d}\hat{h}_\pm(\boldsymbol{\eta}/\beta)$ is small and proportional to β. The cut-off function Ψ in (55) is introduced to restrict $\hat{h}_\pm(\boldsymbol{\eta}/\beta)$ from the entire space to the Brillouin zone $[-\pi,\pi]^d$ and its properties are listed in (190). The slowly varying function $\varrho\psi_0(\varrho t)$ is set to be non-zero only for $0 \leq \tau = \varrho t \leq \tau_0$. Its purpose is to provide a transition from the rest solution to a nonzero solution of the NLM, and also to introduce a finite, proportional to ϱ frequency bandwidth, and, consequently, the slow time scale $\tau = \varrho t$, into the excitation current. We refer to currents and waves of the form similar to (55) as *almost single-mode waves*. The concept of an almost single-mode wave is instrumental in studies on nonlinear wave interactions and NLS regimes.

To explain the construction of an NLS corresponding to the NLM we introduce first an abstract nonlinear equation for a two-component vector valued amplitude $\vec{V}$

$$\partial_t\vec{V} = -\mathrm{i}\mathcal{L}\vec{V} + \alpha F^{(3)}(\vec{V}) - \vec{J}_V \tag{56}$$

where $\mathcal{L}$ is a linear differential operator with constant coefficients, $F^{(3)}(\vec{V})$ is a cubic nonlinearity with the simplest possible structure and $J_V = 0$ when $t < 0$ and $t > \tau_0/\varrho$. The vector $\vec{V}$ in (56) includes two components which correspond to two modes $\pm\mathbf{k}_*$ excited by a real-valued almost single-mode current regime. Our goal is to construct $\mathcal{L}$ and $F^{(3)}$ and choose J_V so that the sum of the linear and the first nonlinear responses associated with (56) would approximate well the directly excited modal coefficients $\tilde{U}_{\zeta,n_0}(\mathbf{k},t)$ when $t > \tau_0/\varrho$. The correspondence

Table 5. Simplified form of the relation between the NLS as it approximates the NLM.

NLM	Mechanism of correspondence	NLS
Solution $\mathbf{U}(\mathbf{r}, t)$	Matching modal coefficients	Solution $Z(\mathbf{r}, t)$
Phase $\omega_n(\mathbf{k}_* + \boldsymbol{\eta})\frac{\tau}{\varrho}$	Taylor polynomial	Phase $\gamma_{(2)}(\boldsymbol{\eta})\frac{\tau}{\varrho}$
Modal coefficient $\tilde{U}_{\zeta,n_0}(\mathbf{k}, t)$	$\tilde{U}_{\zeta,n_0}(\mathbf{k}_* + \boldsymbol{\eta}, t) = \hat{Z}_\zeta(\boldsymbol{\eta}, t)$	Fourier coefficient $\hat{Z}_\zeta(\boldsymbol{\eta}, t)$
Excitation current $\mathbf{J}(\mathbf{r}, t)$	$\mathbf{J}$ is of the form $\psi(\varrho t)\boldsymbol{\Psi}(\mathbf{r}, t)$, where ψ is a cutoff function, and $\boldsymbol{\Psi}(\mathbf{r}, t)$ is determined by the initial data $h(\beta r)$	Initial data $h(\beta r)$

between NLS and NLM is shown in table 5. We rigorously show that equations providing accurate approximations to the NLM are of the form (56), in particular, they are the classical NLS or ENLS for higher order approximations. An analysis shows that for excitation currents as in (55) only the modes close to $\pm\mathbf{k}_*$ interact nonlinearly with themselves strongly enough to determine the nonlinear evolution, whereas all other nonlinear interactions are generically negligible. Note that the difference between the NLS and NLM is obvious even when the nonlinearity is absent, since the NLM is an equation with variable coefficients for six-component vector fields which includes only first-order spatial derivatives whereas the NLS has two components (reducible to one by complex conjugation) with constant coefficients and with second-order spatial derivatives.

The relation between the NLM and corresponding NLS is as follows. The coefficients of the NLS can be explicitly written in terms of the Bloch dispersion relations, the eigenfunctions and the cubic susceptibility. Then the approximate solution $\mathbf{U}_Z(\mathbf{r}, t)$ of the NLM is constructed based on solutions $Z_\pm(\mathbf{r}, t)$ to the NLS (36), (37) by formula (40) where the leading, directly excited component does not include modes that are not present in the excitation current

$$\mathbf{U}_{Z,n}^{\mathrm{dir}}(\mathbf{r}, t) = 0, \ \ n \neq n_0 \tag{57}$$

and the component in the excited band is given by the following fundamental formula

$$\mathbf{U}_{Z,n_0}^{\mathrm{dir}}(\mathbf{r}, t) = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \Psi(\eta) \times [\hat{Z}_+(\boldsymbol{\eta}, t)\tilde{\mathbf{G}}_{+,n_0}(\mathbf{r}, \mathbf{k}_* + \boldsymbol{\eta}) + \hat{Z}_-(\boldsymbol{\eta}, t)\tilde{\mathbf{G}}_{-,n_0}(\mathbf{r}, -\mathbf{k}_* + \boldsymbol{\eta})]\,\mathrm{d}\boldsymbol{\eta} \tag{58}$$

where $\hat{Z}_\pm(\mathbf{k}, t)$ is the Fourier transform of $Z_\pm(\mathbf{r}, t)$, $t \geq \tau_0/\varrho$. The cut-off function $\Psi(\boldsymbol{\eta})$ is introduced to select only $\boldsymbol{\eta}$ from a fixed small vicinity of $\mathbf{k}_*$ in the Brillouin zone. The components $\mathbf{U}_{Z,n}(\mathbf{r}, t)$ with $n \neq n_0$ of $\mathbf{U}_Z(\mathbf{r}, t)$ are included in the indirectly excited part $\mathbf{U}_Z^{\mathrm{ind}}(\mathbf{r}, t)$; they are much smaller and are described in the end of this subsection. Formula (58) shows that the dynamics of the solution of the Maxwell equation on time intervals of order $1/\varrho$ is reduced to the dynamics of solutions to the NLS or ENLS. Formula (58) also shows that the time evolution of the pair of the coefficients $\tilde{U}_{\zeta,n_0}(\mathbf{k}, t)$, $\zeta = \pm 1$, determined by the almost single-mode excitations, is described by solutions $\hat{Z}_\pm$ to the NLS.

The relation between the modal coefficients of the approximate solution $\mathbf{U}_Z(\mathbf{r}, t)$ of the NLS and the exact solution $\mathbf{U}(\mathbf{r}, t)$ of the nonlinear Maxwell equation is represented by the formula

$$\tilde{U}_{\zeta,n_0}(\zeta\mathbf{k}_* + \boldsymbol{\eta}, t) = \hat{Z}_\zeta(\boldsymbol{\eta}, t) + O(\beta) + O(\varrho) \tag{59}$$

Equality (59) holds when (15) is satisfied and $|\boldsymbol{\eta}| \leq \pi_0$.

The coefficients of the NLS can be found as follows. Using the analytic expansion (11) of the solution of (3) we obtain the following representation for the modal coefficients

$$\tilde{U}_{\zeta,n}(\mathbf{k}, t) = \tilde{U}_{\zeta,n}^{(0)}(\mathbf{k}, t) + \alpha\tilde{U}_{\zeta,n}^{(1)}(\mathbf{k}, t) + O(\alpha^2), \ \ \frac{\tau_0}{\varrho} \leq t < \frac{\tau_*}{\varrho} \tag{60}$$

The first-order term of the power expansion (60) of $\tilde{U}_{\bar{n}}(\mathbf{k}, t)$ is given by the modal coefficient $\tilde{U}_{\bar{n}}^{(1)}(\mathbf{k}, t)$ of the FNLR determined by (14). We also have a similar expansion for $Z_\zeta(\mathbf{r}, t)$ and its Fourier transform $\hat{Z}_\zeta(\boldsymbol{\eta}, t)$:

$$\hat{Z}_\zeta(\boldsymbol{\eta}, t) = \hat{Z}_\zeta^{(0)}(\boldsymbol{\eta}, t) + \alpha \hat{Z}_\zeta^{(1)}(\boldsymbol{\eta}, t) + O(\alpha^2), \ \zeta = \pm \tag{61}$$

Then the coefficients to the NLS are determined from the following requirement. The FNLR of the NLS must approximate the FNLR of the NLM with an error O_{FNLR} so that *for all initial data h_ζ* the following two identities hold:

$$\hat{Z}_\zeta^{(0)}(\boldsymbol{\eta}, t) = \tilde{U}_{\zeta,n_0}^{(0)}(\zeta \mathbf{k}_* + \eta, t) + O\left(\frac{\beta^3}{\varrho}\right) \text{ if } |\boldsymbol{\eta}| \leq \pi_0, \ 0 \leq t < \frac{\tau_*}{\varrho} \tag{62}$$

$$\alpha \hat{Z}_\zeta^{(1)}(\boldsymbol{\eta}, t) = \alpha \tilde{U}_{\zeta,n_0}^{(1)}(\zeta \mathbf{k}_* + \boldsymbol{\eta}, t) + O_{\mathrm{FNLR}} \text{ if } |\boldsymbol{\eta}| \leq \pi_0, \ \frac{\tau_0}{\varrho} \leq t < \frac{\tau_*}{\varrho} \tag{63}$$

For instance, if the NLS is second-order (that is given by (36) and (37)) $O_{\mathrm{FNLR}} = O(\beta) + O(\varrho)$. If we use solutions of the fourth-order extended NLS that involve additional terms, the error becomes smaller, in (63) $O_{\mathrm{FNLR}} = O(\beta^3) + O(\beta\varrho)$ and in (62) $O(\beta^3/\varrho)$ is replaced by $O(\beta^5/\varrho)$.

Note that since we use in (58) the exact solution $Z_\pm$ of the NLS and not $Z_\zeta^{(0)} + \alpha Z_\zeta^{(1)}$, the approximation error is much smaller than one may conclude looking only at (62) and (63) (see section 7 for details).

All remaining, indirectly excited modes of the approximate solution are given in terms of the FNLR:

$$\tilde{U}_{Z,\zeta,n}^{\mathrm{ind}}(\mathbf{k}, t) = \tilde{U}_{\zeta,n}^{(1)}(\mathbf{k}, t) \tag{64}$$

and the approximation error

$$\tilde{U}_{\bar{n}}(\mathbf{k}, t) - \tilde{U}_{Z,\bar{n}}^{\mathrm{ind}}(\mathbf{k}, t) = O(\varrho\alpha) \text{ when } n \neq n_0 \text{ or } |\mathbf{k} - \mathbf{k}_*| > \pi_0 \tag{65}$$

Note that indirectly excited modes are much smaller than directly excited, i.e.

$$\tilde{U}_{Z,\zeta,n}^{\mathrm{ind}}(\mathbf{k}, t) = O(\varrho) \tag{66}$$

compared with $\tilde{U}_{Z,n_0}^{\mathrm{dir}} = O(1)$. The relative magnitude of the modes is also shown in table 6. Note that (66) implies that *the indirectly excited modes can be neglected in the cases (46) and (49) but have to be taken into account when higher precision approximation is used*. The analysis in section 7 shows that though we determine the coefficients of the NLS based on the FNLR of the NLM, using *exact* solution Z_ζ of the NLS in (58) allows us to obtain estimates

Table 6. Order of magnitude of excitation current $\mathbf{J}(\mathbf{r}, t)$ before it vanishes for $t \leq \tau_0/\varrho$, the field $\mathbf{U}(\mathbf{r}, t)$, which is an exact solution to the NLM, and its components during the time period $\tau_0/\varrho \leq t \leq \tau_*/\varrho$. Notice that the indirectly excited part of the FNLR is far smaller than the directly excited one.

Order of magnitude of fields and its components for $\tau_0/\varrho \leq t \leq \tau_*/\varrho$ under the classical NLS scaling $\varrho \sim \alpha \sim \beta^2$ in the one-dimensional case	
Excitation current $\mathbf{J}(\mathbf{r}, t)$ for $t \leq \tau_0/\varrho$	$\varrho \sim \beta^2$
Linear response $\mathbf{U}^{(0)}(\mathbf{r}, t)$	1
Directly excited part of the FNLR $\alpha\mathbf{U}^{(1)\mathrm{dir}}(\mathbf{r}, t)$,	$\alpha\varrho^{-1} \sim 1$
Indirectly excited part of the FNLR $\alpha\mathbf{U}^{(1)\mathrm{ind}}(\mathbf{r}, t)$	$\alpha \sim \beta^2$
Exact solution of the NLM $\mathbf{U}(\mathbf{r}, t)$	1

(59). The indirectly excited part of the approximate solution is given by the formula

$$\mathbf{U}_Z^{\mathrm{ind}}(\mathbf{r},t)=\sum_{n=1}^{\infty}\mathbf{U}_{Z,n}^{\mathrm{ind}}(\mathbf{r},t) \tag{67}$$

where

$$\mathbf{U}_{Z,n}^{\mathrm{ind}}(\mathbf{r},t)=\frac{\alpha}{(2\pi)^d}\int_{[-\pi,\pi]^d}$$

$$[\tilde{U}_{+,n}^{(1)}(\mathbf{k}_*+\eta,t)\tilde{\mathbf{G}}_{+,n}(\mathbf{r},\mathbf{k}_*+\boldsymbol{\eta})+\tilde{U}_{-,n}^{(1)}(-\mathbf{k}_*+\boldsymbol{\eta},t)\tilde{\mathbf{G}}_{-,n}(\mathbf{r},-\mathbf{k}_*+\boldsymbol{\eta})]\,\mathrm{d}\boldsymbol{\eta}\ n\neq n_0 \tag{68}$$

$$\mathbf{U}_{Z,n_0}^{\mathrm{ind}}(\mathbf{r},t)=\frac{\alpha}{(2\pi)^d}\int_{[-\pi,\pi]^d}(1-\Psi(\boldsymbol{\eta}))$$

$$[\tilde{U}_{+,n_0}^{(1)}(\mathbf{k}_*+\boldsymbol{\eta},t)\tilde{\mathbf{G}}_{+,n_0}(\mathbf{r},\mathbf{k}_*+\boldsymbol{\eta})+\tilde{U}_{-,n_0}^{(1)}(-\mathbf{k}_*+\boldsymbol{\eta},t)\tilde{\mathbf{G}}_{-,n_0}(\mathbf{r},-\mathbf{k}_*+\boldsymbol{\eta})]\mathrm{d}\boldsymbol{\eta} \tag{69}$$

with $\tilde{U}_{\bar{n}}^{(1)}(\mathbf{k},t)$ being the modal coefficient of the solution $\mathbf{U}^{(1)}(\mathbf{r},t)$ of the linear equation (14) (for an explicit formula see (175), (176) and (166)).

In conclusion, the developed method allows one to find higher order approximations of the solutions of the NLM by solutions of NLS-type equations with a rigorous control of errors on time intervals for which the relations (15) hold. If additional information on the solution of the NLS is available, in particular, if appropriate stability conditions are fulfilled, $\mathbf{U}_Z(\mathbf{r},t)$ approximates the exact solution $\mathbf{U}(\mathbf{r},t)$ well on longer time intervals (see section 7 for details).

1.3 *Wave interactions and multimode NLS regimes*

It is interesting and instructive to look at NLS regimes of nonlinear wave propagation in periodic dielectric media in the context of nonlinear interactions between the eigenmodes of the underlying linear medium. From that perspective an *NLS regime can be characterized as a regime of nonlinear mode interactions when, for a generic mode, its self-interaction (that is interaction with the conjugate mode) significantly dominates the nonlinear interactions with all other modes*. A more accurate description of an NLS regime is based on finer estimations of magnitudes of nonlinear interactions between different modes and their dependence on values of the small parameters α, ϱ and β. In turns out that, in the case of an NLS regime when a generic mode (described by a quasimomentum $\mathbf{k}_*$ and a band index $\bar{n}=(\zeta,n_0)$) is excited, it interacts significantly more strongly with modes from the same band n_0 and with quasimomenta located about $\mathbf{k}_*$ than with all other modes. In addition to that, nonlinear interactions between a mode $((\zeta,n_0),\zeta\mathbf{k}_*)$ and its conjugate mode $((-\zeta,n_0),-\zeta\mathbf{k}_*)$ are much stronger than other mode interactions in this band. We call such a modal pair, occurring often in our analysis, a *doublet* and denote it by

$$\uparrow n_0,\mathbf{k}_*\downarrow=\{(+,n_0,\mathbf{k}_*),\ (-,n_0,-\mathbf{k}_*)\}=\{(\zeta,n_0,\zeta\mathbf{k}_*):\zeta=\pm\} \tag{70}$$

In this article (excluding this subsection) we consider primarily almost single-mode current excitations based on a single doublet $\uparrow n_0,\mathbf{k}_*\downarrow$ formed by a mode $((+,n_0),\mathbf{k}_*)$ together with its conjugate counterpart $((-,n_0),-\mathbf{k}_*)$ that would allow one to produce a real-valued field. The dynamics of a doublet is described by the NLSs (36), (37) or with higher precision by the ENLS (47). A more detailed investigation of nonlinear mode interactions would naturally require the introduction of multimode current excitation involving small vicinities of several doublets $\uparrow n_l,\mathbf{k}_{*l}\downarrow, l=1,\dots,N$, rather than just an almost single-mode excitation and leading to groups of excited modes.

The analysis below suggests a view on the NLS and ENLS as regimes of nonlinear wave propagation when wave modal components admit a decomposition into essentially noninteracting groups. Consequently, the existence, conditions and accuracy of such a decomposition (as well as the derivation of relevant simplified evolution equations of smaller modal groups) become a subject of the theory of ENLSs. *In other words, a 'big picture' characterizing an NLS regime for the electromagnetic wave propagation is that the evolution of components of its modal composition occurs essentially independently for groups of modes with separated carrier frequencies and quasimomenta whereas the interactions inside every single group occur according to a rather universal scenario described by nonlinear Schrödinger-type equations.*

The first and fundamental step in the analysis is to find and classify all the interactions between the modal groups as well as with the rest of modes with estimations of their relative magnitudes. We do this based on the quantitative theory of nonlinear mode interactions and, in particular, with the help of selection rules for stronger interactions studied in [17–19]. The essentials of the analysis are provided below.

1.3.1 Selection rules for stronger wave interactions and NLS regimes. To find the wave decomposition into almost independent components we use the selection rules for stronger interactions [17–19], which are as follows. Consider the modal coefficients $\tilde{U}_{\bar{n}}(\mathbf{k}, t)$ of the wave governed by the NLM. Notice that if $\alpha = 0$ the NLM turns into a linear equation and, according to the classical spectral theory, the modal coefficients $\tilde{U}_{\bar{n}}(\mathbf{k}, t)$ for different $\bar{n}$ and $\mathbf{k}$ evolve independently one from another as in (54). For $\alpha \neq 0$ the cubic nonlinearity introduces interactions between all the modes. In the case when the nonlinear term of the electric polarization has the same spatial period as the underlying linear medium, the first fundamental restriction on any quadruplet of interacting Bloch modes is given by the *phase matching condition*

$$\mathbf{k} = \mathbf{k}' + \mathbf{k}'' + \mathbf{k}''' \bmod(2\pi) \tag{71}$$

where $\mathbf{k}$ is the quasimomentum of the mode which is affected by a triad of modes with the quasimomenta $\mathbf{k}', \mathbf{k}'', \mathbf{k}'''$. We call the triad

$$((\zeta', n'), \mathbf{k}'),\ ((\zeta'', n''), \mathbf{k}''), ((\zeta''', n'''), \mathbf{k}''') \tag{72}$$

the *origin triad* or *origin modes* of the interaction quadruplet and $((\zeta, n), \mathbf{k})$ the *end mode* of the quadruplet. The *interaction quadruplet* is completely defined by its origin triad and its end mode.

If in the excitation current $\mathbf{J}$ of the form (35), (55) or of a more general form described in [17–19], both the parameters α and ϱ are small, and (34) is fulfilled (or, more precisely, (76) holds) then stronger interacting modal quadruplets satisfy also the *group velocity matching* condition

$$\nabla\omega_{\bar{n}'}(\mathbf{k}') = \nabla\omega_{\bar{n}''}(\mathbf{k}'') = \nabla\omega_{\bar{n}'''}(\mathbf{k}''') \tag{73}$$

Note that (73) is a constraint only on the origin triad of the quadruplet. The selection rule (73) is the most important one, since if it is not fulfilled, the magnitude of the interaction is estimated by $O((\varrho/\beta)^{\kappa})$ with arbitrarily large κ, and, in view of (34), is not a strong interaction.

Finally, a modal quadruplet would have even stronger nonlinear interactions if, in addition to the phase and group velocity matching, it satisfies the *frequency matching* condition

$$\omega_{\bar{n}'}(\mathbf{k}') + \omega_{\bar{n}''}(\mathbf{k}'') + \omega_{\bar{n}'''}(\mathbf{k}''') = \omega_{\bar{n}}(\mathbf{k}). \tag{74}$$

For many cases of interest there are modal quadruplets satisfying all three conditions of (71), (73) and (74), [19]. In any case, the selection rules (71), (73) and (74) determine stronger interacting quadruplets of modes with a detailed classification of generic mode interactions provided in [17–19].

1.3.2 Multiple mode excitations and waves. Multimode excitations can be introduced as follows. First we introduce the excitation current as a linear superposition of currents of the form (55), namely

$$\mathbf{J} = \sum_{l=1}^{N} \mathbf{J}_l \tag{75}$$

with every $\mathbf{J}_l$ being an almost single-mode excitation given by (35) and (55) with corresponding $\mathbf{k}_{*l}$ and n_{0l}, $l = 1, \ldots, N$. Consequently, every $\mathbf{J}_l$ excites the corresponding doublet $\uparrow n_l, \mathbf{k}_{*l}\downarrow$. The modal components corresponding to the group B_l of modes with $|\mathbf{k} - \mathbf{k}_{*l}| \precsim \beta$ are directly excited through the linear process, and the amplitudes of the directly excited modes are considerably higher (of the order $O(\alpha^{-1})$ times) than the same for the indirectly excited modes.

We assume in this subsection that the ratio ϱ/β satisfies the condition (34), or, more precisely, that

$$\frac{\varrho}{\beta} \ll \max_{l=1,\ldots,N} \|\omega'_{n_0}(\mathbf{k}_{*l})\|, \;\; N \geq 2 \tag{76}$$

The condition (76) evidently requires the group velocities to be much larger than ϱ/β (this condition is not required in the single-mode case $N = 1$).

To determine finer features of the wave dynamics we pose the following questions.

- Which modes are excited through nonlinear interactions and what are the magnitudes of the amplitudes of such modes?
- Which interactions determine the dynamics of the directly excited modes with a given precision?
- What are the equations that determine the dynamics of the directly excited modes?
- What is the influence of indirectly excited modes on the directly excited modes?

The answers to the above questions depend on the choice of the quasimomenta $\mathbf{k}_{*l}$, $l = 1, \ldots, N$. It turns out that there are special combinations of modes having the strongest interactions and playing the dominant role for nonlinear wave evolution. Such special combinations involve exactly two ($N = 2$) special pairs of modes corresponding to the two values of $\vartheta = \pm 1$ for a two doublets $\uparrow n_0, \mathbf{k}_*\downarrow_\vartheta = \uparrow n_0, \vartheta\mathbf{k}_*\downarrow$. The values $\pm\omega_n(\mathbf{k}_*)$ of the carrier frequencies of the excitation and the quasimomenta $\pm\mathbf{k}_*$ are the same for both doublets. The difference between the doublets is in the value of the group velocity $\vartheta\omega'_n(\mathbf{k}_*)$ which is opposite for alternate doublets with $\vartheta = \pm 1$. Such an excitation and the corresponding wave can be interpreted as *bidirectional* (see also sections 1.4.6 and 5.4). In the case of a bidirectional excitation the wave evolution can be approximated by a four-component system of NLS equations which reduces to a two-component system (128) for the real-valued fields. Note that relevant interactions between the four modes of the bidirectional quadruplet are determined by the selection rules.

1.3.3 Mode-to-mode coupling and almost independence. It turns out that properly defined different types of mode combinations evolve almost independently for long times and high accuracy. In this section we introduce concepts and give a sketch of constructions needed for establishing that almost independence and more.

As in previous subsection we consider the current $\mathbf{J}=\sum_{l=1}^{N}\mathbf{J}_l$ with currents $\mathbf{J}_l$ described there, and denote by B_l of a set of directly excited modes by the current $\mathbf{J}_l$.

$$B_l=\{(\zeta,n,\mathbf{k}):n=n_{0l},\ |\mathbf{k}-\zeta\mathbf{k}_{*l}|\le\pi_0,\ \zeta=+\text{ or }\zeta=-\},\ l=1,\dots,N \tag{77}$$

We consider also: (i) the complement B_l^{C} for every B_l; (ii) B as the union of all B_l; (iii) the complement B^{C}, i.e.

$$B=\bigcup_{l=1,\dots,N}B_l,\ B^{\mathrm{C}}=\left(\bigcup_{l=1,\dots,N}B_l\right)^{\mathrm{C}} \tag{78}$$

Then we introduce a decomposition of the wave $\mathbf{U}$ governed by the NLM based on B_l, namely

$$\mathbf{U}=\mathbf{U}_{B_1}+\ldots+\mathbf{U}_{B_N}+\mathbf{U}_{B^{\mathrm{C}}} \tag{79}$$

where $\mathbf{U}_{B_l}$ is composed of modes from B_l. Using such a decomposition we recast NLM (3) in the equivalent form of the following system of equations:

$$\partial_t\mathbf{U}_{B_1}=-\mathrm{i}\mathbf{M}\mathbf{U}_{B_1}+\alpha\mathcal{F}_{\mathrm{NL}}(\mathbf{U}_{B_1}+\cdots+\mathbf{U}_{B_N}+\mathbf{U}_{B^{\mathrm{C}}})\,|_{B_1}-\mathbf{J}_1 \tag{80}$$

$$\cdots$$

$$\partial_t\mathbf{U}_{B_N}=-\mathrm{i}\mathbf{M}\mathbf{U}_{B_N}+\alpha\mathcal{F}_{\mathrm{NL}}(\mathbf{U}_{B_1}+\cdots+\mathbf{U}_{B_N}+\mathbf{U}_{B^{\mathrm{C}}})\,|_{B_N}-\mathbf{J}_N \tag{81}$$

$$\partial_t\mathbf{U}_{B^{\mathrm{C}}}=-\mathrm{i}\mathbf{M}\mathbf{U}_{B^{\mathrm{C}}}+\alpha\mathcal{F}_{\mathrm{NL}}(\mathbf{U}_{B_1}+\cdots+\mathbf{U}_{B_N}+\mathbf{U}_{B^{\mathrm{C}}})\,|_{B^{\mathrm{C}}} \tag{82}$$

$$\mathbf{U}_{B_1}=0,\dots,\mathbf{U}_{B_N}=0;\quad \mathbf{U}_{B^{\mathrm{C}}}=0\text{ for }t\le 0 \tag{83}$$

Now we define the almost independence of the components $\mathbf{U}_{B_l}$ for different l, $l=1,\dots,N$.

If $\mathbf{U}_{B_1},\dots,\mathbf{U}_{B_N}$ solving the system (80)–(83) were independent then we would be able to drop in the right-hand side of the lth equation in (80)–(83) everything but the corresponding $\mathbf{U}_{B_l}$, and would get the following decoupled system:

$$\partial_t\mathbf{V}_{B_1}=-\mathrm{i}\mathbf{M}\mathbf{V}_{B_1}+\alpha\mathcal{F}_{\mathrm{NL}}(\mathbf{V}_{B_1})\,|_{B_1}-\mathbf{J}_1;\ \mathbf{V}_{B_1}=0\text{ for }t\le 0 \tag{84}$$

$$\cdots$$

$$\partial_t\mathbf{V}_{B_N}=-\mathrm{i}\mathbf{M}\mathbf{V}_{B_N}+\alpha\mathcal{F}_{\mathrm{NL}}(\mathbf{V}_{B_N})\,|_{B_N}-\mathbf{J}_N;\ \mathbf{V}_{B_N}=0\text{ for }t\le 0 \tag{85}$$

$$\partial_t\mathbf{V}_{B^{\mathrm{C}}}=-\mathrm{i}\mathbf{M}\mathbf{V}_{B^{\mathrm{C}}}+\alpha\mathcal{F}_{\mathrm{NL}}(\mathbf{V}_{B_1}+\cdots+\mathbf{V}_{B_N})\,|_{B^{\mathrm{C}}};\mathbf{V}_{B^{\mathrm{C}}}=0\text{ for }t\le 0 \tag{86}$$

In other words, the lth equation in (84)–(85) for $\mathbf{V}_{B_l}$ is obtained from the lth equation for $\mathbf{U}_{B_l}$ by dropping the $\mathbf{U}_{B_j}$, $j\neq l$ and $\mathbf{U}_{B^{\mathrm{C}}}$ in the nonlinear term. Obviously, the first N equations in (84)–(86) can be solved independently, and the very last equation (86) is linear with respect to $\mathbf{V}_{B^{\mathrm{C}}}$ and can be easily solved too.

To find the nonlinear influence of modes from B_l onto themselves we take the lth equation in (84)–(85) and set $\mathbf{V}_{B_l}$ in the nonlinear term $\mathcal{F}_{\mathrm{NL}}$ to be zero that leads to the following linear equation

$$\partial_t\mathbf{V}_{B_l}^{(0)}=-\mathrm{i}\mathbf{M}\mathbf{V}_{B_l}^{(0)}-\mathbf{J}_l;\ \mathbf{V}_{B_l}^{(0)}=0\text{ for }t\le 0 \tag{87}$$

Now we can assess the level of independence or coupling of different $\mathbf{U}_{B_l}$ and $\mathbf{U}_{B^{\mathrm{C}}}$ by comparing them with the corresponding $\mathbf{V}_{B_l}$ and $\mathbf{V}_{B^{\mathrm{C}}}$, and similarly we can compare $\mathbf{V}_{B_l}$ with $\mathbf{V}_{B_l}^{(0)}$ to assess the nonlinear influence of modes B_l onto themselves. Namely, we define the *mode-to-mode coupling* as follows

$$\begin{aligned}\text{mode-to-mode coupling}\quad & B_l^{\mathrm{C}}\to B_l\equiv\mathbf{U}_{B_l}-\mathbf{V}_{B_l}\\ \text{nonlinear mode-to-mode coupling}\quad & B_l\to B_l\equiv\mathbf{V}_{B_l}-\mathbf{V}_{B_l}^{(0)}\end{aligned} \tag{88}$$

Table 7. The magnitude of nonlinear mode-to-mode coupling for unidirectional excitations in the one-dimensional case under classical NLS scaling $\varrho \sim \alpha \sim \beta^2$.

Mode-to-mode coupling	Order of the mode-to-mode coupling for $\tau_0/\varrho \le t \le \tau_*/\varrho$
$\text{doublet}_l \to \text{doublet}_l$	$\alpha/\varrho \sim 1$
$\text{doublet}_l \to \text{doublet}_l^{\mathbf{C}}$	$\varrho \sim \beta^2$
$\text{doublet}_l^{\mathbf{C}} \to \text{doublet}_l$	β^∞

We would like to emphasize that *the definition of mode-to-mode coupling includes the direction of influence* via the corresponding evolution equations (80)–(83) and (84)–(86), and *mode-to-mode coupling is not symmetric*. The analysis of nonlinear evolution requires one to introduce such a direction of influence for nonlinearly interacting modes.

An additional analysis of the equations (84)–(86) also shows that $\mathbf{V}_{B_l}$ can be well approximated by a solution of a corresponding NLS or ENLS systems. In addition to that, estimates similar to (65) for indirectly excited modes $\mathbf{U}_{B^{\mathrm{C}}}$ in one-dimensional case $d = 1$ with the classical NLS scaling $\varrho \sim \alpha \sim \beta^2$ yield that

$$\mathbf{U}_{B^{\mathrm{C}}} - \mathbf{V}_{B^{\mathrm{C}}} = O(\alpha^2) = O(\beta^4) \tag{89}$$

Table 7 shows order of magnitude estimates of the mode-to-mode interactions involving unidirectional excitations and doublets (β^∞ in this table means arbitrarily large power of β).

When the excitations are bidirectional, quadruplets of modes are excited, and in this case magnitudes of nonlinear interactions are as in table 8, which is similar to table 7.

The order of magnitude comparative estimates for the basic system (80)–(83) and its decoupled counterpart (84)–(86) provide additional facts on the interplay between dispersion and nonlinearity. These estimates are collected in table 9, and they are based on the analysis of the exact solution $\mathbf{U}(t)$ of the NLM involving instrumentally: (i) the analytic expansion (11) for $\mathbf{U}(t)$; (ii) representation of the terms of that expansion (11) by oscillatory integrals; (iii) computation of asymptotic approximations and series for these oscillatory integrals as powers of the small parameters α, ϱ and β. Observe that for a generic $\mathbf{W}$, which can be expanded as in (79), the value $\alpha\mathcal{F}_{\mathrm{NL}}(\mathbf{W})|_{B_l}$ differs noticeably from $\alpha\mathcal{F}_{\mathrm{NL}}(\mathbf{W}_{B_l})|_{B_l}$ and the difference is of order β^2. In contrast, in the case when $\mathbf{W}$ is the exact solution $\mathbf{U}(t)$ of the NLM the same difference for $\tau_0/\varrho \le t \le \tau_*/\varrho$ is of order β^∞, which is much smaller. This effect is due to destructive wave interference and wave dispersion for a wave governed exactly by the NLM.

We end this subsection by the following qualitative conclusions on the interplay between dispersive and nonlinear effects:

- dispersive effects balance nonlinear effects when modes interact inside one doublet leading to NLS/ENLS-type dynamics;
- dispersive effects are dominant in interactions between different doublets, and nonlinear effects are less pronounced.

Table 8. Magnitudes of nonlinear mode-to-mode coupling for bidirectional excitations in the one-dimensional case under the classical NLS scaling $\varrho \sim \alpha \sim \beta^2$.

Mode-to-mode coupling	Order of the mode-to-mode coupling for $\tau_0/\varrho \le t \le \tau_*/\varrho$
$\text{quadruplet}_l \to \text{quadruplet}_l$	$\alpha/\varrho \sim 1$
$\text{quadruplet}_l \to \text{quadruplet}_l^{\mathbf{C}}$	$\alpha \sim \beta^2$
$\text{quadruplet}_l^{\mathbf{C}} \to \text{quadruplet}_l$	β^∞

Table 9. The order of magnitude estimates collected here are based on the analysis of the exact solution $\mathbf{U}(t)$ of the NLM involving instrumentally: (i) the analytic expansion (11) for $\mathbf{U}(t)$; (ii) representation of the terms of that expansion (11) by oscillatory integrals; (iii) computation of asymptotic approximations and series for these oscillatory integrals as powers of the small parameters α, ϱ and β.

Comparison of solutions to the basic system and its decoupled counterpart under classical NLS scaling $\varrho \sim \alpha \sim \beta^2$ for $\tau_0/\varrho \leq t \leq \tau_*/\varrho$	
Solutions $\mathbf{U}_{B_l}, \mathbf{V}_{B_l}$	1
Nonlinearity $\alpha\mathcal{F}_{\mathrm{NL}}(\mathbf{U})\mid_{B_l}, \alpha\mathcal{F}_{\mathrm{NL}}(\mathbf{V}_{B_l})\mid_{B_l}$	α
Difference of the values of nonlinearity on generic test functions $\alpha\mathcal{F}_{\mathrm{NL}}(\mathbf{W}_{B_1}+\cdots+\mathbf{W}_{B_N}+\mathbf{W}_{B^C})\mid_{B_l} -\alpha\mathcal{F}_{\mathrm{NL}}(\mathbf{W}_{B_l})\mid_{B_l}$	$\alpha \sim \beta^2$
Difference of the values of nonlinearity applied to solutions $\alpha\mathcal{F}_{\mathrm{NL}}(\mathbf{U}_{B_1}(t)+\cdots+\mathbf{U}_{B_N}(t)+\mathbf{U}_{B^C}(t))\mid_{B_l} -\alpha\mathcal{F}_{\mathrm{NL}}(\mathbf{U}_{B_l}(t))\mid_{B_l}$	β^∞
Difference of solutions $\mathbf{U}_{B_l}(t)-\mathbf{V}_{B_l}(t)$	β^∞

1.3.4 Spectral theory of nonlinear wave propagation. The above discussion suggests that the theory of NLSs, ENLSs and systems of coupled ENLSs can be viewed as the *spectral theory of nonlinear wave propagation*. The word spectral here refers to the property of certain classes of waves to be decomposable into components evolving *almost independently* for long times as described in the previous section. The almost independence, in turn, means that the coupling between the components is small, and, more precisely, that the coupling terms in the relevant exact evolution equations can be classified by powers $\alpha^{l_0}\varrho^{l_1}\beta^{l_2}$. We recall that the small parameters α, ϱ and β introduced in previous sections characterize, respectively, the relative magnitude of nonlinearity, the time and the space scales related to the nonlinear evolutions. The parameter α characterizes the magnitude of the nonlinearity. The next is the small parameter ϱ, which characterizes the degree of time-harmonicity of the excitation wave. And, finally, the third small parameter β characterizes the linear dimensions of a small vicinity of a single or several quasimomenta $\mathbf{k}_{*j}$ involved in the modal decomposition of the wave. *When accounting for different magnitudes of mode interactions as powers $\alpha^{l_0}\varrho^{l_1}\beta^{l_2}$ we come to either the classical NLS, ENLS or a system of ENLSs*. The so obtained equations take into account at the prescribed precision level all relevant nonlinear interactions and with that level of accuracy describe the nonlinear wave evolution. In such a construction, linear spectral theory forms a fundamental basis for the nonlinear one. It yields a system of eigenmodes that evolve independently and set a framework for nonlinear spectral theory.

In this article we focus primarily at almost single-mode excitation currents and only sketch the case of multimode excitations. More detailed studies of waves generated by multimode excitation currents and, in particular, the derivation of the corresponding systems of ENLSs accounting for smaller coupling between essentially noninteracting groups of modes are naturally to be conducted as the next step.

The essence of the above discussion on nonlinear evolution and wave interactions can be formulated in the form of the following *principle of approximate superposition*. Let us call a solution to the NLM *a multiple-mode solution* if it corresponds to an excitation current which is generic and is a sum of almost time-harmonic single-mode excitations. *Then being given a level of accuracy and any multiple-mode solution we can decompose it into the sum of certain single-mode solutions each of which is governed by the NLS or ENLS (can be a system) with a prescribed accuracy.*

More accurate formulation of the principle of approximate superposition is as follows. Let $\mathbf{U}_l$ be a solution of the NLM corresponding to an almost single-mode excitation $\mathbf{J}_l$ around $\mathbf{k}_{*l}$, i.e.

$$\partial_t \mathbf{U}_l = -\mathrm{i}\mathbf{M}\mathbf{U}_l + \alpha\mathcal{F}_{\mathrm{NL}}(\mathbf{U}_l) - \mathbf{J}_l, \; l = 1, \ldots, N \tag{90}$$

Then for a generic collection of $\mathbf{k}_{*l}$ the multiple-mode solution $\mathbf{U}$ corresponding to sum of $\mathbf{J}_l$ satisfies

$$\partial_t \mathbf{U} = -\mathrm{i}\mathbf{M}\mathbf{U} + \alpha\mathcal{F}_{\mathrm{NL}}(\mathbf{U}) - \mathbf{J},\ J = \sum_{l=1}^{N} \mathbf{J}_l \tag{91}$$

and on the time interval (15)

$$\mathbf{U} = \sum_{l=1}^{N} \mathbf{U}_l + O\left(\frac{\varrho}{\beta}\right)^{N_1},\ \text{where } N_1 \text{ can be arbitrarily large} \tag{92}$$

Observe a remarkable *superaccuracy* of the superposition formula in (92). For the typical scaling $\varrho \sim \alpha \sim \beta^{\varkappa_1}$, $\varkappa_1 \geq 2$, as in (29), the approximation error is smaller than any power of α whereas the nonlinearity itself is of order α. The explanation of the superaccuracy follows from an analysis of nonlinear wave interactions.

Notice that the principle of approximate superposition has its natural limitations, and the condition $(n_{l_1}, \mathbf{k}_{*l_1}) \neq (n_{l_1}, \mathbf{k}_{*l_2})$ for $l_1 \neq l_2$ in (75) is absolutely instrumental. For instance, though evidently $2\mathbf{J}_l = \mathbf{J}_l + \mathbf{J}_l$, the solution for $2\mathbf{J}_l$ is evidently not $2\mathbf{U}_l = \mathbf{U}_l + \mathbf{U}_l$ since $\mathbf{U}_l$ is a solution of a nonlinear equation which is well-approximated by the NLS. So, to have (92) with arbitrary large N_1 the proper genericity condition has to include

$$\nabla\omega_{\bar{n}_{l_2}}(\mathbf{k}_{*l_2}) \neq \nabla\omega_{\bar{n}_{l_1}}(\mathbf{k}_{*l_1}) \text{ for } l_2 \neq l_1 \tag{93}$$

1.4 *Extended nonlinear Schrödinger equations*

As discussed above, the NLS describes the evolution of a doublet of directly excited modes of the NLM. More accurate higher order approximations of solutions to the NLM can be obtained by constructing an extended NLS (ENLS) instead of the classical NLS. Extended NLSs are widely used in nonlinear optics (see [23–26]). The corrective terms in the ENLS originate from several sources, resulting in relatively smaller alterations of solutions compared to the basic (classical) NLS. Note that the nonlinearity $\alpha|Z|^2 Z$ in the classical NLS gives an alteration of the linear Schrödinger equation of order $O(\alpha)$ and the error of approximation (when time t is $O(1)$) of the NLM by a linear Schrödinger equation is $O(\alpha)$ too. One though has to take into account that an $O(\alpha)$ alteration of the equation leads on time intervals of length $O(\varrho^{-1})$, which we consider here, to alterations of solutions of order $O(\alpha\varrho^{-1})$. In the case of classical NLS scaling (26) $O(\alpha) = O(\varrho) = O(\beta^2)$ and solutions of the classical NLS give approximations of solutions of the NLM with the error $O(\beta)$. We consider in this article two types of the ENLS: third and fourth order. Using a third-order ENLS improves the error estimate in the case of classical NLS scaling (26) from $O(\beta)$ to $O(\beta^2)$ and solutions of fourth-order ENLSs approximate solutions of the NLM with accuracy $O(\beta^3)$.

Here is a complete list of all sources of the additional corrective terms that are required to be added to the NLS to improve the accuracy of approximation with estimations of their magnitude:

- cubic and the fourth-order polynomial approximations of the dispersion relation $\omega_{n_0}(\mathbf{k})$ at $\mathbf{k}_*$ with the corrective terms magnitude $O(\beta^3)$ and $O(\beta^4)$, respectively;
- the first-order approximation of the cubic susceptibility accounting for its frequency dependence (see section 6.3) with the corrective terms magnitude $O(\alpha\varrho)$;
- polynomial approximation of the modal susceptibility $\check{Q}_{\bar{n}}(\vec{k})$ in (179) at $\vec{k}_*$ with the corrective terms magnitude $O(\alpha\beta)$ or $O(\alpha\beta^2)$;

- non-frequency-matched interactions between waves propagating in opposite directions (see section 5.4) with the correction term magnitude $O(\alpha\varrho)$;
- fifth-order nonlinear terms in the expansion of the nonlinearity in the NLM with the corrective terms magnitude $O(\alpha^2)$;
- the interband interaction terms with the corrective terms magnitude $O(\alpha\varrho)$.

If all the corrective terms from the above list are taken into account then the accuracy of the approximation by the fourth-order ENLS of the NLM is estimated by:

$$\mathbf{U} - \mathbf{U}_Z = O(\beta^3) + O(\beta\varrho) + O(\varrho^2) \tag{94}$$

In the case of classical NLS scaling $\varrho \sim \alpha \sim \beta^2$ and a time interval length $O(\varrho^{-1})$, we find that the neglected terms in the right-hand side of (94) correspond to alterations of the equations of order $O(\varrho\beta^3) = O(\beta^5)$, whereas the classical NLS nonlinearity itself is of order β^2 and the introduced above additional corrective terms in ENLS are of order β^3 or β^4.

Let us introduce the following short notation for the linear Schrödinger operator of the order ν:

$$\mathcal{L}_+^{[\nu]} Z = \gamma_{(\nu)}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]Z,\ \mathcal{L}_-^{[\nu]} Z = -\gamma_{(\nu)}[\mathrm{i}\vec{\nabla}_{\mathbf{r}}]Z \tag{95}$$

where the so-called symbol (characteristic polynomial) $\gamma_{(\nu)}(\boldsymbol{\eta})$ of the differential operator $\gamma_{(\nu)}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]$ is the Taylor polynomial of order ν of the dispersion relation $\omega_{n_0}(\mathbf{k})$ at $\mathbf{k} = \mathbf{k}_*$. For instance, for $\nu = 2$

$$\gamma_{(2)}(\boldsymbol{\eta}) = \omega_{n_0}(\mathbf{k}_*) + \omega'_{n_0}(\mathbf{k}_*)(\eta) + \frac{1}{2}\omega''_{n_0}(\mathbf{k}_*)(\boldsymbol{\eta}^2) \tag{96}$$

We always consider the situation where the FNLR is applicable, that is the time interval satisfies (15), i.e.

$$\frac{\tau_0}{\varrho} \le t \le \frac{\tau_*}{\varrho} \text{ where } \frac{\tau_0}{\tau_*} < 1 \text{ isfixed} \tag{97}$$

We consider the cases $\nu = 2$, $\nu = 3$ and $\nu = 4$. The resulting ENLSs and the approximation error estimates are the same in both the dispersive and weakly dispersive cases (28) and (30). In the error estimates we assume that (29) or (31) holds with some fixed value of $\varkappa_1$. Often from general error estimates that include the three parameters α, β, ϱ we deduce in the case of the classical NLS a scaling (26) simpler estimates in terms of a single parameter β as a consequence.

For illustration we give the form of a typical ENLS of order ν (for simplicity skipping some corrective terms)

$$\partial_t Z_+ = -\mathrm{i}\mathcal{L}_+^{[\nu]} Z_+ + \alpha_\pi p_+^{[\nu-2]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}](Z_+^2 Z_-),\ \alpha_\pi = 3\alpha(2\pi)^{2d} \tag{98}$$

$$\partial_t Z_- = -\mathrm{i}\mathcal{L}_-^{[\nu]} Z_- + \alpha_\pi p_-^{[\nu-2]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}](Z_-^2 Z_+) \tag{99}$$

where the nonlinearity $p_+^{[\nu-2]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}](Z_+^2 Z_-)$ includes spatial derivatives of Z_+ and Z_- of order up to $\nu - 2$.

Note that if (i) the excitation currents are real-valued, (ii) equation (6) holds and (iii) the polarization tensors have real coefficients, then the equation for Z_+ is obtained by the complex conjugation of the equation for Z_- and $Z_- = Z_+^*$. Moreover, we can use (38) and reduce the system for two equations to one equation for Z_+.

The ENLSs which we describe below are *universal*; they do not depend on the a relation between ϱ, α and β, in particular on the exponents $\varkappa_0$ in (16) and $\varkappa_1$ in (29) or (31). From the universal ENLS one may deduce a reduced ENLS for a particular scaling (see section 1.4.7 for examples of such reduction). The reduced ENLS may depend on the choice of $\varkappa_0$ and $\varkappa_1$.

Remark When we discuss the magnitude of the terms in the NLS, one has to take into account that we study the NLM and NLS on intervals of time t of order $1/\varrho$. Integration of the equation with respect to t leads to a factor $1/\varrho$ in the contribution of the corresponding terms to the exact solution of the NLS. This also follows from the fact that the change of variables $t = \tau/\varrho$ (which introduces slow time τ of order one) leads to multiplication of the right-hand sides of (98) and (99) by $1/\varrho$. For example, in the case of the classical NLS scaling addition of the NLS nonlinearity which has order $O(\alpha)$ leads at times $t \sim 1/\varrho$ to a change of a solution of the linear equation of order $O(\alpha/\varrho)$ which is of order one as can be seen from (22). In our error estimates, for example (94), the effects of integration are already taken into account. Therefore, variation of solutions is $1/\varrho$ times greater than the corresponding variation of coefficients. In this section we will systematically use this correspondence without further reference; we give some details only when it is necessary, as in (129). To simplify the discussion of the magnitude of the nonlinear terms we everywhere in this section assume that (17) holds, namely $\alpha \sim \varrho$. Sometimes, for a further simplification, we consider classical NLS scaling, that is $\alpha \sim \varrho$, $\varrho \sim \beta^2$. ♦

1.4.1 The second-order ENLS.

If the order of the linear part $\nu = 2$ then the ENLSs take the form

$$\partial_t Z_+ = -\mathrm{i}\mathcal{L}_+^{[2]} Z_+ + \alpha_\pi p_+^{[0]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}](Z_+^2 Z_-), \ \alpha_\pi = 3\alpha(2\pi)^{2d} \tag{100}$$

$$\partial_t Z_- = -\mathrm{i}\mathcal{L}_-^{[2]} Z_- + \alpha_\pi p_-^{[0]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}](Z_-^2 Z_+) \tag{101}$$

with the initial conditions

$$Z_+(\mathbf{r}, t)|_{t=0} = h_+(\beta\mathbf{r}), \ Z_-(\mathbf{r}, t)\ |_{t=0} = h_-(\beta\mathbf{r}) \tag{102}$$

where

$$h_-(\beta\mathbf{r}) = h_+^*(\beta\mathbf{r}) \tag{103}$$

The linear operator

$$\mathcal{L}_\zeta^{[\nu]} = \zeta\gamma_{(\nu)}[-\mathrm{i}\zeta\vec{\nabla}_{\mathbf{r}}], \ \nu = 2, \ \zeta = \pm \tag{104}$$

is the second-order linear differential operator with constant coefficients given by formulas involving $\omega_{n_0}(\mathbf{k})$ and its derivatives at $\mathbf{k}_*$ which are similar to (20) (see equation (297) for general case and details). The action of $p_\pm^{[0]} = p_\pm^{[\sigma]}$ with $\sigma = 0$ is just the multiplication by a constant, that is

$$p_+^{[0]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}](Z_+^2 Z_-) = Q_\pm Z_+^2 Z_- \tag{105}$$

where $Q_\pm$ is determined by the modal susceptibility (179) (see equation (264) for details) and in this case we obtain the classical NLS (36). The order of approximation is given by the formula

$$\mathbf{U} - \mathbf{U}_Z = O(\beta) + O(\varrho) \tag{106}$$

Remark When together with fulfilment of (103) and (6) the nonlinearity in NLM is real-valued for real-valued vector fields, we have

$$Q_- = Q_+^*, \ Z_-(\mathbf{r}, t) = Z_+^*(\mathbf{r}, t) \tag{107}$$

Therefore we can use (38) and (100), and (101) is equivalent to the NLS in its classical form:

$$\partial_t Z_+ = -\mathrm{i}\mathcal{L}_+^{[2]} Z_+ + \alpha_\pi Q_+ |Z_+|^2 Z_+ \tag{108}$$

1.4.2 The third-order ENLS. For the classical NLS scaling (26) the leading term in the error estimate (106) is $O(\beta)$ compared with $O(\varrho) = O(\beta^2)$. To reduce $O(\beta)$ to $O(\beta^2)$ and thus get a higher accuracy of approximation we take $\nu = 3$ and obtain the following system of two third-order equations similar to (100), (101):

$$\partial_t Z_\zeta + \mathrm{i}\mathcal{L}_\zeta^{[3]} Z_\zeta = \alpha_\pi p_\zeta^{[1]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]\big(Z_\zeta^2 Z_{-\zeta}\big),\ \zeta = \pm \tag{109}$$

with the initial conditions (102) (see also [22, pp. 44–45, 27–34], where similar equations are studied). Now $\gamma_{(3)}(-\mathrm{i}\partial_1)$ in (104) is the third-order linear operator with the symbol which is the third-degree Taylor polynomial of the dispersion relation $\omega_{n_0}(\mathbf{k})$ at $\mathbf{k}_*$. For $\sigma = 1$ the polynomial

$$p_\pm^{[1]}(\boldsymbol{\eta}) = p_\pm^{[0]}(\boldsymbol{\eta}) + p_{1,\pm}(\boldsymbol{\eta}) \tag{110}$$

originates from the Taylor approximation of order one for the modal susceptibility $\check{Q}_{\vec{n}}(\vec{k})$ in (179) at the point $\vec{k}_*$ determined by $\mathbf{k}_*$. The zero-order term is given by (105). The action of $p_{1,\pm}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]$ on the product $Z_+^2 Z_-$ is defined by the formula

$$\begin{aligned} p_{1,+}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}](Z_+^2 Z_-) = {} & -\mathrm{i}Z_+ Z_-(a_{11,+} + a_{12,+}) \cdot \nabla_{\mathbf{r}}(Z_+) - \mathrm{i}Z_+^2 a_{13,+} \cdot \nabla_{\mathbf{r}}(Z_-) \\ & + Q_+ Z_+^2 Z_- \end{aligned} \tag{111}$$

where $a_{11,+}$, $a_{12,+}$ and $a_{13,+}$ are constant vectors explicitly given in terms of the gradient of $\check{Q}_{\vec{n}}$ at $\vec{k}_*$ by formula (267) which also defines $p_-^{[1]}[-\mathrm{i}\nabla_{\mathbf{r}}]$. Note that the order of the factors Z_+ and Z_- in the notation $p_+^{[1]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}](Z_+^2 Z_-)$ is important, see (435). The corrective terms $\alpha_\pi p_+^{[1]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}](Z_+^2 Z_-)$ can be considered as nonlinear corrections to the linear operator $\gamma_{(3)}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]$. Note that in the case of real-valued fields using (38) the first-order part of $p_+^{[1]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}](Z_+^2 Z_-)$ can be rewritten in the following commonly used form (see [22] pp. 44–45):

$$\begin{aligned} & -\mathrm{i}Z_+ Z_-(a_{11,+} + a_{12,+}) \cdot \vec{\nabla}_{\mathbf{r}}(Z_+) - \mathrm{i}Z_+^2 a_{13,+} \cdot \vec{\nabla}_{\mathbf{r}}(Z_-) \\ & = -\mathrm{i}|Z_+|^2(a_{11,+} + a_{12,+} - a_{13,+}) \cdot \vec{\nabla}_{\mathbf{r}}(Z_+) - \mathrm{i}Z_+ a_{13,+} \cdot \vec{\nabla}_{\mathbf{r}}(|Z_+|^2) \end{aligned}$$

The error of approximation in the case $\nu = 3$ is

$$\mathbf{U} - \mathbf{U}_Z = O(\beta^2) + O(\varrho) \tag{112}$$

The improvement $O(\beta^2)$ in (112) compared with $O(\beta)$ in (106) is obtained by a taking the variability of $\check{Q}_{\vec{n}}(\vec{k})$ into account and by a more precise approximation of $\omega_{n_0}(\mathbf{k})$.

In particular, for the classical NLS scaling (26) the error in (112) is $O(\beta^2)$. According to (102)

$$\vec{\nabla}_{\mathbf{r}} Z_+ = O(\beta),\ \vec{\nabla}_{\mathbf{r}}^2 Z_\zeta = O(\beta^2),\ \vec{\nabla}_{\mathbf{r}}^3 Z_+ = O(\beta^3) \tag{113}$$

Therefore the third-order terms added in $\mathcal{L}_+^{[3]}$ and the first-order terms added in (111) to the ENLS yield corrections of solutions of order $\varrho^{-1}O(\beta^3) = O(\beta)$ and $\alpha\varrho^{-1}O(\beta) = O(\beta)$; they are generically nonzero and much larger than the approximation error $O(\beta^2)$. Note that if the terms of order $O(\beta^3)$ in (109) were thrown away, we would arrive at the classical second-order NLS. Therefore, the difference of solutions of the NLM and the classical NLS really is of order $O(\beta)$ and is represented by the additional terms in the ENLS (109). Hence, *the corrections introduced into the NLSs capture the actual properties of solutions of the NLM and they are necessary if one wants to approximate the solutions to the NLM with a higher accuracy than the classical NLS.*

1.4.3 The fourth-order ENLS. From the very form of the error estimate (112), one can see that to improve the error estimate we have to make smaller each one of the two terms $O(\beta^2)$ and $O(\varrho)$, which are of the same order under the classical NLS scaling $\varrho \sim \alpha \sim \beta^2$ (and may be of different magnitude in a more general situation). To reduce $O(\beta^2)$ we have to approximate better the modal susceptibility $\breve{Q}_{\vec{n}}(\vec{k})$ about $\mathbf{k} = \mathbf{k}_*$, and to this end we use the second-degree Taylor polynomial of $\breve{Q}_{\vec{n}}$ instead of the first-degree one, that yields the second-order linear operator $p_\zeta^{[2]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]$ (see (262) for an explicit formula). To reduce $O(\varrho)$ we have to take into account finer effects related to the convolution integrals in the nonlinearity, which, in turn, are reflected in the frequency dependence of the susceptibility. Since $\alpha \sim \varrho$, the fifth-rank nonlinearity with the coefficient α^2 in the expansion of the nonlinearity in the NLM also contributes to this term and to make it smaller we have to take the quintic term into account.

1.4.4 Corrections related to the frequency dependence of the susceptibility. To take into account the first-order corrections due to the frequency dependence of the susceptibility tensor (discussed in detail in section 6), some terms involving time derivatives must be added to the ENLS (109) yielding the following ENLSs

$$\begin{aligned}\left(\partial_t + \mathrm{i}\mathcal{L}_\zeta^{[4]}\right) Z_\zeta &= \alpha_\pi p_\zeta^{[2]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]\left(Z_\zeta^2 Z_{-\zeta}\right) - \alpha_\pi \delta_{1,\zeta} Z_\zeta Z_{-\zeta} \\ &\times \left(\partial_t + \mathrm{i}\mathcal{L}_\zeta^{[4]}\right) Z_\zeta - \alpha_\pi \delta_{2,\zeta}\, Z_\zeta^2 \left(\partial_t + \mathrm{i}\mathcal{L}_{-\zeta}^{[4]}\right) Z_{-\zeta},\ \zeta = \pm \end{aligned} \tag{114}$$

with the initial conditions (102) (here we consider the case when the fifth-order term in the nonlinearity in the NLM is absent; a general case is considered a little later). The coefficients $\delta_{1,\pm}$, $\delta_{2,\pm}$ in (114) are proportional to the derivatives of the susceptibility with respect to the frequency and defined by (442) and (393). The Fourier transforms of the new terms in (114) create the same FNLR as the terms in (440) (for details see sections 6.2 and 8.3.2). The approximation error is of order

$$\mathbf{U} - \mathbf{U}_Z = O(\beta^3) + O(\beta\varrho) + O(\varrho) \tag{115}$$

The parameter ϱ does not enter into (114), but it is important for the matching of the initial data for the NLS with the excitation currents for the NLM (see section 5.2 for details) and determines the slow time scale in the NLM. Note that the last term in (115) corresponds to interband interactions and is discussed below.

Now we briefly discuss the relative magnitude of the terms in (114). It follows from (114) that

$$\left(\partial_t + \mathrm{i}\mathcal{L}_\zeta^{[4]}\right) Z_\zeta = O(\alpha) \tag{116}$$

Hence, the correction terms due to the frequency dependence have magnitude $O(\alpha^2)$. This agrees with the scaling $\varrho \sim \alpha$ in (15). Now we compare the contribution of the corrective terms with the terms of order $O(\alpha^2)$ that come from other sources. When the fifth- and higher order terms in the expansion of $\mathcal{F}_{\mathrm{NL}}(\mathbf{U})$ are much smaller than $O(\alpha^2)$, namely (401) holds, they can be neglected. When the fifth-order term in $\mathcal{F}_{\mathrm{NL}}(\mathbf{U})$ is exactly of order $O(\alpha^2)$, it has to be taken into account by adding the terms of the form $Q_{5,\zeta} Z_{-\zeta}^2 Z_\zeta^3$ to (114) (an explicit formula for $Q_{5,\pm}$ is given in (410)) thus reducing the error of the approximation from this source to $O(\beta\alpha)$.

1.4.4.1 *Simplification of the system.* Now we can simplify (114). We consider the case (38). First, we write (114) in the form

$$(1 + \alpha_\pi \delta_{1,\zeta} |Z_\zeta|^2)\left[\partial_t + \mathrm{i}\mathcal{L}_\zeta^{[4]}\right] Z_\zeta + \alpha_\pi \delta_{2,\zeta} Z_\zeta^2 \left(\left(\partial_t + \mathrm{i}\mathcal{L}_\zeta^{[4]}\right) Z_\zeta\right)^* = \alpha_\pi p_\zeta^{[2]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]\left(Z_\zeta^2 Z_\zeta^*\right) \tag{117}$$

and solve for $(\partial_t + \mathrm{i}\mathcal{L}_\zeta^{[4]})Z_\zeta$ obtaining an equation equivalent to (117) where we keep terms of order α_π and α_π^2. According to (113) we can neglect spatial derivatives in terms with the factor α_π^2 (this requires, strictly speaking, some regularity of solutions of the ENLS (see section 7 for references)). Therefore (117) can be rewritten as:

$$\left[\partial_t + \mathrm{i}\mathcal{L}_\zeta^{[4]}\right]Z_\zeta = \alpha_\pi\left(p_\zeta^{[2]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]\left(Z_\zeta^2 Z_\zeta^*\right)\right) - \alpha_\pi^2\delta_{1,\zeta}Q_\zeta|Z_\zeta|^4 Z_\zeta$$
$$-\alpha_\pi^2\delta_{2,\zeta}Q_\zeta^*|Z_\zeta|^4 Z_\zeta + O(\alpha^3) + O(\beta\alpha^2)$$

Consequently, we introduce the following equation with a quintic nonlinearity

$$\left[\partial_t + \mathrm{i}\mathcal{L}_\zeta^{[4]}\right]Z_\zeta = \alpha_\pi\left(p_\zeta^{[2]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]\left(Z_\zeta^2 Z_\zeta^*\right)\right) + \alpha_\pi^2\delta_{5,\zeta}|Z_\zeta|^4 Z_\zeta \tag{118}$$

$$\delta_{5,\zeta} = -\delta_{2,\zeta}Q_\zeta^* - \delta_{1,\zeta}Q_\zeta,\ \zeta = \pm \tag{119}$$

and with the initial condition (102). The solution of this equation approximates the solution of (114). Solutions of (118) approximate solutions of the NLM with the same order of accuracy as solutions of (114). Hence, (58) gives an approximate solution to the NLM with the error estimate (115). Note that to take into account the fifth-order term in the expansion of $\mathcal{F}_{\mathrm{NL}}(\mathbf{U})$ we have to use the coefficients $Q_{5,\pm}$ defined by (410), which effect the values of $\delta_{5,+}$, $\delta_{5,-}$ in (118). Namely, the values of the coefficients that take into account the fifth-order terms of the NLM are

$$\delta_{5,\zeta} = -\delta_{2,\zeta}Q_\zeta^* - \delta_{1,\zeta}Q_\zeta + Q_{5,\zeta}, \quad \zeta = \pm \tag{120}$$

with $Q_{5,\pm}$ and $\delta_{1,\pm}$, $\delta_{2,\pm}$ respectively defined by (410) and (442).

It is interesting that the both refinements coming from the frequency dependence of the cubic susceptibility and the fifth-order susceptibility are taken care of by the same fifth-order term $\delta_{5,\pm}|Z_\pm|^4\, Z_\pm$ *in the NLS.* In conclusion, to take into account these effects we take in (58) the solution $Z_\pm$ of (118).

If the quintic terms of the NLM are taken into account as in (120), the excitation currents of NLM are formed as in section 7 and $\varrho \sim \alpha$, the estimate of error of approximation by solutions of (118) still takes the form (115). The third term in (115) has a different nature. It is caused by interband interactions; we discuss this in the section right after this discussion. When (118) is extended to take care of the interband interactions, (115) takes the form of (94). In particular, for the classical NLS scaling (26) the error is $O(\beta^3)$. Let us compare (118) with the classical NLS. According to (113), the corrective terms

$$\alpha_\pi p_\zeta^{[2]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]\left(Z_\zeta^2 Z_\zeta^*\right) \tag{121}$$

involving the second-order derivatives which we have added here are estimated by $O(\alpha\beta^2) = O(\beta^4)$. The quintic corrective term $\alpha_\pi^2\delta_{5,\zeta}|Z_\zeta|^4 Z_\zeta$ is estimated by $O(\alpha^2) = O(\beta^4)$ too, and its effect $\varrho^{-1}O(\beta^4) = O(\beta^2)$ on the solution is larger than the difference $O(\beta^3)$ of NLM and ENLS solutions in (94).

1.4.4.2 *Effect of interband interactions.* In the one-dimensional case of classical NLS scaling the term $O(\varrho)$ in (115) has magnitude $O(\beta^2)$ and gives a leading contribution to the error. This term originates from the interband interactions, i.e. non-frequency-matched (non-FM) interactions that involve indirectly excited modes. The significance and exact contribution of these interactions to the NLM can be found when all higher order terms in the analytic expansion (160) are taken into account. We can construct the ENLS system which takes the effect of such interactions into account and admits an improved error estimate, namely $O(\varrho)$ in (115) is replaced by $O(\varrho^2)$ yielding (94), i.e. the total error is reduced from $O(\beta^2)$ to $O(\beta^3)$. These ENLSs in addition to (118) have to include another pair of scalar nonlinear

Schrödinger-type equations for auxiliary functions $Z_{+,1}$ and $Z_{-,1}$ with zero initial data; there are additional nonlinear Schrödinger-type nonlinear terms which couple these two equations with (118) forming a four-component system. For example, a cubic term with Z_+^3 is included in the equation for $Z_{+,1}$ and a cubic term with $(Z_+^*)^2 Z_{+,1}$ in the equation for $Z_{+,1}$. An additional two equations for $Z_{+,1}$ and $Z_{-,1}$ modify the dynamics of Z_+ and Z_- but (58) preserves its form and includes only Z_+ and Z_- . The coefficients at the coupling cubic terms describe nonlinear interactions between spectral bands related to the third harmonic generation. Since a detailed explanation and introduction of the coupling coefficients would require new notations and techniques which are beyond the scope of this paper we leave it for a future article.

1.4.5 Complex initial data. There are (see [35]) situations when complex electromagnetic vector fields are of interest and useful [35]. In this case the excitation currents are still given essentially by (55), but now $\hat{h}_\zeta$ and $\hat{h}_{-\zeta}$ are unrelated, and consequently the current can be complex-valued. This also may happen if (6) does not hold. In this case, in contrast to (193), the conjugation property does not have to hold and generically we may have

$$\hat{h}_+(\mathbf{q}) \neq \hat{h}_-(-\mathbf{q})^* \tag{122}$$

In this case the initial condition (102) does not involve the restriction (103). Consequently, we cannot assume that $Z_- = Z_+^*$, in (114). This system can also be reduced to the system (118) with a quintic nonlinearity. The estimate (115) holds in the complex-valued case too.

Note that if

$$\hat{h}_+(\mathbf{q}) \neq 0, \hat{h}_-(\mathbf{q}) = 0 \tag{123}$$

the solutions of (114) and (118) with $\zeta = -$ equal zero:

$$Z_-(\mathbf{r}, t) = 0 \tag{124}$$

Substituting $Z_- = 0$ into (114), we observe that $Z_+(\mathbf{r}, t)$ becomes a solution of the *linear* Schrödinger equation. *This fact shows that the nonlinearity in the classical NLS stems from the interaction of two modes of the doublet* $\{(+, n_0, \mathbf{k}_*), (-, n_0, -\mathbf{k}_*)\}$ *and when one of the modes is not initially excited the nonlinear interaction disappears at the prescribed accuracy level.*

1.4.6 Bidirectional waves. If the linearly excited waves propagate in both directions $\pm\nabla\omega_{n_0}(\mathbf{k}_*)$ the non-FM interactions between two wavepackets are of the same order as the first-order susceptibility corrections in (114). The corresponding interactions involve four modes $\tilde{U}_{\zeta,n_0}(\pm\mathbf{k}_* + \boldsymbol{\eta}, t)$, $\zeta = \pm 1$, and their dynamics is approximated by the ENLS solutions Z_ζ^ϑ, $\vartheta = \pm$. The ENLS system in the general complex currents case consists of four coupled equations. Let us consider here the system in the simplest case when the excitation currents and the nonlinearity are real and we use (38) (for the general case see section 5.4, in particular (353)). In this case $Z_{-\zeta}^\pm(\mathbf{r}, t) = Z_\zeta^\pm(\mathbf{r}, t)^*$ and the system reduces to the following two equations:

$$\begin{aligned}&\left[\partial_t + \mathrm{i}\gamma_{(4)}[-\mathrm{i}\vartheta\vec{\nabla}_{\mathbf{r}}]\right] Z_+^\vartheta + \alpha_\pi \delta_{\times,+}^\vartheta (|Z_+^{-\vartheta}|^2 Z_+^{-\vartheta *})\\ &\quad = -\alpha_\pi \delta_{1,+}^\vartheta Z_+^\vartheta Z_+^{\vartheta *}\left[\partial_t + \mathrm{i}\gamma_{(4)}[-\mathrm{i}\vartheta\vec{\nabla}_{\mathbf{r}}]\right] Z_+^\vartheta - \alpha_\pi \delta_{2,+}^\vartheta (Z_+^\vartheta)^2 \left[\partial_t - \mathrm{i}\gamma_{(4)}[\mathrm{i}\vartheta\nabla_{\mathbf{r}}]\right] Z_+^{\vartheta *}\\ &\qquad + \alpha_\pi p_+^{\vartheta,[2]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]((Z_+^\vartheta)^2 Z_+^{\vartheta *}), \ \ \vartheta = \pm\end{aligned} \tag{125}$$

We can approximate this system similarly to (118) by the system

$$\begin{aligned}&[\partial_t + \mathrm{i}\gamma_{(4)}[-\mathrm{i}\vartheta\vec{\nabla}_{\mathbf{r}}]] Z_+^\vartheta = \alpha_\pi p_+^{\vartheta,[2]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]((Z_+^\vartheta)^2 Z_+^{\vartheta *})\\ &-\alpha_\pi \delta_{\times,+}^\vartheta (|Z_+^{-\vartheta}|^2 Z_+^{-\vartheta *}) + \alpha_\pi^2 \delta_{5,+}^\vartheta |Z_+^\vartheta|^4 Z_+^\vartheta, \quad \vartheta = \pm\end{aligned} \tag{126}$$

where

$$\delta_{5,+}^{+} = -\delta_{2,+}^{+} Q_{+}^{+*} - \delta_{1,+}^{+} Q_{+}^{+}, \quad \delta_{5,+}^{-} = -\delta_{2,+}^{-} Q_{+}^{-*} - \delta_{1,+}^{-} Q_{+}^{-} \tag{127}$$

and $\delta_{\times,\zeta}^{\pm}$ are coefficients defined by (349), (345), and (346).

If (i) the fifth-order nonlinear terms in the NLM are taken into account by (120), (ii) (97) holds, (iii) excitation currents are formed as in section 7 and (iv) $\varrho \sim \alpha$, then the approximation error estimate (115) holds. To obtain an improved estimate (94) the effect of interband interactions has to be taken into account by adding four equations for $Z_{\zeta,1}^{\vartheta}$ with cubic coupling terms.

Note that the substitution $Z_{+}^{\vartheta} = \mathrm{e}^{-\mathrm{i}\gamma_0\tau/\varrho} z_{+}^{\vartheta}$, where $\gamma_0 = \omega_{n_0}(\mathbf{k}_*)$, $\tau = \varrho t$, transforms the system (126) into a similar one, namely

$$\left[\partial_\tau + \frac{\mathrm{i}}{\varrho}\gamma_{(4)}^{0}[-\mathrm{i}\vartheta\vec{\nabla}_{\mathbf{r}}]\right] z_{+}^{\vartheta} = \frac{\alpha_\pi}{\varrho}[p_{+}^{\vartheta,[2]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]((z_{+}^{\vartheta})^2 z_{+}^{\vartheta*}) - \delta_{\times,+}^{\vartheta}\mathrm{e}^{\mathrm{i}\gamma_0\tau/\varrho}(|z_{+}^{-\vartheta}|^2 z_{+}^{-\vartheta*})$$
$$+ \alpha_\pi \delta_{5,+}^{\vartheta}|z_{+}^{\vartheta}|^4 z_{+}^{\vartheta}] \tag{128}$$

with the differential operator $\gamma_{(3)}^{0}[\mathrm{i}\nabla_{\mathbf{r}}]$ having no zero-order terms. This system has the oscillatory coefficients $\alpha_\pi \delta_{\times,+}^{\pm}\mathrm{e}^{\mathrm{i}\gamma_0\tau/\varrho}$ accounting for the effect of the non-FM interactions. Integration over τ of these coefficients produces the factor $\varrho/\omega_{n_0}(\mathbf{k}_*)$ and we get for $\tau' \leq \tau_*$:

$$\alpha_\pi \int_0^{\tau'} \delta_{\times,+}^{-}\mathrm{e}^{\mathrm{i}\gamma_0\tau/\varrho}(|z_{+}^{+}|^2 z_{+}^{+*})\,\mathrm{d}\tau = O(\alpha\varrho) \tag{129}$$

thus showing that the coupling interactions are suppressed due to the frequency mismatch and $\alpha_\pi \delta_{\times,+}^{+}(|Z_{+}^{-}|^2 Z_{+}^{-*})$ in the case of the classical NLS scaling after the integration have the same order of magnitude $O(\alpha\varrho) = O(\beta^4)$ as $\alpha_\pi^2 \delta_{5,+}^{+}|Z_{+}^{+}|^4 Z_{+}^{+} = O(\alpha^2) = O(\beta^4)$ and fourth-order derivatives in $\gamma_{(4)}^{0}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]Z_{+}^{+}$ which are also $O(\beta^4)$.

Remark For a derivation based on anharmonic Maxwell–Lorenz system of coupled-mode equations which describe bidirectional propagation of waves in one-dimensional periodic structures, see [36] and references therein. ⧫

1.4.7 Other scalings and the reduction of ENLS. Recall that when deriving the NLS and ENLS and providing the related error estimates, we allow an arbitrary power dependence between the parameters ϱ and β. The condition (15) on ϱ and α also has the form (16), we take here $\varkappa_0 = 1$, that is $\alpha \sim \varrho$. The properties of the ENLS in different ranges of the parameters imply corresponding properties for the NLM. The ENLSs themselves can be reduced to simpler equations by formally throwing away higher order terms. Estimating the order of terms in the ENLS one has to take into account the discussion at the beginning of section 1.4.

1.4.7.1 *Example of a strongly dispersive scaling.* For example, let us consider the particular case

$$\alpha \sim \varrho \sim \beta^3 \tag{130}$$

implying strong dispersion

$$\theta^{-1} = \frac{\beta^2}{\varrho} \gg 1 \tag{131}$$

The right-hand side of (94) is now $O(\beta^3)$. The terms of order $O(\alpha\varrho) = O(\beta^6)$ and$O(\alpha^2) = O(\beta^6)$ in (126) after multiplication by $\varrho^{-1} = \beta^3$ are now of the same order as the error and can be neglected. After discarding the higher order terms in (126) we get the reduced equation

$$\left[\partial_t + \mathrm{i}\gamma_{(4)}[-\mathrm{i}\vartheta\vec{\nabla}_{\mathbf{r}}]\right]Z_+^{\vartheta} = \alpha_\pi p_+^{\vartheta,[2]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]((Z_+^{\vartheta})^2 Z_+^{\vartheta *}) \tag{132}$$

and the equation for Z_+^+ is now decoupled from the equation for Z_+^- at the level of accuracy $O(\beta^3)$ on the time interval of length τ_*/β^3. Note that the coupling terms with the inter-band component mentioned at the end of section 1.4.4 also are of order $O(\alpha\varrho)$ and can be neglected.

1.4.7.2 *Example of a weakly dispersive scaling.* Let us consider the particular case when (30) holds:

$$\varrho \sim \beta, \ \alpha \sim \varrho \tag{133}$$

First, consider the third-order ENLS (109). The error term in (112) is now of order

$$O(\beta^2) + O(\varrho) = O(\beta) \tag{134}$$

Since (113) holds, the second and third derivatives in $\mathcal{L}_+^{[3]} Z_+$ have order β^2 and β^3, respectively; they are $O(\beta^2)$ and have to be thrown away. The first derivative in $\alpha_\pi p_+^{[1]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}](Z_+^2 Z_-)$ is also $O(\alpha\beta) = O(\beta^2)$. The reduced equation takes the form of a first-order hyperbolic equation

$$\partial_t Z_+ + \mathrm{i}\mathcal{L}_+^{[1]} Z_+ = \alpha_\pi p_+^{[0]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}](Z_+^2 Z_-) \tag{135}$$

In the case of the space dimension $d = 1$, $\mathbf{r} = x$, the reduced equation takes the form

$$\partial_t Z_+ + \mathrm{i}\gamma_0 Z_+ + \frac{\beta\gamma_1}{\varrho}\partial_x Z_+ = \alpha_\pi Q_+ |Z_+|^2 Z_+ \tag{136}$$

and a similar equation for $Z_- = Z_+^*$. This system approximates the NLM with the accuracy $O(\beta)$ in the case of the scaling (133). As we have pointed out earlier, the condition $\beta^2/\varrho \ll 1$ in (30) implies that the dispersive effects are small, which agrees with the form of equation (135) which does not include dispersive terms.

Remark The described reduction of the universal ENLS to the reduced ENLS in the case of particular scaling relations between parameters α, β and ϱ is quite simple. The nontrivial part is the validity of the error estimates in the whole range of parameters, which guarantees that the reduced equations approximate well the exact solutions of the NLM itself. The estimates imply that the differences between different reduced ENLSs correspond to actual differences between different classes of solutions of the NLM which are generated by different initial excitations. ⧫

2. Modal decompositions and power series expansions of the linear and the first nonlinear responses

The very form (58) of the approximation $\mathbf{U}_{Z,n_0}(\mathbf{r}, t)$ is based on the modal decomposition. Recall that one of our goals is the construction of excitation currents producing waves governed essentially by NLSs. This construction is carried out in terms of the modal decomposition of all fields which is absolutely instrumental to the analysis of nonlinear wave propagation [17]. We are particularly interested in approximations for propagating waves as the quantities α, ϱ and β approach zero, and these approximations are constructed based on relevant asymptotic expansions of the involved fields.

2.1 *Bloch modal decomposition*

We systematically use modal decompositions based on the Bloch eigenmodes $\tilde{\mathbf{G}}_{\bar{n}}(\mathbf{r},\mathbf{k})$ of the linear Maxwell operator $\mathbf{M}$ in (3):

$$\mathbf{M}\tilde{\mathbf{G}}_{\bar{n}}(\mathbf{r},\mathbf{k}) = \omega_{\bar{n}}(\mathbf{k})\tilde{\mathbf{G}}_{\bar{n}}(\mathbf{r},\mathbf{k}) \tag{137}$$

where

$$\bar{n} = (\zeta, n),\ n = 1, 2\ldots,\ \zeta = \pm 1;\ \omega_{\bar{n}}(\mathbf{k}) = \omega_{\zeta,n}(\mathbf{k}) = \zeta\omega_n(\mathbf{k}),\ \omega_{n+1}(\mathbf{k}) \geq \omega_n(\mathbf{k}) \geq 0 \tag{138}$$

n being the band number and $\mathbf{k}$ being the quasimomentum taking values in the Brillouin zone. For notational simplicity we consider the cubic lattice with the lattice constant $L = 1$ and, consequently, with the Brillouin zone being the cube $[-\pi, \pi]^3$. Note that for given n and $\mathbf{k}$ there are exactly two eigenvalues $\omega_{\bar{n}}(\mathbf{k}) = \zeta\omega_n(\mathbf{k})$, $\zeta = \pm 1$. Schematic graphs of dispersion relations are given in figure 2.

The inversion symmetry condition (2) for the dispersion relations, which is

$$\omega_n(-\mathbf{k}) = \omega_n(\mathbf{k}),\ n = 1, 2, \ldots \tag{139}$$

readily implies the following properties of its first-, second- and higher order differentials

$$\omega_n'(-\mathbf{k}) = -\omega_n'(\mathbf{k}),\ \omega_n''(-\mathbf{k}) = \omega_n''(\mathbf{k}), \ldots,\ \omega_n^{(j)}(-\mathbf{k}) = (-1)^j\omega_n^{(j)}(\mathbf{k}),\ j = 1, 2, \ldots \tag{140}$$

The eigenmodes $\tilde{\mathbf{G}}_{\bar{n}}(\mathbf{r},\mathbf{k})$ are six-component vector fields satisfying the following relations [17]:

$$\tilde{\mathbf{G}}_{\bar{n}}(\mathbf{r},\mathbf{k}) = \begin{pmatrix} \tilde{\mathbf{G}}_{D,\bar{n}}(\mathbf{r},\mathbf{k}) \\ \tilde{\mathbf{G}}_{B,\bar{n}}(\mathbf{r},\mathbf{k}) \end{pmatrix},\ \vec{\nabla}_{\mathbf{r}} \cdot \tilde{\mathbf{G}}_{D,\bar{n}}(\mathbf{r},\mathbf{k}) = \vec{\nabla}_{\mathbf{r}} \cdot \tilde{\mathbf{G}}_{B,\bar{n}}(\mathbf{r},\mathbf{k}) = 0,\ \mathbf{r} \text{ in } [0,1]^3 \tag{141}$$

$$\tilde{\mathbf{G}}_{\bar{n}}(\mathbf{r}+\mathbf{m},\mathbf{k}) = \exp\ \{\mathrm{i}\mathbf{k}\cdot\mathbf{m}\}\ \tilde{\mathbf{G}}_{\bar{n}}(\mathbf{r},\mathbf{k}),\ \mathbf{m} \text{ in } \mathbf{Z}^3 \tag{142}$$

The properties of $\tilde{\mathbf{G}}_{\zeta,n}(\mathbf{r},\mathbf{k})$ and $\omega_n(\mathbf{k})$ are discussed in detail in [17]. We introduce the scalar product

$$(\mathbf{U},\mathbf{V})_{\mathcal{H}} = \int_{[0,1]^d} \mathbf{U}(\mathbf{r})\cdot\sigma_\varepsilon(\mathbf{r})\mathbf{V}(\mathbf{r})^*\ \mathrm{d}\mathbf{r},\ \sigma_\varepsilon(\mathbf{r}) = \begin{bmatrix} \varepsilon^{-1}(\mathbf{r}) & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \tag{143}$$

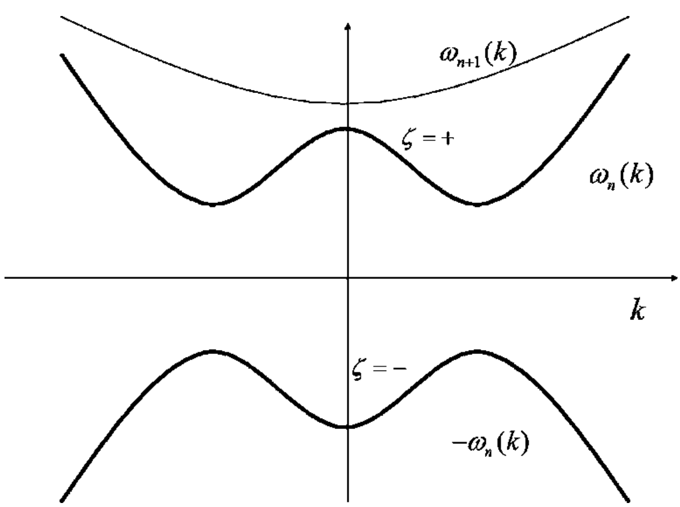

Figure 2. Schematic graphs of the dispersion relations corresponding to a pair of conjugate bands $\zeta\omega_n(k)$, $\zeta = \pm$, which are inversion symmetric.

and assume that $\tilde{\mathbf{G}}_{\zeta,n}(\mathbf{r},\mathbf{k})$ are orthonormal in $\mathcal{H}$:

$$\|\tilde{\mathbf{G}}_{\bar{n}}(\,\cdot\,,\mathbf{k})\|_{\mathcal{H}} = (\tilde{\mathbf{G}}_{\bar{n}}(\cdot,\mathbf{k}),\tilde{\mathbf{G}}_{\bar{n}}(\cdot,\mathbf{k}))_{\mathcal{H}}^{1/2} = 1 \tag{144}$$

Notice that if the condition (6) holds then the complex conjugate of every eigenmode $\tilde{\mathbf{G}}_{\zeta,n}(\mathbf{r},\mathbf{k})$ coincides with the eigenmode $\tilde{\mathbf{G}}_{-\zeta,n}(\mathbf{r},-\mathbf{k})$, i.e.

$$[\tilde{\mathbf{G}}_{\zeta,n}(\mathbf{r},\mathbf{k})]^* = \tilde{\mathbf{G}}_{-\zeta,n}(\mathbf{r},-\mathbf{k}) \tag{145}$$

under the assumption $\mathrm{Im}\varepsilon(\mathbf{r}) = \{\mathrm{Im}\varepsilon_{jm}(\mathbf{r})\}_{j,m=1}^3 = 0$.

Let us consider now a solution $\mathbf{U}(\mathbf{r},t)$ to the NLM (3) and its Floquet–Bloch modal decomposition [17]

$$\mathbf{U}(\mathbf{r},t) = \sum_{\bar{n}} \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \tilde{\mathbf{U}}_{\bar{n}}(\mathbf{r},\mathbf{k},t)\,\mathrm{d}\mathbf{k},\ \tilde{\mathbf{U}}_{\bar{n}}(\mathbf{r},\mathbf{k},t) = \tilde{U}_{\bar{n}}(\mathbf{k},t)\tilde{\mathbf{G}}_{\bar{n}}(\mathbf{r},\mathbf{k}) \tag{146}$$

where $\mathbf{U}_{\bar{n}}(\mathbf{r},\mathbf{k},t)$ are the modal components and $\tilde{U}_{\bar{n}}(\mathbf{k},t)$ are the (scalar) modal coefficients given by the formula

$$\tilde{U}_{\bar{n}}(\mathbf{k},t) = \int_{\mathbf{R}^d} \mathbf{U}(\mathbf{r},t)\cdot\sigma_\varepsilon(\mathbf{r})\tilde{\mathbf{G}}_{\bar{n}}^*(\mathbf{r},\mathbf{k})\ \mathrm{d}\mathbf{r} \tag{147}$$

The property (145) implies the following relations for the modal coefficients of the complex conjugate fields

$$(\widetilde{U^*})_{\zeta,n}(\mathbf{k},t) = [\tilde{U}_{-\zeta,n}(-\mathbf{k},t)]^* \tag{148}$$

The Floquet–Bloch transform $\tilde{\mathbf{U}}$ of $\mathbf{U}$, which involves all modes, is defined by the formula

$$\tilde{\mathbf{U}}(\mathbf{r},\mathbf{k},t) = \sum_{\bar{n}} \tilde{\mathbf{U}}_{\bar{n}}(\mathbf{r},\mathbf{k},t) = \sum_{\bar{n}} \tilde{U}_{\bar{n}}(\mathbf{k},t)\tilde{\mathbf{G}}_{\bar{n}}(\mathbf{r},\mathbf{k}) \tag{149}$$

It is often convenient to write the coefficients $\tilde{U}_{\bar{n}}(\mathbf{k},t)$ in a special form, namely

$$\tilde{U}_{\bar{n}}(\mathbf{k},t) = \tilde{u}_{\bar{n}}(\mathbf{k},\tau)\mathrm{e}^{-\mathrm{i}\omega_{\bar{n}}(\mathbf{k})t}, \quad \tau = \varrho t \tag{150}$$

factoring out the carrier frequency $\omega_{\bar{n}}(\mathbf{k})$. This equality defines the modal coefficient $\tilde{u}_{\bar{n}}(\mathbf{k},\tau)$ as a function of slow time τ. Similarly to (149) we define

$$\tilde{\mathbf{u}}(\mathbf{r},\mathbf{k},\tau) = \sum_{\bar{n}} \tilde{\mathbf{u}}_{\bar{n}}(\mathbf{r},\mathbf{k},\tau) = \sum_{\bar{n}} \tilde{u}_{\bar{n}}(\mathbf{k},\tau)\tilde{\mathbf{G}}_{\bar{n}}(\mathbf{r},\mathbf{k}) \tag{151}$$

2.2 *Nonlinearity and related power expansions*

The nonlinear term $\mathcal{F}_{\mathrm{NL}}(\mathbf{U})$ in the NLM (3) is given by the formula [17–20]

$$\mathcal{F}_{\mathrm{NL}}(\mathbf{U}) = \mathcal{F}_{\mathrm{NL}}(\mathbf{U};\alpha) = \begin{bmatrix} \mathbf{0} \\ \nabla\times\mathbf{S}_D(\mathbf{r},t;\mathbf{D};\alpha) \end{bmatrix},\ \mathbf{U}(\mathbf{r},t) = \begin{bmatrix} \mathbf{D}(\mathbf{r},t) \\ \mathbf{B}(\mathbf{r},t) \end{bmatrix} \tag{152}$$

where

$$\mathbf{S}_D(\mathbf{r},t;\mathbf{D}) = \mathbf{S}_D^{(3)}(\mathbf{r},t;\mathbf{D}) + \alpha\mathbf{S}_D^{(5)}(\mathbf{r},t;\mathbf{D}) + \alpha^2\mathbf{S}_D^{(7)}(\mathbf{r},t;\mathbf{D}) + \cdots \tag{153}$$

is a series of causal integral operators $\mathbf{S}_D^{(2n+1)}$ that are determined based on the response functions from (8). Notice that the representation (153) consists of only odd-order terms as it is typical for dielectric media allowing NLS regimes. The dominant cubic nonlinearity is given by the causal integral

$$\mathbf{S}_D^{(3)}(\mathbf{r},t;\mathbf{D}) = \int_{-\infty}^t\int_{-\infty}^t\int_{-\infty}^t \mathbf{R}_D^{(3)}(\mathbf{r};t-t_1,t-t_2,t-t_3)\vdots\prod_{j=1}^3 \mathbf{D}(\mathbf{r},t_j)\,\mathrm{d}t_j \tag{154}$$

where the trilinear tensorial operator $\mathbf{R}_D^{(3)}$ is assumed to be symmetric:

$$\mathbf{R}_D^{(3)}(\mathbf{r}; t-t_1, t-t_2, t-t_3) \vdots \mathbf{D}_1\mathbf{D}_2\mathbf{D}_3 = \cdots = \mathbf{R}_D^{(3)}(\mathbf{r}; t-t_2, t-t_1, t-t_3) \vdots \mathbf{D}_2\mathbf{D}_1\mathbf{D}_3 \quad (155)$$

An alternative and often used representation of the polarization is through its frequency dependent susceptibility tensor $\chi_D^{(3)}$,

$$\chi_D^{(3)}(\mathbf{r}; \omega_1, \omega_2, \omega_3) = \int_0^\infty \int_0^\infty \int_0^\infty \mathbf{R}_D^{(3)}(\mathbf{r}; t_1, t_2, t_3)\, \mathrm{e}^{\{\mathrm{i}(\omega_1 t_1 + \omega_2 t_2 + \omega_3 t_3)\}}\, \mathrm{d}t_1\, \mathrm{d}t_2\, \mathrm{d}t_3 \quad (156)$$

Note that the standard frequency-dependent susceptibility tensor $\chi^{(3)}(\mathbf{r}; \boldsymbol{\omega})$ is determined in terms of the nonlinear polarization $\mathbf{P}_{\mathrm{NL}}(\mathbf{r}, t; \mathbf{E}(\cdot))$ of the medium by a formula similar to (156) (see [21]). The tensor $\chi_D^{(3)}(\mathbf{r}; \boldsymbol{\omega})$ (which acts on $\mathbf{D}$) is expressed in terms of $\chi^{(3)}(\mathbf{r}; \boldsymbol{\omega})$ (which acts on $\mathbf{E}$) and the dielectric tensor $\varepsilon(\mathbf{r})$ by the following formula (see [17, 20] for details):

$$\chi_D^{(3)}(\mathbf{r}; \boldsymbol{\omega}) \vdots \prod_{j=1}^{3} \mathbf{D}_j = 4\pi \varepsilon^{-1}(\mathbf{r}) \chi^{(3)}(\mathbf{r}; \boldsymbol{\omega}) \vdots \prod_{j=1}^{3} [\varepsilon^{-1}(\mathbf{r})\mathbf{D}_j] \quad (157)$$

Let us consider the power series expansion (11) for the exact solution $\mathbf{U}$ to the NLM (3) with the current $\mathbf{J}$ satisfying the relations (9) and (35), i.e.

$$\mathbf{U} = \mathbf{U}^{(0)} + \alpha\mathbf{U}^{(1)} + \alpha^2\mathbf{U}^{(2)} + \cdots, \ \mathbf{J} = \mathbf{J}^{(0)} + \alpha\mathbf{J}^{(1)} + \cdots, \quad (158)$$

$$\mathbf{J}^{(j)}(\mathbf{r}, t) = 0 \text{ if } t \le 0 \quad \text{or} \quad t \ge \frac{\tau_0}{\varrho}, \ j = 0, 1, \ldots \quad (159)$$

For every amplitude $\tilde{u}_{\bar{n}}(\mathbf{k}, \tau)$ defined by (146) and (150) the series corresponding to (158) becomes

$$\tilde{u}_{\bar{n}}(\mathbf{k}, \tau) = \sum_{m=0}^{\infty} \tilde{u}_{\bar{n}}^{(m)}(\mathbf{k}, \tau)\alpha^m \quad (160)$$

Power expansions for the amplitudes $\tilde{u}_{\bar{n}}(\mathbf{k}, \tau)$ as well as for other quantities of interest with respect to the small parameter α are given by convergent Taylor series. The expansions for the amplitudes $\tilde{u}_{\bar{n}}(\mathbf{k}, \tau)$ with respect to the small parameters ϱ and β are of a more complicated nature related to almost time-harmonic expansions and asymptotic expansions for oscillatory integrals (see equation (169) and also section 6.2).

We recall that the current $\mathbf{J}^{(1)}$ in (158) is introduced to provide proper transformation of the initial data for the NLS into the excitation current (see section 5.2 for details). The expansion (158) determines the *linear medium response* $\mathbf{U}^{(0)}$ and *the first nonlinear response* $\mathbf{U}^{(1)}$ satisfying, respectively, the evolution equations (13) and (14), i.e.

$$\partial_t \mathbf{U}^{(0)} = -\mathrm{i}\mathbf{M}\mathbf{U}^{(0)} - \mathbf{J}^{(0)}; \ \mathbf{U}^{(0)}(t) = 0 \text{ for } t \le 0 \quad (161)$$

$$\begin{aligned} \partial_t \mathbf{U}^{(1)} &= -\mathrm{i}\mathbf{M}\mathbf{U}^{(1)} + \mathcal{F}_{\mathrm{NL}}^{(1)}(\mathbf{U}^{(0)}) - \mathbf{J}^{(1)}; \ \mathbf{U}^{(1)}(t) = 0 \text{ for } t \le 0 \\ \mathcal{F}_{\mathrm{NL}}^{(1)}(\mathbf{U}^{(0)}) &= \mathcal{F}_{\mathrm{NL}}(\mathbf{U}^{(0)}; \alpha)|_{\alpha=0} \end{aligned} \quad (162)$$

We now introduce the currents $\mathbf{J}^{(0)}$ and $\mathbf{J}^{(1)}$ satisfying the conditions (159) by their modal coefficients as follows:

$$\begin{aligned}&\tilde{\mathbf{J}}_{\bar{n}}^{(j)}(\mathbf{r},\mathbf{k},t)=\tilde{J}_{\bar{n}}^{(j)}(\mathbf{k},t)\tilde{\mathbf{G}}_{\bar{n}}(\mathbf{r},\mathbf{k}),\ \tilde{J}_{\bar{n}}^{(j)}(\mathbf{k},t)=\mathrm{e}_{\bar{n}}^{-\mathrm{i}\omega_{\bar{n}}(\mathbf{k})t}\tilde{j}_{\bar{n}}^{(j)}(\mathbf{k},\tau),\ \tau=\varrho t\\&\tilde{j}_{\bar{n}}^{(j)}(\mathbf{k},\tau)=0\text{ if }\tau\leq 0\quad\text{or}\quad\tau\geq\tau_0,\ j=0,1\end{aligned}\tag{163}$$

For the currents $\mathbf{J}^{(0)}$ and $\mathbf{J}^{(1)}$ to be real, in view of (148), their modal coefficients should satisfy the relations

$$\left[j_{\zeta,n}^{(j)}(\mathbf{k})\right]^{*}=j_{-\zeta,n}^{(j)}(-\mathbf{k}),\ j=0,1\tag{164}$$

From (161), (162), (163) and (160) we get the following representation for the modal forms of the first two terms $\mathbf{U}^{(0)}$ and $\mathbf{U}^{(1)}$ of the power expansion (158):

$$\tilde{\mathbf{U}}_{\bar{n}}^{(j)}(\mathbf{r},\mathbf{k},t)=\tilde{U}_{\bar{n}}^{(j)}(\mathbf{k},t)\tilde{\mathbf{G}}_{\bar{n}}(\mathbf{r},\mathbf{k}),\ \tilde{U}_{\bar{n}}^{(j)}(\mathbf{k},t)=\tilde{u}_{\bar{n}}^{(j)}(\mathbf{k},\tau)\mathrm{e}^{-\mathrm{i}\omega_{\bar{n}}(\mathbf{k})t},\ \tau=\varrho t,\ j=0,1\tag{165}$$

where

$$\tilde{u}_{\bar{n}}^{(0)}(\mathbf{k},\tau)=-\int_0^{\tau}\tilde{j}_{\bar{n}}^{(0)}(\mathbf{k},\tau)\,\mathrm{d}\tau_1\tag{166}$$

$$\tilde{U}_{\bar{n}}^{(1)}(\mathbf{k},\tau)=\frac{1}{\varrho}\int_0^{\tau}\mathrm{e}^{-\mathrm{i}\omega_{\bar{n}}(\mathbf{k})\frac{(\tau-\tau_1)}{\varrho}}\left\{\left[\mathcal{F}_{\mathrm{NL}}^{(0)}\left(\mathbf{U}^{(0)}\right)\right]_{\bar{n}}(\mathbf{k},\tau)-\tilde{J}_{\bar{n}}^{(1)}(\mathbf{k},t)\right\}\mathrm{d}\tau_1\tag{167}$$

Similarly to (151) we introduce

$$\tilde{\mathbf{u}}^{(0)}(\mathbf{r},\mathbf{k},\tau)=\sum_{\bar{n}}\tilde{\mathbf{u}}_{\bar{n}}^{(0)}(\mathbf{r},\mathbf{k},\tau)=\sum_{\bar{n}}\tilde{u}_{\bar{n}}^{(0)}\tilde{\mathbf{G}}_{\bar{n}}(\mathbf{r},\mathbf{k})\tag{168}$$

2.2.1 Structured asymptotic expansions.

We are interested in asymptotic approximations with respect to ϱ,β of the coefficients $\tilde{u}_{\bar{n}}^{(j)}(\mathbf{k},\tau)=\tilde{u}_{\bar{n}}^{(j)}(\mathbf{k},\tau;\varrho,\beta)$ of the expansion (158) given by (165). We primarily study the modal amplitudes $\tilde{u}_{\bar{n}}^{(1)}(\mathbf{k},\tau)=\tilde{u}_{\bar{n}}^{(1)}(\mathbf{k},\tau;\varrho,\beta)$ of the FNLR $\mathbf{U}^{(1)}$ for small ϱ and β. Our analysis shows that the dependence on ϱ and β is more complicated than on α. There are three different types of asymptotic expansions involved in the description of the dependence of $\tilde{u}_{\bar{n}}^{(1)}(\mathbf{k},\tau;\varrho,\beta)$ on ϱ, β; the first type involves powers of ϱ; the second one involves powers of β; and the third type involves powers of $\beta^{\nu+1}/\varrho$ if $\theta^{-1}=\beta^2/\varrho\leq 1$, where $\nu=2,3,4$ is the order of the NLS or ENLS. The resulting expansion of the modal amplitudes of the NLM has the form of a *structured power asymptotic series.* In the weakly dispersive case $\theta^{-1}=\beta^2/\varrho\leq 1$ we have the following expansions:

$$\begin{aligned}\tilde{u}_{\zeta,n_0}^{(1)}(\mathbf{k},\tau;\varrho,\beta)&=\frac{1}{\varrho}\sum_{l_1=0}^{N_1}\sum_{l_2=0}^{N_2}\sum_{l_3=0}^{N_3}c_{l_1,l_2,l_3}^{\mathrm{NLM}}(\mathbf{k},\tau;\zeta,n_0)\varrho^{l_1}\beta^{l_2}\left(\frac{\beta^{\nu+1}}{\varrho}\right)^{l_3}\\&\quad+\frac{1}{\varrho}\left[O(\beta\varrho^{N_1})+O(\beta^{N_2+1})+O\left(\left(\frac{\beta^{\nu+1}}{\varrho}\right)^{N_3+1}\right)\right]\end{aligned}\tag{169}$$

The powers $(\beta^{\nu+1}/\varrho)^{l_3}$ (that is $(\beta^4/\varrho)^{l_3}$ for $\nu=3$) come from the expansion in (273). *We would like to emphasize that the form (169) for $\tilde{u}_{\zeta,n_0}^{(1)}(\mathbf{k},\tau;\varrho,\beta)$ is not imposed as an ansatz, but it follows from the analysis of the interaction integrals, and it describes properties of exact solutions to the NLM.* Powers ϱ^{l_1} stem from the asymptotic expansions for almost monochromatic waves (for details see section 6). Some expressions in the interaction integrals admit regular Taylor expansions (see section 4.1.1) that lead to the powers β^{l_2}. The complexity of the expression (169) reflects the complexity of exact solutions to the NLM. *The type of dependence in (169) shows that formal asymptotic expansions of solutions with respect to* *powers of independent parameters ϱ and β are not very useful since such expansions would involve negative powers of the small parameter ϱ.* In addition, this form of dependence implies that if one prescribes power relations of the form (16), (29) or (31), the resulting expansions in powers of one remaining parameter would strongly depend on the choice of $\varkappa_1$ and $\varkappa_0$, whereas the higher order approximating ENLS, which we introduce, are *universal*—they do not depend on $\varkappa_1$ and $\varkappa_0$. Introduction of a specific power dependence $\varrho = \beta^{\varkappa_1}$ selects from the universal ENLS, which is described in section 1.4, a specific reduced ENLS depending on the choice of $\varkappa_1$ (see section 1.4.7 for examples).

Our strategy for approximating $\tilde{u}^{(1)}_{\zeta,n_0}$ by a solution of the NLS can be described as follows. We consider solutions $V_\zeta(\mathbf{r}, t)$ of the NLS, take their Fourier transform $\hat{V}_\zeta(\boldsymbol{\eta}, t)$ and similarly to (150) introduce slowly varying coefficients $\hat{v}_\zeta(\boldsymbol{\eta}, \tau)$

$$\hat{V}_\zeta(\boldsymbol{\eta}, t) = \hat{v}_\zeta(\boldsymbol{\eta}, \tau) e^{-\mathrm{i}\zeta\gamma_{(\nu)}(\zeta\boldsymbol{\eta})t}, \ \tau = \varrho t \tag{170}$$

Similarly to (160) we introduce asymptotic expansions in α

$$\hat{v}_\zeta(\boldsymbol{\eta}, \tau) = \hat{v}^{(0)}_\zeta(\boldsymbol{\eta}, \tau) + \alpha\hat{v}^{(1)}_\zeta(\boldsymbol{\eta}, \tau) + \cdots \tag{171}$$

(see section 5.2 for details). We expand the Fourier transform of a solution to the NLS similarly to (169): and respective expansion for $\hat{v}^{(1)}_\zeta$

$$\hat{v}^{(1)}_\zeta(\boldsymbol{\eta}, \tau; \varrho, \beta) = \varrho^{-1} \sum_{l_1=0}^{N_1} \sum_{l_2=0}^{N_2} \sum_{l_3=0}^{N_3} c^{\mathrm{NLS}}_{l_1,l_2,l_3}(\boldsymbol{\eta}, \tau; \zeta) \varrho^{l_1} \beta^{l_2} \left(\frac{\beta^{\nu+1}}{\varrho}\right)^{l_3}$$
$$+ \frac{1}{\varrho}\left[O\left(\varrho^{N_1+1}\right) + O\left(\beta^{N_2+1}\right) + O\left(\left(\frac{\beta^{\nu+1}}{\varrho}\right)^{N_3+1}\right)\right] \tag{172}$$

The coefficients $C^{\mathrm{NLS}}_{l_1,l_2,l_3}$ of the expansion depend on the choice of parameters $p^{[\nu-2]}_{\pm}$ and $\delta_{1,\pm}$ in the NLS (114). These parameters are chosen so that the following conditions are satisfied:

$$C^{\mathrm{NLS}}_{l_1,l_2,l_3}(\boldsymbol{\eta}, \tau; \zeta) = C^{\mathrm{NLM}}_{l_1,l_2,l_3}(\mathbf{k}_* + \boldsymbol{\eta}, \tau; \zeta, n_0), \ l_1 \le N_1, \ l_2 \le N_2, \ l_3 \le N_3 \tag{173}$$

To satisfy the conditions we choose, in a proper way, the excitation currents for the NLM based on initial data for the NLS. In particular, if $\nu = 3$ we take $N_1 = 0$, $N_2 = 1$, $N_3 = 0$. The details of the related analysis are rather technical and are considered in subsequent sections. *Recall again that the form and coefficients of (169) are the result of explicitly defined transformations of the interaction integral and should be considered as a result of the analysis rather than a starting point.* Similarly, fulfilment of (173) follows from our choice of excitation currents and coefficients of the NLS, and is based on analysis of the interaction integrals.

The case of strong dispersion is similar, but expansions (169) and (172) would include powers of ϱ/β^2.

2.2.2 First nonlinear response and modal susceptibility. For the current $\mathbf{J}^{(1)}$ of the form (163) and $\varrho \to 0$ the FNLR $\mathbf{U}^{(1)}$ determined by (167) can be represented as the following series based on the time-harmonic expansion (see section 6.2 for details)

$$\mathbf{U}^{(1)} = \mathbf{U}^{(1,0)} + \varrho\mathbf{U}^{(1,1)} + \varrho^2\mathbf{U}^{(1,2)} + \cdots \tag{174}$$

$$\tilde{u}^{(1)}_{\bar{n}} = \tilde{u}^{(1,0)}_{\bar{n}} + \varrho\tilde{u}^{(1,1)}_{\bar{n}} + \varrho^2\tilde{u}^{(1,2)}_{\bar{n}} + \cdots \tag{175}$$

Notice that (174) and (175) are not the Taylor series, and the quantities $\mathbf{U}^{(1,s)}$ and $\tilde{u}^{(1,s)}_{\bar{n}}$, $s = 0, 1, \ldots$ in (174) and (175) are represented as oscillatory integrals which depend on ϱ themselves. As we will see in section 4.3 and 4.1.2 respectively if $\varrho \to 0$ we have $\tilde{u}^{(1,s)}_{\bar{n}} \sim \varrho^{d-1}$ for the dispersive case and $\tilde{u}^{(1,s)}_{\bar{n}} \sim \varrho^{-1}$ for the weakly dispersive case. Since we are interested in the FNLR $\mathbf{U}^{(1)}$ for small ϱ we look first at the dominant term $\mathbf{U}^{(1,0)}$ in the series (174). We refer to $\mathbf{U}^{(1,0)}$ as the *time-harmonic FNLR*. Using the formula (167) and the time-harmonic approximation of $\mathcal{F}_{\mathrm{NL}}^{(0)}$ together with the definitions of $\tilde{\mathbf{G}}_{\bar{n}}(\mathbf{r},\mathbf{k})$, the susceptibility $\chi_D^{(3)}$, the inner product $(\cdot,\cdot)_{\mathcal{H}}$ by respectively (141), (156) and (143) we get the following integral representation for $\tilde{u}_{\bar{n}}^{(1,0)}$ (see [17] and section 6.2 for details)

$$\begin{gathered}\tilde{u}_{\bar{n}}^{(1,0)}(\mathbf{k},\tau)=\frac{1}{\varrho}\sum_{\bar{n}',\bar{n}'',\bar{n}'''}\int_0^\tau\int_{\substack{[-\pi,\pi]^{2d}\\ \mathbf{k}'+\mathbf{k}''+\mathbf{k}'''=\mathbf{k}}}\exp\left\{\mathrm{i}\phi_{\bar{n}}(\vec{k})\frac{\tau_1}{\varrho}\right\}\\ \breve{Q}_{\bar{n}}(\vec{k})\tilde{u}_{\bar{n}'}^{(0)}(\mathbf{k}',\tau_1)\tilde{u}_{\bar{n}''}^{(0)}(\mathbf{k}'',\tau_1)\tilde{u}_{\bar{n}'''}^{(0)}(\mathbf{k}''',\tau_1)\,\mathrm{d}\mathbf{k}'\mathrm{d}\mathbf{k}''\mathrm{d}\tau_1-\tilde{u}_{\bar{n}}^{(1)}(\mathbf{J}_1;\mathbf{k},\tau)\end{gathered}\tag{176}$$

where

$$\begin{aligned}\vec{k}&=(\mathbf{k},\mathbf{k}',\mathbf{k}'',\mathbf{k}'''),\ \vec{\zeta}=(\zeta,\zeta',\zeta'',\zeta''')\\ \vec{n}&=(\bar{n},\bar{n}',\bar{n}'',\bar{n}''')=((\zeta,n),(\zeta',n'),(\zeta'',n''),(\zeta''',n'''))\end{aligned}\tag{177}$$

$$\phi_{\vec{n}}(\vec{k})=\zeta\omega_n(\mathbf{k})-\zeta'\omega_{n'}(\mathbf{k}')-\zeta''\omega_{n''}(\mathbf{k}'')-\zeta'''\omega_{n'''}(\mathbf{k}''')\tag{178}$$

$$\breve{Q}_{\vec{n}}(\vec{k})=\frac{1}{(2\pi)^{2d}}\left(\begin{bmatrix}\mathbf{0}\\ Q_{\chi_D^{(3)}}\end{bmatrix},\tilde{\mathbf{G}}_{\bar{n}}(\mathbf{r},\mathbf{k})\right)_{\mathcal{H}}\tag{179}$$

$$Q_{\chi_D^{(3)}}=\nabla\times\chi_D^{(3)}(\omega_{\bar{n}'}(\mathbf{k}'),\omega_{\bar{n}''}(\mathbf{k}''),\omega_{\bar{n}'''}(\mathbf{k}'''))\,\vdots\,\tilde{\mathbf{G}}_{D,\bar{n}'}(\mathbf{r},\mathbf{k}')\tilde{\mathbf{G}}_{D,\bar{n}''}(\mathbf{r},\mathbf{k}'')\tilde{\mathbf{G}}_{D,\bar{n}'''}(\mathbf{r},\mathbf{k}'''),$$

$\tilde{u}_{\bar{n}}^{(0)}$ are defined in (166) and

$$\begin{aligned}\tilde{u}_{\zeta,n_0}^{(1)}(\mathbf{J}_1;\mathbf{k},\tau)&=\frac{1}{\varrho}\int_0^\tau\tilde{j}_{\zeta,n_0}^{(1)}(\mathbf{k},\tau_1)\,\mathrm{d}\tau_1\\ \tilde{u}_{\bar{n}}^{(1)}(\mathbf{J}_1;\mathbf{k},\tau)&=0\text{ for }n\neq n_0\end{aligned}\tag{180}$$

The quantity $\breve{Q}_{\vec{n}}(\vec{k})$ given by the integral (179) plays an important role in the approximation analysis and we refer to it as the *modal susceptibility*. An estimate for the difference $\tilde{u}_{\bar{n}}^{(1)}(\mathbf{k},\tau)-\tilde{u}_{\bar{n}}^{(1,0)}(\mathbf{k},\tau)$ is given by (395).

Note that though the formula (179) uses a specific form of $\mathcal{F}_{\mathrm{NL}}$ in (152), that particular form it is not essential for our analysis. For example, if the **D**-component of $\mathcal{F}_{\mathrm{NL}}$ were not zero, or if $\chi_D^{(3)}$ acted also on the **B**-component of the vector **U**, all steps and conclusions of our analysis would remain the same.

We will also use the following notation which allows us to rewrite (176) in a shorter way:

$$\begin{gathered}\tilde{F}_{\bar{n}}\left[\left(\mathbf{u}^{(0)}\right)^3\right](\mathbf{k},\tau)=\sum_{\bar{n}',\bar{n}'',\bar{n}'''}\int_0^\tau\int_{\substack{[-\pi,\pi]^{2d}\\ \mathbf{k}'+\mathbf{k}''+\mathbf{k}'''=\mathbf{k}}}\\ \times\,\mathrm{e}^{\mathrm{i}\phi_{\vec{n}}(\vec{k})\frac{\tau_1}{\varrho}}\frac{\left(\nabla\times I_{\chi,\vec{n}}\left(\tilde{\mathbf{u}}^{(0)}\right),\tilde{\mathbf{G}}_{\bar{n}}(\mathbf{r},\mathbf{k})\right)_{\mathcal{H}}}{\varrho(2\pi)^{2d}}\mathrm{d}\mathbf{k}'\mathrm{d}\mathbf{k}''\mathrm{d}\tau_1\\ I_{\chi,\vec{n}}\left(\tilde{\mathbf{u}}^{(0)}\right)=\chi_{D,B}^{(3)}(\mathbf{r};\omega_{\bar{n}'}(\mathbf{k}'),\omega_{\bar{n}''}(\mathbf{k}''),\omega_{\bar{n}'''}(\mathbf{k}'''))\,\vdots\,\tilde{\mathbf{u}}_{\bar{n}'}^{(0)}(\mathbf{r},\mathbf{k}',\tau_1)\tilde{\mathbf{u}}_{\bar{n}''}^{(0)}(\mathbf{r},\mathbf{k}'',\tau_1)\tilde{\mathbf{u}}_{\bar{n}'''}^{(0)}(\mathbf{r},\mathbf{k}''',\tau_1)\end{gathered}\tag{181}$$

where $\tilde{\mathbf{u}}^{(0)}$ is defined by (168),

$$\tilde{\mathbf{u}}_{\bar{n}}^{(0)}(\mathbf{r},\mathbf{k},\tau)=\tilde{u}_{\bar{n}}^{(0)}(\mathbf{k},\tau)\tilde{\mathbf{G}}_{\bar{n}}(\mathbf{r},\mathbf{k}),\ \tilde{\mathbf{u}}_{\bar{n}^{(i)}}=\begin{bmatrix}\tilde{\mathbf{u}}_{D,\bar{n}^{(i)}}\\ \tilde{\mathbf{u}}_{B,\bar{n}^{(i)}}\end{bmatrix}\tag{182}$$

and

$$\chi^{(3)}_{D,B}(\mathbf{r};\omega_1,\omega_2,\omega_3)\vdots\tilde{\mathbf{u}}_{\bar{n}'}\tilde{\mathbf{u}}_{\bar{n}''}\tilde{\mathbf{u}}_{\bar{n}'''}=\begin{bmatrix}\mathbf{0}\\ \chi^{(3)}_{D}(\mathbf{r};\omega_1,\omega_2,\omega_3)\vdots\tilde{\mathbf{u}}_{D,\bar{n}'}\tilde{\mathbf{u}}_{D,\bar{n}''}\tilde{\mathbf{u}}_{D,\bar{n}'''}\end{bmatrix} \tag{183}$$

is a tensor obtained from $\chi^{(3)}_{D}$. This tensor acts not in the three-dimensional $\mathbf{D}$-space, but in the six-dimensional $(\mathbf{D},\mathbf{B})$-space. It acts on the $\mathbf{D}$-components of $\tilde{\mathbf{u}}^{(0)}_{\bar{n}'}$ taking values in the $\mathbf{B}$-component as in (179). Using (181) we can rewrite (176) in the form:

$$\tilde{u}^{(1,0)}_{\bar{n}}(\mathbf{k},\tau)=\tilde{F}_{\bar{n}}[\left(\mathbf{u}^{(0)}\right)^3](\mathbf{k},\tau)-\tilde{u}^{(1)}_{\bar{n}}(\mathbf{J}_1;\mathbf{k},\tau) \tag{184}$$

Below we analyse (176) using the approach of [17–20] in the case when the frequency of the excitation current (163) is in a fixed band n_0, the quasimomentum $\mathbf{k}$ is in a vicinity of a fixed quasimomentum $\mathbf{k}_*$ and the excitation current is almost time-harmonic (such currents are described in detail in the following section). The term $\tilde{u}^{(1)}_{\bar{n}}(\mathbf{J}_1;\mathbf{k},\tau)$ in (176), as one can see from (180), is due to the excitation current $\mathbf{J}_1$ with amplitudes $\tilde{j}^{(1)}_{\bar{n}}$. This current is introduced to transform the initial data for the NLS into a proper excitation current with maximal accuracy (see section 5.2 for details). The modal components of $j^{(1)}_{\bar{n}}$ are defined by the following formula

$$\tilde{j}^{(1)}_{\zeta,n_0}(\zeta\mathbf{k}_*+\boldsymbol{\eta},\tau)=\exp\left\{\mathrm{i}\omega_{\zeta,n_0}(\zeta\mathbf{k}_*+\boldsymbol{\eta})\frac{\tau_1}{\varrho}\right\}\hat{J}^{(1)}_{Z,\zeta}(\boldsymbol{\eta},t),\ t=\frac{\tau}{\varrho},\ \zeta=\pm=\pm1 \tag{185}$$

with $\hat{J}^{(1)}_{Z,\zeta}$ being defined by (330).

2.3 *Almost time-harmonic excitations*

The concept of an almost time-harmonic excitation is central to the theory of nonlinear mode interactions. An abstract form for an almost time-harmonic function $a(t)$ is given by formula (10). In this section we give a precise definition of an almost time-harmonic excitation current which generates nonlinear Schrödinger-type regimes for the NLM. The first step in setting up the current $\mathbf{J}$ as defined by (35) or (158) is to assume that its modal composition (163), (164) involves only a single spectral band with the index n_0, i.e.

$$\begin{aligned}\tilde{\mathbf{J}}^{(j)}_{n_0}(\mathbf{r},\mathbf{k},t)&=\tilde{j}^{(j)}_{+,n_0}(\mathbf{k},\tau)\tilde{\mathbf{G}}_{+,n_0}(\mathbf{r},\mathbf{k})\mathrm{e}^{-\mathrm{i}\omega_{n_0}(\mathbf{k})t}+\tilde{j}^{(j)}_{-,n_0}(\mathbf{k},\tau)\tilde{\mathbf{G}}_{-,n_0}(\mathbf{r},\mathbf{k})\mathrm{e}^{\mathrm{i}\omega_{n_0}(\mathbf{k})t}\\ \tau&=\varrho t,\ \tilde{\mathbf{J}}^{(j)}_{n}(\mathbf{r},\mathbf{k},t)=0,\ n\neq n_0,\ j=0,1;\ \tilde{\mathbf{J}}^{(j)}_{n}(\mathbf{r},\mathbf{k},t)=0\text{ for all }n\text{ if }j>2\end{aligned} \tag{186}$$

The second step in the construction of the current $\mathbf{J}$ is to pick a single quasimomentum $\mathbf{k}_*$ and to compose $\mathbf{J}$ of only the modes with quasimomenta $\mathbf{k}$ in a β-vicinity of $\pm\mathbf{k}_*$. The choice of the form of the excitation current defines a correspondence between initial data for the NLS and excitation currents for the NLM. Suppose that we are given two scalar functions $h_\pm(\mathbf{r})$ and assume that these $h_\pm(\beta\mathbf{r})$ are the initial data for the NLS such as (36), (37). In particular they satisfy the relation

$$h_-(\mathbf{r})=h^*_+(\mathbf{r}) \tag{187}$$

We define the current amplitudes $\tilde{j}^{(0)}_{\zeta,n_0}(\mathbf{k},\tau)$ in (186) by

$$\tilde{j}^{(0)}_{\zeta,n_0}(\mathbf{k},\tau)=-\varrho\beta^{-d}\psi_0(\tau)\Psi(\mathbf{k}-\zeta\mathbf{k}_*)\hat{h}_\zeta\left(\frac{\mathbf{k}-\zeta\mathbf{k}_*}{\beta}\right),\ \tau=\varrho t \tag{188}$$

The function $\hat{h}_\zeta$ is defined as the Fourier transform of the initial data h_ζ by the formula

$$\hat{h}_\zeta(\mathbf{q})=\frac{1}{(2\pi)^d}\int\mathrm{e}^{-\mathrm{i}\mathbf{q}\cdot\mathbf{r}}h_\zeta(\mathbf{r})\,\mathrm{d}\mathbf{r} \tag{189}$$

To restrict $\hat{h}_\zeta(\mathbf{q})$ from the entire $\mathbf{R}^d$ to a vicinity of $\mathbf{k}_*$ in the Brillouin zone we use a smooth cut-off function $\Psi(\boldsymbol{\eta})$, $\boldsymbol{\eta} = \mathbf{k} - \mathbf{k}_* \in \mathbf{R}^d$, satisfying the relations

$$\begin{gathered} 0 \leq \Psi(\boldsymbol{\eta}) \leq 1,\ \Psi(-\boldsymbol{\eta}) = \Psi(\boldsymbol{\eta}) \\ \Psi(\boldsymbol{\eta}) = 1 \text{ for } |\boldsymbol{\eta}| \leq \pi_0/2,\ \Psi(\boldsymbol{\eta}) = 0 \text{ for } |\boldsymbol{\eta}| \geq \pi_0 \end{gathered} \tag{190}$$

where π_0 is a sufficiently small constant which satisfies the inequalities $0 < \pi_0 < \pi/2$. In (188) $\psi_0(\tau)$ is a smooth function of the slow time τ such that

$$0 \leq \psi_0(\tau) \leq 1,\ \psi_0(\tau) = 0,\ t \leq 0 \quad \text{and} \quad t \geq \tau_0 > 0,\ \int_{-\infty}^{\infty} \psi_0(\tau) = 1 \tag{191}$$

This function provides fulfilment of (9). The normalization factor $\varrho\beta^{-d}$ in (188) makes the linear response of the media to be of order one for any choice of small ϱ, β. Notice that (189) and (187) imply that

$$\hat{h}_{-\zeta}(\mathbf{q}) = \hat{h}_\zeta(-\mathbf{q})^*,\ \zeta = \pm 1 \tag{192}$$

which together with (186) and (188) yield

$$\tilde{j}^{(0)}_{-\zeta,n}(-\mathbf{k},\tau) = \left[\tilde{j}^{(0)}_{\zeta,n}(\mathbf{k},\tau)\right]^* \tag{193}$$

In addition, (189), (187) and (145) imply that $\mathbf{J}^{(0)}(\mathbf{r},t)$ is real-valued (notice that (145) is satisfied due the condition (6)). Observe also that it follows from (186) and (188) that the current $\mathbf{J}^{(0)}(\mathbf{r},t)$ is (i) real-valued, (ii) almost time-harmonic and (iii) composed of modes from a single band n_0 and quasimomenta in a β-vicinity of $\pm\mathbf{k}_*$. We call an excitation current defined by (186)–(188) *unidirectional* since the group velocities $\nabla(-\omega_{n_0}(\mathbf{k}_*))$ and $\nabla(\omega_{n_0}(-\mathbf{k}_*))$ corresponding to both terms in (186) coincide thanks to (140).

As to the corrective current $\mathbf{J}^{(1)}$ in (186), its modal amplitudes are defined by the formula (185).

Remark In fact, our approach can be extended to excitation currents involving several $\mathbf{k}_*$ and n. For such currents the NLM generically can be reduced with high precision to several uncoupled NLSs, and we discuss this case in section 1.2. ♦

Remark Note that the magnitude of the inverse Fourier transform $h_\zeta(\beta\mathbf{r})$ of $\beta^{-d}\hat{h}_\zeta(\mathbf{q}/\beta)$ in (189) does not depend on β, implying boundedness of the maximum of its the absolute value. To obtain boundedness in a different norm one has to introduce an additional dependence on β into $h_\zeta(\beta\mathbf{r})$. For example, the integral of $|\beta^{d/2}h_\zeta(\beta\mathbf{r})|^2$ is bounded uniformly in β. ♦

2.4 *Linear response for the NLM*

It follows from (165), (166) and (186) that:

$$\begin{aligned} &\tilde{\mathbf{U}}^{(0)}(\mathbf{k},t) = \tilde{\mathbf{U}}^{(0)}_{+,n_0}(\mathbf{k},t) + \tilde{\mathbf{U}}^{(0)}_{-,n_0}(\mathbf{k},t) \\ &\tilde{\mathbf{U}}^{(0)}_{\zeta,n_0}(\mathbf{k},t) = \tilde{u}^{(0)}_{\zeta,n_0}(\mathbf{k},\tau)\tilde{\mathbf{G}}_{\zeta,n_0}(\mathbf{r},\mathbf{k})\mathrm{e}^{-\mathrm{i}\zeta\omega_{n_0}(\mathbf{k})t},\ \zeta = \pm \\ &\tilde{\mathbf{U}}^{(0)}_{\zeta,n}(\mathbf{k},t) = 0,\ n \neq n_0,\ \tau = \varrho t \end{aligned} \tag{194}$$

In addition to that, if we introduce

$$\psi(\tau) = \int_0^\tau \psi_0(\tau)\,\mathrm{d}\tau_1 \tag{195}$$

then using (166) and (188) we get

$$\tilde{u}^{(0)}_{\zeta,n_0}(\mathbf{k},\tau)=\psi(\tau)\beta^{-d}\Psi(\mathbf{k}-\zeta\mathbf{k}_*)\hat{h}_\zeta\left(\frac{\mathbf{k}-\zeta\mathbf{k}_*}{\beta}\right),\ \tau=\varrho t \tag{196}$$

and, in particular,

$$\tilde{u}^{(0)}_{\zeta,n}(\mathbf{k},\tau)=0 \text{ if either } |\mathbf{k}-\zeta\mathbf{k}_*|\geq\pi_0,\quad \text{or}\quad \tau=\varrho t\leq 0,\quad \text{or } n\neq n_0 \tag{197}$$

By (165) and (196)

$$\tilde{U}^{(0)}_{\bar{n}_0}(\mathbf{k},t)=\tilde{u}^{(0)}_{\zeta,n_0}(\mathbf{k},\tau)\mathrm{e}^{-\mathrm{i}\zeta\omega_{n_0}(\mathbf{k})t}=\psi(\tau)\Psi(\mathbf{k}-\zeta\mathbf{k}_*)\beta^{-d}\hat{h}_\zeta\left(\frac{\mathbf{k}-\zeta\mathbf{k}_*}{\beta}\right)\mathrm{e}^{-\mathrm{i}\zeta\omega_{n_0}(\mathbf{k})t} \tag{198}$$

Remark It follows from (194), (196), (197) and (198) that for the chosen currents the linear medium response $\mathbf{U}^{(0)}$ is composed of only the modes from a single band n_0 corresponding to the carrier wave frequency $\omega_{n_0}(\mathbf{k})$ together with the opposite band corresponding to $-\omega_{n_0}(\mathbf{k})$ with wave numbers in a β-vicinity of respectively two wavevectors $\pm\mathbf{k}_*$. We call such an excitation in a vicinity of $\pm\omega_{n_0}(\pm\mathbf{k})$ a *doublet*, see Figure 3. Note that the group velocities of the two components of $\mathbf{U}^{(0)}$ corresponding to the points $\zeta\mathbf{k}_*$, $\zeta=\pm1$ and the bands $\zeta\omega_{n_0}(\mathbf{k})$ are equal to $\zeta\nabla\omega_{n_0}(\zeta\mathbf{k}_*)$ for $\zeta=\pm1$, and these group velocities are the same in view of (140). Hence, a doublet is a *unidirectional* excitation. Consequently, if β is small $\mathbf{U}^{(0)}$ is a real-valued almost time-harmonic wavepacket propagating in the direction of $\nabla\omega_{n_0}(\mathbf{k}_*)$ (see [18] for a discussion of the group velocity of wavepackets in photonic crystals). ⧫

The formula (196) suggests introduction of a local variable $\boldsymbol{\eta}$ and its scaled version $\mathbf{q}$ at $\zeta\mathbf{k}_*$ by the following formulas

$$\boldsymbol{\eta}=\beta\mathbf{q}=\mathbf{k}-\zeta\mathbf{k}_* \tag{199}$$

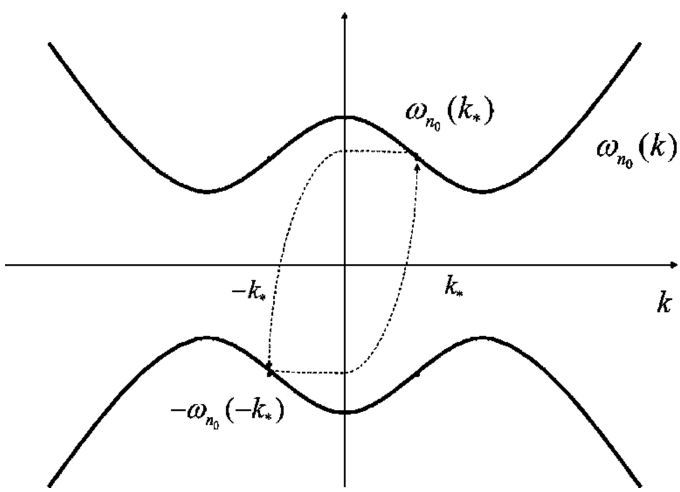

Figure 3. A real-valued excitation current based on the band number n_0 and the quasimomentum k_* directly excites a pair of modes with quasimomenta ζk_* and frequencies $\zeta\omega_{n_0}(\zeta k_*)$, $\zeta=\pm$, forming a doublet. The modes in the doublet have a strong nonlinear interaction.

Let us approximate $\omega_{n_0}(\mathbf{k}_*+\boldsymbol{\eta})$ in a vicinity of $\eta=0$ by its Taylor polynomial $\gamma_{(\nu)}(\boldsymbol{\eta})$ of the degree ν (see sections 8.1 and 8.2 for notation)

$$\gamma_{(\nu)}(\boldsymbol{\eta})=\gamma_{(\nu)}(\mathbf{k}_*;\boldsymbol{\eta})=\sum_{j=0}^{\nu}\frac{1}{j!}\omega_{n_0}^{(j)}(\mathbf{k}_*)(\boldsymbol{\eta}^j) \tag{200}$$

In particular, for $\nu=2$

$$\gamma_{(2)}(\boldsymbol{\eta})=\omega_{n_0}(\mathbf{k}_*)+\omega'_{n_0}(\mathbf{k}_*)(\boldsymbol{\eta})+\frac{1}{2}\omega''_{n_0}(\mathbf{k}_*)(\boldsymbol{\eta}^2) \tag{201}$$

where $\omega'_{n_0}(\mathbf{k}_*)$ and $\omega''_{n_0}(\mathbf{k}_*)$ are respectively linear and quadratic forms, i.e. a vector and a matrix, i.e.

$$\omega'_{n_0}(\mathbf{k}_*)(\eta)=\omega'_{n_0}(\mathbf{k}_*)\cdot\eta,\ \omega''_{n_0}(\mathbf{k}_*)(\boldsymbol{\eta}^2)=\boldsymbol{\eta}\cdot\omega''_{n_0}(\mathbf{k}_*)\boldsymbol{\eta} \tag{202}$$

Note that (2), (140) and (200) imply that $\omega_{n_0}(\zeta\mathbf{k}_*+\boldsymbol{\eta})=\omega_{n_0}(\mathbf{k}_*+\zeta\eta)$ and a similar property for its Taylor polynomial, namely

$$\gamma_{(\nu)}(\zeta\mathbf{k}_*;\boldsymbol{\eta})=\gamma_{(\nu)}(\mathbf{k}_*;\zeta\boldsymbol{\eta})=\sum_{j=0}^{\nu}\frac{1}{j!}\omega_{n_0}^{(j)}(\mathbf{k}_*)(\zeta\boldsymbol{\eta}^j) \tag{203}$$

The following Taylor remainder estimation holds

$$|\omega_{n_0}(\zeta\mathbf{k}_*+\beta\mathbf{q})-\gamma_{(\nu)}(\zeta\beta\mathbf{q})|\le C\beta^{\nu+1}|\mathbf{q}|^{\nu+1},\ \mathbf{q}\in\mathbf{R}^d \tag{204}$$

After the change of variables (199) the linear response takes the form

$$\tilde{U}^{(0)}_{\zeta,n_0}(\zeta\mathbf{k}_*+\beta\mathbf{q},t)=\psi(\tau)\Psi(\beta\mathbf{q})\beta^{-d}\hat{h}_\zeta(\mathbf{q})\mathrm{e}^{-\mathrm{i}\zeta\gamma_{(\nu)}(\zeta\beta\mathbf{q})t},\ \tau=\varrho t \tag{205}$$

with ψ, Ψ and $\hat{h}_\zeta$ defined by (191), (190) and (189). Notice that for $\nu=2$ the phase function $\gamma_{(2)}(\boldsymbol{\eta})$ in (205) is a quadratic polynomial which is identical to the phase function of the linear response of the relevant linear Schrödinger equation (see (299) and (300)).

Let us consider now the properties of the functions $\hat{h}_\zeta(\mathbf{q})$. We take two smooth functions $\hat{h}_\zeta(\mathbf{q})$, $\zeta=\pm$, defined for all $q\in\mathbf{R}^d$ that decay for large $|\mathbf{q}|$ faster than any negative power, i.e.

$$|\hat{h}_\zeta(\mathbf{q})|\le C_{N_\Psi}(1+|\mathbf{q}|)^{-N_\Psi},\ \zeta=\pm,\ \text{with arbitrarily large } N_\Psi>0 \tag{206}$$

The condition (206) is used primarily to show that as $\beta\to 0$ the function Ψ in (188) does not affect the asymptotic expansions we derive below. Notice that the wavevectors (quasimomenta) $\mathbf{k}$ needed to compose a solution to the NLS via its Fourier transform vary over the entire space $\mathbf{R}^d$, whereas for the NLM we use quasimomenta from the Brillouin zone. Since $\Psi(\beta\mathbf{q})\hat{h}_\zeta(\mathbf{q})=0$ when $|\beta\mathbf{q}|\ge\pi_0$ we can consider the right-hand side of (205) as a regular function in the Brillouin zone. Using the above notation we rewrite (205) in the form

$$\begin{gathered}\tilde{U}^{(0)}_{\bar{n}}(\zeta\mathbf{k}_*+\beta\mathbf{q},t)=\tilde{u}^{(0)}_{\bar{n}}(\zeta\mathbf{k}_*+\beta\mathbf{q},\tau)\mathrm{e}^{-\mathrm{i}\zeta\gamma_{(\nu)}(\zeta\beta\mathbf{q})t},\ \tau=\varrho t\\ \tilde{u}^{(0)}_{\bar{n}}(\zeta\mathbf{k}_*+\beta\mathbf{q},\tau)=\psi(\tau)\Psi(\beta\mathbf{q})\beta^{-d}\hat{h}_\zeta(\mathbf{q}),\ \bar{n}=(\zeta,n_0)\end{gathered} \tag{207}$$

Notice that by (190), $\Psi(\beta\mathbf{q})=1$ for $|\beta\mathbf{q}|\le\pi_0$, and it follows from (206) that

$$|\Psi(\beta\mathbf{q})\hat{h}_\zeta(\mathbf{q})-\hat{h}_\zeta(\mathbf{q})|\le C'_{N_\Psi}\beta^{N_\Psi},\ |\mathbf{q}|\ge\frac{\pi_0}{2\beta},\ \text{with arbitrarily large } N_\Psi>0 \tag{208}$$

Consequently, we have the representations for all $\mathbf{q}$

$$\tilde{u}^{(0)}_{\bar{n}}(\zeta\mathbf{k}_*+\beta\mathbf{q},\tau)=\psi(\tau)\hat{h}_\zeta(\mathbf{q})+O(\beta^{N_\Psi})O(|\hat{h}|) \tag{209}$$

where $N_\Psi>0$ can be arbitrarily large.

3. Asymptotic expansions for the FNLR for the Maxwell equations

From (176), (194) and (196) it follows that the FNLR for every $\bar{n} = (\zeta, n)$ is a sum of only eight non-zero terms:

$$\tilde{u}_{\bar{n}}^{(1,0)}(\mathbf{k}, \tau) = \sum_{\zeta', \zeta'', \zeta'''} I_{\bar{n}, \zeta', \zeta'', \zeta'''}(\mathbf{k}, \tau) - \tilde{u}_{\bar{n}}^{(1)}(\mathbf{J}_1; \mathbf{k}, \tau) \tag{210}$$

where the interaction integrals $I_{\bar{n}, \zeta', \zeta'', \zeta'''}$ have the following representations

$$I_{\bar{n}, \zeta', \zeta'', \zeta'''}(\mathbf{k}, \tau) = \frac{1}{\varrho} \int_0^\tau \int_{\substack{[-\pi, \pi]^{2d} \\ \mathbf{k}'+\mathbf{k}''+\mathbf{k}''' = \mathbf{k}}} \exp\left\{ \mathrm{i}\phi_{\vec{n}}(\vec{k}) \frac{\tau_1}{\varrho} \right\} \mathcal{A}_{\vec{n}}(\vec{k}, \tau_1) \, \mathrm{d}\mathbf{k}' \mathrm{d}\mathbf{k}'' \mathrm{d}\tau_1 \tag{211}$$

with

$$\mathcal{A}_{\vec{n}}(\vec{k}, \tau) = \breve{Q}_{\vec{n}}(\vec{k}) \tilde{u}_{\zeta', n_0}^{(0)}(\mathbf{k}', \tau) \tilde{u}_{\zeta'', n_0}^{(0)}(\mathbf{k}'', \tau) \tilde{u}_{\zeta''', n_0}^{(0)}(\mathbf{k}''', \tau) \tag{212}$$

and $\vec{k}$ defined by (178). The term $\tilde{u}_{\bar{n}}^{(1)}(\mathbf{J}_1; \mathbf{k}, \tau)$ is defined by (180). Note that indices $\vec{n}$ involved in the representation (211) satisfy the relation

$$n' = n'' = n''' = n_0, \;\; \vec{n} = ((\zeta, n), (\zeta', n_0), (\zeta'', n_0), (\zeta''', n_0)) \tag{213}$$

Observe also that for $n \neq n_0$ the integral (211) describes the nonlinear impact on the indirectly excited modes. It can be non-zero, though, as we discussed in section 1.2 and show later in section 3.2, it is small compared to $n = n_0$ since it is not frequency matched.

Since the tensors $\mathbf{R}_D^{(3)}$ in (153) are symmetric, the coefficients $\breve{Q}_{\vec{n}}(\mathbf{k}, \mathbf{k}', \mathbf{k}'', \mathbf{k}''')$ are symmetric with respect to the interchange of $(\zeta^{(i)}, \mathbf{k}^{(i)})$ and $(\zeta^{(l)}, \mathbf{k}^{(l)})$ if relations (213) hold. Consequently, we have

$$I_{\bar{n}, \zeta', \zeta'', \zeta'''} = I_{\bar{n}, \zeta'', \zeta', \zeta'''} = I_{\bar{n}, \zeta'', \zeta''', \zeta'} \tag{214}$$

It follows from (197) that the integrands in the right-hand side of (211) are non-zero only when

$$|\mathbf{k}' - \zeta' \mathbf{k}_*| \leq \pi_0, \;\; |\mathbf{k}'' - \zeta'' \mathbf{k}_*| \leq \pi_0, \;\; |\mathbf{k}''' - \zeta''' \mathbf{k}_*| \leq \pi_0 \tag{215}$$

Observe that since the very form of the integral $I_{\bar{n}, \zeta', \zeta'', \zeta'''}$ (211) obeys the phase matching condition (71) through its domain of integration, the *four-wave interactions may occur only if*

$$|\zeta' \mathbf{k}_* + \zeta'' \mathbf{k}_* + \zeta''' \mathbf{k}_* - \mathbf{k}| \leq 3\pi_0 \tag{216}$$

We assume that $\mathbf{k}_*$ is a generic point in the following sense.

Genericity condition. A point $\mathbf{k}_*$ is called *generic* if it satisfies the relations

$$3\mathbf{k}_* \neq \mathbf{k}_* (\mathrm{mod}\; 2\pi); \tag{217}$$

$$|3\omega_{n_0}(\mathbf{k}_*) - \omega_n(3\mathbf{k}_*)| \neq 0, \;\; |\omega'_{n_0}(\mathbf{k}_*) \pm \omega'_n(3\mathbf{k}_*)| \neq 0, \;\; n = 1, 2, \ldots$$

$$|\omega_{n_0}(\mathbf{k}_*) - \omega_n(\mathbf{k}_*)| \neq 0, \;\; |\omega'_{n_0}(\mathbf{k}_*) \pm \omega'_n(\mathbf{k}_*)| \neq 0, \;\; n \neq n_0 \tag{218}$$

According to (188), we compose the currents $\mathbf{J}$ from eigenmodes $\{(\zeta, n_0), \mathbf{k}\}$ satisfying the following condition

$$|\mathbf{k} - \zeta \mathbf{k}_*| \leq \pi_0, \zeta = \pm 1 \tag{219}$$

where π_0 is small constant. In fact when $\beta \to 0$, π_0 in (219) can be replaced for almost time-harmonic waves (188) with even a smaller number $\beta\pi_0$.

We call a mode $(n, \mathbf{k})$ *indirectly excited* if

$$\tilde{u}^{(0)}_{\zeta,n}(\mathbf{k}, \tau) = 0 \text{ for all } \tau \quad \text{and} \quad \zeta = \pm \tag{220}$$

According to (197), the modes with $n \neq n_0$ or $|\mathbf{k} - \zeta \mathbf{k}_*| > \pi_0$ are indirectly excited. All other modes are called *directly excited*, obviously directly excited modes must satisfy (219). *In other words, directly excited modes are excited through the linear mechanism whereas indirectly excited ones are excited only through the nonlinear mechanism*. Hence, based on the medium linear and the FNLRs all the eigenmodes labelled with $\{(\zeta, n), \mathbf{k}\}$ can be naturally partitioned into two classes: the eigenmodes that are involved in the composition of the probing excitation current and eigenmodes that are not. The first class coincides with the directly excited modes and the second with the indirectly excited. The linear response of the medium obviously involves only the eigenmodes presented in the source (current), i.e. ones satisfying the condition (219), and nothing else. If we look now at the FNLR we find that eigenmodes which do not satisfy the condition (219) generically are also presented in its composition though with much smaller amplitudes.

As was shown in [17–19] stronger interactions must satisfy the group velocity matching (73) and the frequency matching (74) conditions. It follows from (2) that

$$(\nabla \omega_n)(\zeta \mathbf{k}_*) = \zeta \nabla \omega_n(\mathbf{k}_*), \ \zeta = \pm 1 \tag{221}$$

Since (213) holds, the group velocity matching condition (73) at the points $\zeta^{(i)}\mathbf{k}_*$ takes the form

$$\nabla[\zeta' \omega_{n_0}(\zeta' \mathbf{k}_*)] = \nabla[\zeta'' \omega_{n_0}(\zeta'' \mathbf{k}_*)], \nabla[\zeta' \omega_{n_0}(\zeta' \mathbf{k}_*)] = \nabla[\zeta''' \omega_{n_0}(\zeta''' \mathbf{k}_*)] \tag{222}$$

and by (221) it is always fulfilled. The *frequency matching (FM) condition* (74) can be written in the form

$$\phi_{\vec{n}}(\vec{k}) = \zeta \omega_n(\mathbf{k}) - \zeta' \omega_{n_0}(\mathbf{k}') - \zeta'' \omega_{n_0}(\mathbf{k}'') - \zeta''' \omega_{n_0}(\mathbf{k}''') = 0 \tag{223}$$

where (213) is assumed. Rather often fulfilment of the equality (223) is called the *phase matching* condition (see [12, 37]), but we prefer to call it the *frequency matching* condition and reserve the term phase matching condition for the condition (71)).

At the points $\zeta^{(i)}\mathbf{k}_*$ according to (71) the relations (74) and (223) take the form

$$\zeta \omega_n(\zeta' \mathbf{k}_* + \zeta'' \mathbf{k}_* + \zeta''' \mathbf{k}_*) = \zeta' \omega_{n_0}(\zeta' \mathbf{k}_*) + \zeta'' \omega_{n_0}(\zeta'' \mathbf{k}_*) + \zeta''' \omega_{n_0}(\zeta''' \mathbf{k}_*) \tag{224}$$

Note now that the sum $\zeta' + \zeta'' + \zeta'''$ equals either ± 1 or ± 3. From (218) we obtain that $\zeta' + \zeta'' + \zeta'''$ cannot be ± 3, and, hence,

$$n = n_0, \ \zeta' \mathbf{k}_* + \zeta'' \mathbf{k}_* + \zeta''' \mathbf{k}_* = \pm \mathbf{k}_* \tag{225}$$

The inequality (216) implies that for frequency matched interactions we have

$$|\mathbf{k} - \zeta \mathbf{k}_*| \leq 3\pi_0 \tag{226}$$

Finally, condition (224) together with (2) and (218) imply that

$$n = n' = n'' = n''' = n_0, \ \zeta' + \zeta'' + \zeta''' = \zeta \tag{227}$$

We will refer to a situation when the multi-index $\vec{n}$ satisfies the relation (227) as the *frequency matched case (FM case)*, and to a situation when at least one of the relations (227) does not hold as to *non-frequency-matched case (non-FM case). Observe that for frequency matched interactions, i.e. for the FM case, all significant mode interactions are restricted to a single band $n = n_0$*. It is convenient to introduce the interaction phase $\phi_{\vec{n}}(\vec{k})$ for the special situation of the FM case for which $\zeta''' = -\zeta$, $\zeta' = \zeta'' = \zeta$, namely when

$$\vec{n} = \vec{n}_0 = ((\zeta, n_0), (\zeta, n_0), (\zeta, n_0), (-\zeta, n_0)) \tag{228}$$

The phase interaction function takes the following form

$$\phi_{\vec{n}_0}(\vec{k}) = \zeta\left[\omega_{n_0}(\mathbf{k}) - \omega_{n_0}(\mathbf{k}') - \omega_{n_0}(\mathbf{k}'') + \omega_{n_0}(\mathbf{k}''')\right] \tag{229}$$

We would like to remark that it turns out that the interaction integrals (211) in the non-FM case are much smaller than in the FM case. *Consequently, more significant nonlinear mode interactions are expected to be frequency matched*. The non-FM and FM cases will be discussed in detail in the following two subsections.

3.1 *Frequency matched interactions*

In this section we consider the interaction integrals $I_{\vec{n},\zeta',\zeta'',\zeta'''}$ in the frequency matched (FM) case, i.e. if relations (227) are fulfilled. In section 3.2 we consider the same interaction integrals in the non-FM case, i.e. when (227) does not hold. Comparing both cases we will see in particular that the interactions in the FM case are stronger than those in the non-FM case.

Assuming that the relations (227) hold we introduce the following change of variables

$$\mathbf{k} - \zeta\mathbf{k}_* = \beta\mathbf{q},\ k' - \zeta'\mathbf{k}_* = \beta\mathbf{q}',\mathbf{k}''-\zeta''\mathbf{k}_* = \beta\mathbf{q}'',\ k''' - \zeta'''\mathbf{k}_* = \beta\mathbf{q}''' \tag{230}$$

or in shorter notation

$$\vec{k} = \vec{\zeta}\vec{k}_* + \beta\vec{q} \tag{231}$$

where we use notation (177) and

$$\vec{q} = (\mathbf{q}, \mathbf{q}', \mathbf{q}'', \mathbf{q}'''),\ \vec{\zeta}\vec{q} = (\zeta\mathbf{q}, \zeta'\mathbf{q}', \zeta''\mathbf{q}'', \zeta'''\mathbf{q}''') \tag{232}$$

Note that if (227) holds, the following two equalities are equivalent:

$$\mathbf{k}' + \mathbf{k}'' + \mathbf{k}''' = \mathbf{k} \text{ is equivalent to } \mathbf{q}' + \mathbf{q}'' + \mathbf{q}''' = \mathbf{q} \tag{233}$$

Obviously, there are three combinations of $\zeta', \zeta'', \zeta'''$ satisfying (227) with $\zeta = 1$ and three more combinations with $\zeta = -1$. It follows from (227) that two of the numbers $\zeta', \zeta'', \zeta'''$ have to coincide with ζ, and the third one equals $-\zeta$. Let us fix ζ and assume that $\zeta''' = -\zeta$, $\zeta' = \zeta'' = \zeta$, and denote

$$\vec{\zeta}_0 = (\zeta, \zeta, \zeta, -\zeta) \tag{234}$$

Two other integrals with $\zeta' = -\zeta$, $\zeta'' = -\zeta$ can be reduced to the above case with the help of the equalities (214), namely

$$I_{\vec{n},\zeta,\zeta,-\zeta} = I_{\vec{n},\zeta,-\zeta,\zeta} = I_{\vec{n},-\zeta,\zeta,\zeta} \tag{235}$$

Using (210) and (211) together with (235) and taking into account the estimates of non-FM terms provided in the section 3.2 we obtain the following important representation related to the FNLR

$$\tilde{u}_{\vec{n}}^{(1,0)}(\mathbf{k}, \tau) = 3I_{\vec{n},\zeta,\zeta,-\zeta}(\mathbf{k}, \tau) - \tilde{u}_{\vec{n}}^{(1)}(\mathbf{J}_1; \mathbf{k}, \tau) + O(\varrho)O\left(\mathbf{u}^{(1)}\right) \tag{236}$$

with the integral $I_{\vec{n},\zeta,\zeta,-\zeta}(\mathbf{k}, \tau)$ given by

$$\begin{aligned} &I_{\vec{n},\zeta,\zeta,-\zeta}(\mathbf{k}, \tau) \\ &= \frac{1}{\varrho}\int_0^\tau \int_{\substack{[-\pi,\pi]^{2d} \\ \mathbf{k}' + \mathbf{k}'' + \mathbf{k}''' = \mathbf{k}}} \exp\left\{\mathrm{i}\phi_{\vec{n}_0}(\vec{k})\frac{\tau_1}{\varrho}\right\} \mathcal{A}_{\vec{n}_0}(\vec{k}, \tau_1)\, \mathrm{d}\mathbf{k}'\mathrm{d}\mathbf{k}''\mathrm{d}\tau_1 \end{aligned} \tag{237}$$

where $\mathcal{A}_{\vec{n}}$, $\tilde{u}_{\zeta,n_0}^{(0)}$ and $\phi_{\vec{n}_0}$ are defined respectively by (212), (196) and (229), and $\vec{n}_0$ is defined by (228).

Let us substitute for the factors $\tilde{u}^{(1)}_{\zeta,n_0}$ in (212) their expressions in terms of the currents $\tilde{j}^{(0)}_{n_0,\zeta}(\mathbf{k},\tau)$ and, consequently, the initial data $\hat{h}_\zeta$. The equalities (237), (188) and (180) together with (230), (232) and (234) yield the following expression for the interaction integral in terms of $\hat{h}_\zeta$:

$$\beta^d I_{\bar{n},\zeta,\zeta,-\zeta}(\zeta\mathbf{k}_* + \beta\mathbf{q},\tau) = \tfrac{1}{\varrho}\int_0^\tau \int_{\mathbf{q}'+\mathbf{q}''+\mathbf{q}'''=\mathbf{q}}$$

$$\exp\left\{\mathrm{i}\phi_{\vec{n}_0}(\vec{\zeta}\vec{k}_* + \beta\vec{q})\tfrac{\tau_1}{\varrho}\right\}\psi^3(\tau_1)A_1(\beta\vec{q}^0)\hat{h}_\zeta(\mathbf{q}')\hat{h}_\zeta(\mathbf{q}'')\hat{h}_{-\zeta}(\mathbf{q}''')\,\mathrm{d}\mathbf{q}'\,\mathrm{d}\mathbf{q}''\mathrm{d}\tau_1 \tag{238}$$

where

$$A_1(\beta\vec{q}) = \breve{Q}_{\vec{n}_0}((\vec{\zeta}_0\mathbf{k}_* + \beta\vec{q}))\Psi^3(\beta\vec{q}) \tag{239}$$

$$\Psi^3(\beta\vec{q}) = \Psi(\beta\mathbf{q}')\Psi(\beta\mathbf{q}'')\Psi(\beta\mathbf{q}''') \tag{240}$$

$\vec{n}_0$ is given by (228), $\hat{h}_\zeta(\mathbf{q})$ is the same as in (189) and $\breve{Q}_{\vec{n}_0}$ is defined by (179). Notice that the domain of integration of the integral (238) allows $\mathbf{q}'$ and $\mathbf{q}''$ to vary over the entire space $\mathbb{R}^d$ rather than restricting them to just $[-\pi,\pi]^d$, which can be done since the function $\Psi^3(\beta\vec{q})$ by its definition (240) and (190) is zero if either $\mathbf{q}'$ or $\mathbf{q}''$ is outside of $[-\pi,\pi]^d$ for $\beta \le 1$. Recall that the function $\Psi(\mathbf{k})$ was introduced to do exactly that in order to resolve the difference in setting of the NLS with the quasimomentum $\mathbf{k}$ varying in the entire space $\mathbb{R}^d$ and the NLM for periodic medium with $\mathbf{k}$ varying in $[-\pi,\pi]^d$.

Remark The Floquet–Bloch representation of a single-mode function

$$\mathbf{J}_n(\mathbf{r},t) = \frac{1}{(2\pi)^d}\int_{[-\pi,\pi]^d}\tilde{j}_n(\mathbf{k})\tilde{\mathbf{G}}_n(\mathbf{r},\mathbf{k})\,\mathrm{d}\mathbf{k} \tag{241}$$

with the coefficient $\tilde{j}_n(\mathbf{k})$ satisfying $\tilde{j}_n(\mathbf{k}) = 0$ for $|\mathbf{k}-\mathbf{k}_*| \ge \pi_0$ can be rewritten in the form

$$\mathbf{J}_n(\mathbf{r},t) = \beta^d\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\tilde{j}_n(\mathbf{k}_*+\beta\mathbf{q})\tilde{\mathbf{G}}_n(\mathbf{r},\mathbf{k}_*+\beta\mathbf{q})\,\mathrm{d}\mathbf{q} \tag{242}$$

This identity shows that the coefficient $\tilde{j}_n(\mathbf{k}_*+\beta\mathbf{q})$ has to have the scaling factor β^d to determine a function $\mathbf{J}_n(\mathbf{r},t)$ which is bounded uniformly in β. Exactly this kind of expression is written in the left-hand side of (238). ♦

3.2 *Non-frequency-matched interactions*

There are two different possibilities for the non-frequency-matched (non-FM) case, i.e. when (213) holds and (224) does not, which are described by the following two alternatives:

$$\zeta' + \zeta'' + \zeta''' = -\zeta,\ n = n_0 \tag{243}$$

or

$$\zeta' + \zeta'' + \zeta''' = 3\zeta \quad \text{or} \quad \zeta' + \zeta'' + \zeta''' = -3\zeta \quad \text{or} \quad n \ne n_0. \tag{244}$$

In the case (244) when $\mathbf{k} = \pm 3\mathbf{k}_* + \beta\mathbf{q}$ or $n \ne n_0$ (225) does not hold, the FNLR is nonzero, but the linear response is zero. In the case when (243) holds according to (216) the FNLR with $n = n_0$, $|-\zeta\mathbf{k}_* - \mathbf{k}| \le 3\pi_0$ is nonzero, but by (197) the linear response $\tilde{u}^{(0)}_{\zeta,n}(\mathbf{k},\tau)$ is zero when the linearly excited wave is unidirectional. Using (371) we obtain

$$\tilde{u}^{(1)}_{\bar{n}}(\mathbf{k},t) = \tilde{u}^{(1,0)}_{\bar{n}}(\mathbf{k},t) + O(\varrho)O\left(\mathbf{u}^{(1)}\right),\ \frac{\tau_0}{\varrho} \le t \le \frac{\tau_*}{\varrho} \tag{245}$$

where $\tilde{u}_{\bar{n}}^{(1,0)}$ is the time-harmonic approximation for $\tilde{u}_{\bar{n}}^{(1)}$ defined by (210), (211) and, in view of (181), we get

$$\tilde{u}_{\bar{n}}^{(1,0)} = \tilde{F}_{\bar{n}}\left[\left(\mathbf{u}^{(0)}\right)^3\right](\mathbf{k},\tau) \tag{246}$$

Let us show that for the both non-FM cases (243) and (244) the FNLR $\tilde{u}_{\bar{n}}^{(1)}(\mathbf{k},t)$ is much smaller than it its counterpart for the FM case, and the following estimation holds:

$$\tilde{u}_{\bar{n}}^{(1)}(\mathbf{k},t) = O(\varrho)O\left(\left|\mathbf{U}^{(1)}\right|\right),\ \frac{\tau_0}{\varrho} \le t \le \frac{\tau_*}{\varrho} \tag{247}$$

If the estimation (247) holds then the FNLR $\tilde{u}_{\bar{n}}^{(1)}(\mathbf{k},t)$ in the non-FM case evidently is smaller by the factor ϱ than the FNLR for the FM case in (238) which is of order $O(|\mathbf{U}^{(1)}|)$. Let us look first at how the non-FM condition affects the magnitude of the interactions as described by the interaction integral $I_{\bar{n},\zeta',\zeta'',\zeta'''}$ in (211). Observe that if relation (227) does not hold, the FM condition (224) for the phase $\phi_{\bar{n}}(\vec{k})$ does not hold either. Using this fact and integrating the interaction integral (211) by parts m_2+1 times as in [17, 20] we obtain

$$I_{\bar{n},\zeta',\zeta'',\zeta'''}(\mathbf{k},\tau) = \frac{1}{\varrho}\sum_{m=1}^{m_2}\varrho^m K_m(\mathbf{k},\tau) + O(\varrho^{m_2}) \tag{248}$$

The integral

$$K_m(\mathbf{k},\tau) = \int_{\substack{[-\pi,\pi]^{2d} \\ \mathbf{k}'+\mathbf{k}''+\mathbf{k}'''=\mathbf{k}}} \exp\left\{\mathrm{i}\phi_{\bar{n}}(\vec{k})\frac{\tau}{\varrho}\right\}\mathcal{A}_{\bar{n}m}(\vec{k},\tau)\,\mathrm{d}\mathbf{k}'\,\mathrm{d}\mathbf{k}'' \tag{249}$$

is similar to the integral in (238), but since it has a factor ϱ^m in (248) it may affect only the higher order approximations. Note that since the operators K_m in (248) do not involve integration with respect to τ they produce expressions in the relevant extended NLS involving the time derivatives. The dominant term K_1 is given by

$$K_1(\mathbf{k},\tau) = \int_{\substack{[-\pi,\pi]^{2d} \\ \mathbf{k}'+\mathbf{k}''+\mathbf{k}'''=\mathbf{k}}} \frac{\mathrm{e}^{\mathrm{i}\phi_{\bar{n}}(\vec{k})\frac{\tau}{\varrho}}}{\mathrm{i}\phi_{\bar{n}}(\vec{k})}\mathcal{A}_{\bar{n}}(\vec{k},\tau)\,d\mathbf{k}'d\mathbf{k}'' \tag{250}$$

The integral with respect to $\mathrm{d}\mathbf{k}'\,\mathrm{d}\mathbf{k}''$ is similar to $I_{\bar{n},\zeta,\zeta,-\zeta}$ but without the factor $1/\varrho$, therefore $K_1(\mathbf{k},\tau)$ is of order ϱ times (238) that is $O(1)$.

3.2.1 Approximation of indirectly excited modes. For unidirectional excitation currents the indirectly excited modes $((\zeta,n_0),\mathbf{k})$, i.e. the ones not satisfying relation (219), have zero linear response, i.e. $\tilde{u}_{\bar{n}}^{(0)}(\mathbf{k},t)=0$, and in both cases (243) or (244) the principal part of the corresponding amplitudes $\tilde{u}_{\bar{n}}(\mathbf{k},t)$ is given by (245). In particular, relation (247) holds, in view of (248). For given $h_\pm$, for the indirectly excited modes when $n\neq n_0$ or $n=n_0$ and $|\mathbf{k}-\mathbf{k}_*|\geq \pi_0$ we set

$$\tilde{U}_{Z,\bar{n}}(\mathbf{k},t) = \alpha\mathrm{e}^{-\mathrm{i}\omega_{\bar{n}}(\mathbf{k})t}\tilde{u}_{\bar{n}}^{(1,0)}(\mathbf{k},t) \tag{251}$$

with the FNLR amplitudes $\tilde{u}_{\bar{n}}^{(1,0)}$ being defined by (210) and (211). Then using (246) we can recast $\tilde{U}_{Z,\bar{n}}(\mathbf{k},t)$ in the form

$$\tilde{U}_{Z,\bar{n}}(\mathbf{k},t) = \alpha\mathrm{e}^{-\mathrm{i}\omega_{\bar{n}}(\mathbf{k})(\frac{\tau}{\varrho}t)}\bar{u}_{Z,\bar{n}}(\mathbf{k},t),\ \tilde{u}_{Z,\bar{n}}(\mathbf{k},t) = \tilde{F}_{\bar{n}}\left[\left(\mathbf{u}^{(0)}\right)^3\right](\mathbf{k},\tau) \tag{252}$$

with formula (245) providing an estimate for the difference between the modal coefficient of the exact solution of the NLM and the approximation $\bar{u}_{Z,\bar{n}}(\mathbf{k},t)$.

The above discussion shows that the amplitudes of indirectly excited modes are determined mainly by the FNLR. The amplitudes of indirectly excited modes are relatively small and are of the order $O(\varrho\alpha)O(\mathbf{U}^{(1)})$.

4. Asymptotic analysis of the FNLR

In this section we write expansions of the FNLR of the NLM in β up to the order $\sigma = \nu - 2$ with given ν using Taylor expansions of the integrands of the interaction integral (238). More delicate are expansions involving the phase function. The corresponding asymptotic expansions in the weakly dispersive case involve powers of $\beta^{\nu+1}/\varrho$; in the strongly dispersive case (28) we use expansions with respect to the parameter $\theta = \varrho/\beta^2$. Then we analyse the nonlinear interaction integrals and the corresponding interaction phases $\phi_{\vec{n}}(\vec{k})$ and $\phi_{\vec{n}_0}(\vec{k})$ defined by (223) and (229) and obtain relatively simple asymptotic expansions for the modal form of the FNLR. Those modal expansions can be directly related to the FNLR of a proper NLS and, at the same time, provide a basis for estimating the difference between a solution to the NLM and its nonlinear Schrödinger approximation.

In the weakly dispersive case, when (30) holds, the asymptotic analysis of a solution to the NLM is carried out straightforwardly based on the Taylor expansion of the oscillating factor. In the strongly dispersive case, when (28) holds, we apply the stationary phase method.

4.1 *Interaction integrals*

With the interaction integral $I_{\vec{n},\zeta',\zeta'',\zeta'''}$ defined by (211) with

$$\vec{\zeta} = (\zeta, \zeta', \zeta'', \zeta''') = \vec{\zeta}_0 = (\zeta, \zeta, \zeta, -\zeta) \quad \text{and} \quad n = n_0 \tag{253}$$

as in (234), (237) turns into the integral $I_{\vec{n},\zeta,\zeta,-\zeta}$ defined by (238). To study the integral (238) we approximate the phase function $\phi_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_* + \beta\vec{q})$ defined by (229) by a polynomial phase function $\Phi^{(\nu)}$ of degree ν (for instance, quadratic when $\nu = 2$):

$$\Phi^{(\nu)}(\vec{\zeta}_0, \beta\vec{q}) = \zeta\left[\gamma_{(\nu)}(\zeta\beta\mathbf{q}) - \gamma_{(\nu)}(\zeta\beta\mathbf{q}') - \gamma_{(\nu)}(\zeta\beta\mathbf{q}'') + \gamma_{(\nu)}(-\zeta\beta\mathbf{q}''')\right] \tag{254}$$

The polynomial $\gamma_{(\nu)}$ is defined by (200) and (203) based on $\omega_{n_0}(\mathbf{k})$. Let us consider now the phase function (254) under the constraint (233) and denote

$$\mathbf{q}'''(\vec{q}) = \mathbf{q} - \mathbf{q}' - \mathbf{q}'', \vec{q}^{\,0} = (\mathbf{q}, \mathbf{q}', \mathbf{q}'', \mathbf{q} - \mathbf{q}' - \mathbf{q}'') \tag{255}$$

that is

$$\Phi^{(\nu)}(\vec{\zeta}_0, \beta\vec{q}) = \zeta[\gamma_{(\nu)}(\zeta\beta\mathbf{q}) - \gamma_{(\nu)}(\zeta\beta\mathbf{q}') - \gamma_{(\nu)}(\zeta\beta\mathbf{q}'') + \gamma_{(\nu)}(-\zeta\beta(\mathbf{q} - \mathbf{q}' - \mathbf{q}''))] \tag{256}$$

An important property of the phase function $\Phi^{(\nu)}$ in the frequency matched case is given by the following formula:

$$\begin{aligned}\frac{1}{\beta^2}\Phi^{(\nu)}(\vec{\zeta}_0, \beta\vec{q}) = {} & \frac{\zeta}{2}[\mathbf{q}\cdot\omega''_{n_0}\mathbf{q} - \mathbf{q}'\cdot\omega''_{n_0}\mathbf{q}' - \mathbf{q}''\cdot\omega''_{n_0}\mathbf{q}'' - \mathbf{q}'''\cdot\omega''_{n_0}\mathbf{q}'''] \\ & + O(\beta|\vec{q}-\vec{q}^{\,\flat}|^2|\vec{q}|), \omega''_{n_0} = \omega''_{n_0}(\mathbf{k}_* + \mathbf{q}),\ \vec{q}^{\,\flat} = (\mathbf{q}, \mathbf{q}, \mathbf{q}, -\mathbf{q})\end{aligned} \tag{257}$$

This formula shows that the principal part of the phase function $1/\beta^2\Phi^{(\nu)}(\vec{\zeta}_0, \beta\vec{q})$ does not depend on β and is the same as the principal part of a similar phase function for the NLS. This explains the importance of the parameter θ in (24). Note that for $\nu = 2$ only the quadratic part of $\Phi^{(2)}$ can be nonzero.

4.1.1 Approximation of the modal susceptibility. In this section we introduce the expansion producing powers β^{l_2} in the structured power series (169). To get an expansion for the interaction integral in (238) we need to have an expansion for function $A_1(\beta\vec{q})$ as defined by (239). To have an expansion for $A_1(\beta\vec{q})$ we need, in turn, an expansion for one of its factors, namely the modal susceptibility $\check{Q}_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_* + \beta\vec{q})$ defined by (179). In fact, the coefficients of the expansion of the modal susceptibility determine the coefficients of a corresponding NLS. The resulting approximation polynomials in $\vec{q}$ applied in the Fourier representation lead to differential operators which are present in the NLS (see section 8.3.1).

We use the Taylor expansion in β for the modal susceptibility $\check{Q}_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_* + \beta\vec{q})$, namely

$$\begin{aligned}\check{Q}_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_* + \beta\vec{q}) &= \check{Q}_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_*) + \beta\check{Q}'_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_*, \vec{q}) + \frac{\beta^2}{2}\check{Q}''_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_*, \vec{q}) + \cdots \\ &\quad + \frac{\beta^{\nu-1}}{(\nu-1)!}\check{Q}^{(\sigma)}_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_*, \vec{q}) + O(\beta^{\sigma+1})\end{aligned} \tag{258}$$

where $\check{Q}^{(j)}_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_*, \vec{q})$ is a j-linear symmetric form of $\vec{q}$, in particular

$$\check{Q}'_{\vec{n}}(\vec{\zeta}_0\mathbf{k}_*, \vec{q}) = \nabla_{\vec{q}}\check{Q}_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_*)\cdot\vec{q},\ \ \check{Q}''_{\vec{n}}(\vec{\zeta}_0\mathbf{k}_*, \vec{q}) = \nabla^2_{\vec{q}}\check{Q}_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_*)\,\dot{:}\,(\vec{q}^{\,2}), \ldots \tag{259}$$

We introduce the Taylor polynomial $p^{[\sigma]}_{\mathrm{T},\zeta}(\beta\vec{q})$ of $\check{Q}_{\vec{n}_0}$ of degree σ by the formula

$$\begin{gathered}p^{[\sigma]}_{\mathrm{T},\zeta}(\beta\vec{q}) = \sum_{j=0}^{\sigma}\frac{1}{j!}\check{Q}^{(j)}_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_*, \beta\vec{q}) = \sum_{j=0}^{\sigma}\frac{\beta^j}{j!}\check{Q}^{(j)}_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_*, \vec{q}) \\ \vec{\zeta}_0 = (\zeta, \zeta, \zeta, -\zeta),\ \vec{n}_0 = ((\zeta, n), (\zeta, n_0), (\zeta, n_0), (-\zeta, n_0))\end{gathered} \tag{260}$$

Now we consider vectors and polynomials with a smaller number of variables, namely we eliminate $\mathbf{q}$ using the relation $\mathbf{q} = \mathbf{q}' + \mathbf{q}'' + \mathbf{q}'''$. Given a vector $\vec{q} = (\mathbf{q}, \mathbf{q}', \mathbf{q}'', \mathbf{q}''')$ we introduce vectors

$$\vec{q}^{\,\sharp} = (\mathbf{q}' + \mathbf{q}'' + \mathbf{q}''', \mathbf{q}', \mathbf{q}'', \mathbf{q}'''),\ \vec{q}^{\,\star} = (\mathbf{q}', \mathbf{q}'', \mathbf{q}''') \tag{261}$$

and the polynomial

$$p^{[\sigma]}_{\zeta}(\vec{q}^{\,\star}) = p^{[\sigma]}_{\mathrm{T},\zeta}(\vec{q}^{\,\sharp}),\ \zeta = \pm \tag{262}$$

Observe that the polynomial $p^{[\sigma]}_{\zeta}(\beta\vec{q}^{\,\star})$ defined by (262) has the following m-homogeneous terms

$$p^{[\sigma]}_{\zeta}(\beta\vec{q}^{\,\star}) = \sum_{j=0}^{\sigma}\beta^m p_{m,\zeta}(\vec{q}^{\,\star}),\ \ p_{m,\zeta}(\vec{q}^{\,\star}) = \frac{1}{m!}\check{Q}^{(m)}_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_*, \vec{q}^{\,\sharp}) \tag{263}$$

which evidently depend only on $\mathbf{q}', \mathbf{q}'', \mathbf{q}'''$. In particular, for $\sigma = 0$ we have

$$p_{0,\zeta} = p^{[0]}_{\zeta} = \check{Q}_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_*) = Q_\zeta = Q_\pm,\ \zeta = \pm 1 \tag{264}$$

Formulas (263) and (259) imply the following representation for the linear form $p_{1,\zeta}(\vec{q}^{\,\star})$:

$$p_{1,\zeta}(\vec{q}^{\,\star}) = \nabla_{\vec{q}^{\,\star}}\check{Q}_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_*)\cdot\vec{q}^{\,\star} + \nabla_{\mathbf{q}}\check{Q}_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_*)\cdot(\mathbf{q}' + \mathbf{q}'' + \mathbf{q}''') \tag{265}$$

Recasting formula (265) in terms of the gradients $\nabla_{\mathbf{q}'}$, $\nabla_{\mathbf{q}''}$ we get

$$p_{1,\zeta}(\vec{q}^{\,\star}) = \nabla_{\mathbf{q}'}\breve{Q}_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_*)\cdot\mathbf{q}' + \nabla_{\mathbf{q}''}\breve{Q}_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_*)\cdot\mathbf{q}'' + \nabla_{\mathbf{q}'''}\breve{Q}_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_*)\cdot\mathbf{q}''' + \nabla_{\mathbf{q}}\breve{Q}_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_*)\cdot(\mathbf{q}' + \mathbf{q}'' + \mathbf{q}''')$$

implying

$$p_{1,\zeta}(\vec{q}^{\,\star}) = a_{11,\zeta}\cdot\mathbf{q}' + a_{12,\zeta}\cdot\mathbf{q}'' + a_{13,\zeta}\cdot\mathbf{q}''' \tag{266}$$

with vectors $a_{11,\zeta}$, $a_{12,\zeta}$, $a_{13,\zeta}$ defined by

$$\begin{aligned} a_{11,\zeta} &= \nabla_{\mathbf{q}'}\breve{Q}_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_*) + \nabla_{\mathbf{q}}\breve{Q}_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_*),\ \ a_{12,\zeta} = \nabla_{\mathbf{q}''}\breve{Q}_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_*) + \nabla_{\mathbf{q}}\breve{Q}_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_*) \\ a_{13,\zeta} &= \nabla_{\mathbf{q}'''}\breve{Q}_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_*) + \nabla_{\mathbf{q}}\breve{Q}_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_*) \end{aligned} \tag{267}$$

The quadratic polynomial $p_{2,\zeta}(\vec{q}^{\,\star})$ has the following representation

$$\begin{aligned} p_{2,\zeta}(\vec{q}^{\,\star}) &= \nabla^2_{\vec{q}^{\,\star}}\breve{Q}_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_*)\,\vdots\,(\vec{q}^{\,\star})^2 + 2\nabla_{\vec{q}^{\,\star}}\nabla_{\mathbf{q}}\breve{Q}_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_*)\,\vdots\,(\mathbf{q}' + \mathbf{q}'' + \mathbf{q}''')(\vec{q}^{\,\star}) \\ &\quad + \nabla^2_{\mathbf{q}}\breve{Q}_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_*)\,\vdots\,(\mathbf{q}' + \mathbf{q}'' + \mathbf{q}''')^2 \end{aligned} \tag{268}$$

4.1.2 Asymptotic expansion in the weakly dispersive case. We now study the asymptotic expansions of the interaction integral (238) in the weakly dispersive case when (30) holds or, in other words, when the dispersion parameter θ satisfies the inequality

$$\theta = \frac{\varrho}{\beta^2} \geq \theta_0 \text{ with a fixed } \theta_0 > 0 \tag{269}$$

In fact, the most interesting is the borderline case which corresponds to the classical nonlinear Schrödinger scaling (26), namely $\alpha \sim \varrho \sim \beta^2$. To get the integral expansions in β up to the order $\sigma = \nu - 2$ we need the corresponding expansions for the involved integrands $A_1(\beta\vec{q})$ which can be found as follows. Using (239) and (258) and taking into account that for small β, in view of (190), $\Psi(\beta\mathbf{q}) = 1$ we obtain

$$A_1(\beta\vec{q}) = p^{[\sigma]}_{\mathrm{T},\zeta}(\beta\vec{q}) + O(\beta^{\sigma+1}),\ \sigma = 0, \ldots, \nu - 2 \tag{270}$$

We have

$$\phi_{\vec{n}}(\vec{\zeta}_0\vec{k}_* + \beta\vec{q}) = \Phi^{(\nu)}(\vec{\zeta}_0, \beta\vec{q}) + O(\beta^{\nu+1}) \tag{271}$$

and

$$\exp\left\{\mathrm{i}\phi_{\vec{n}}(\vec{\zeta}_0\vec{k}_* + \beta\vec{q})\frac{\tau_1}{\varrho}\right\} = \exp\left\{\mathrm{i}\Phi^{(\nu)}(\vec{\zeta}_0, \beta\vec{q}^{\,0})\frac{\tau_1}{\varrho}\right\}\exp\left\{\mathrm{i}\frac{\tau_1}{\varrho}O(\beta^{\nu+1})\right\} \tag{272}$$

where we have a standard series expansion

$$\exp\left\{\mathrm{i}\frac{\tau_1}{\varrho}O(\beta^{\nu+1})\right\} = 1 + \mathrm{i}\frac{\tau_1\beta^{\nu+1}}{\varrho}O(1) + \cdots \tag{273}$$

This expansion leads to (169). We conclude that in the interaction integral (238) we can write

$$\exp\left\{\mathrm{i}\phi_{\vec{n}}(\vec{\zeta}_0\vec{k}_* + \beta\vec{q}^{\,0})\frac{\tau_1}{\varrho}\right\} = \exp\left\{\mathrm{i}\Phi^{(\nu)}(\vec{\zeta}_0, \beta\vec{q}^{\,0})\frac{\tau_1}{\varrho}\right\}\left\{1 + O\left(\frac{\beta^{\nu+1}}{\varrho}\right)\right\} \tag{274}$$

Recall that, by (269), in the weakly dispersive case $\beta^{\nu+1}\tau_1/\varrho \ll 1$ when $\nu \geq 2$. Using (270) we infer from (238) that

$$\beta^d I_{\bar{n},\zeta,\zeta,-\zeta}(\zeta\mathbf{k}_* + \beta\mathbf{q}, \tau) = \frac{1}{\varrho}\int_0^\tau \int_{\mathbb{R}^{2d}} \exp\left\{\mathrm{i}\Phi^{(\nu)}\big(\vec{\zeta}_0, \beta\vec{q}^{\,0}\big)\frac{\tau_1}{\varrho}\right\}\left\{1 + O\left(\frac{\beta^{\nu+1}}{\varrho}\right)\right\}$$
$$\times\psi^3(\tau_1)\big(p^{[\sigma]}_{\mathrm{T},\zeta}(\beta\vec{q}) + O(\beta^{\sigma+1})\big)(\hat{h}_\zeta(\mathbf{q}')\hat{h}_\zeta(\mathbf{q}'')$$
$$\times\,\hat{h}_{-\zeta}(\mathbf{q}'''(\vec{q})))\mathrm{d}\mathbf{q}'\mathrm{d}\mathbf{q}''\mathrm{d}\tau_1 + O\left(\frac{\beta^{N_\Psi - d}}{\varrho}\right) \tag{275}$$

Remark The term $O(\beta^{N_\Psi - d}/\varrho)$ in (275) arises from replacing Ψ by 1. When all $|\mathbf{q}'|$, $|\mathbf{q}''|$, $|\mathbf{q}'''|$ are smaller than $\pi_0 2\beta$ we have $\Psi(\beta\mathbf{q}') = \Psi(\beta\mathbf{q}'') = \Psi(\beta\mathbf{q}''') = 1$ and this replacement creates no error at all. When one of arguments, for example $|\mathbf{q}'''|$, is greater than $\pi_0/2\beta$ then, by (206) $\hat{h}(\mathbf{q}''')$ is very small for large values of arguments. Using these observations we obtain the term $O(\beta^{N_\Psi - d}/\varrho)$ in (275). Note that since N_Ψ can be chosen arbitrary large and the relations (29) or (31) hold we assume that N_Ψ is large enough to yield the following inequalities

$$O\left(\frac{\beta^{N_\Psi - d}}{\varrho}\right) \ll O\left(\frac{\beta^{\sigma+1}}{\varrho}\right),\ O\left(\frac{\beta^{N_\Psi - d}}{\varrho}\right) \ll O\left(\frac{\beta^{\nu+1}}{\varrho^2}\right) \tag{276}$$

Under this condition this term in (275) is negligible. Hence, if one replaces Ψ by 1 the error is negligible. ♦

4.2 *Approximation of the interaction integral*

Here we approximate the interaction integral $I_{\bar{n},\zeta,\zeta,-\zeta}$ defined by (238) by a simpler integral $I^{(\sigma)}_{\bar{n},\zeta,\zeta,-\zeta}$ which is introduced below. Using this approximation we will be able to relate solutions to the NLM and the NLS.

The approximation $I^{(\sigma)}_{\bar{n},\zeta,\zeta,-\zeta}$, $\sigma = 0, 1, 2$, is constructed by the following alterations in the integral $I_{\bar{n},\zeta,\zeta,-\zeta}$ represented by (238). In the expression $A_1(\beta\vec{q})$ defined by (239) the modal susceptibility $\check{Q}_{\vec{n}_0}((\vec{\zeta}_0\mathbf{k}_* + \beta\vec{q}))$ is replaced with the polynomial $p^{[\sigma]}_{\mathrm{T},\zeta}(\beta\vec{q})$ defined by (260), and the cut-off function Ψ defined by (190) is replaced by 1. Thus, we introduce the integral $I^{(\sigma)}_{\bar{n},\zeta,\zeta,-\zeta}$ by the following formula

$$\beta^d I^{(\sigma)}_{\bar{n},\zeta,\zeta,-\zeta}(\zeta\mathbf{k}_* + \beta\mathbf{q}, \tau) = \frac{1}{\varrho}\int_0^\tau \mathrm{d}\tau_1 \int_{\mathbb{R}^{2d}} \exp\left\{\mathrm{i}\Phi^{(\nu)}\big(\vec{\zeta}_0, \beta\vec{q}^{\,0}\big)\frac{\tau_1}{\varrho}\right\} p^{[\sigma]}_{\mathrm{T},\zeta}(\beta\vec{q}^{\,0})$$
$$\times\psi^3(\tau_1)\hat{h}_\zeta(\mathbf{q}')\hat{h}_\zeta(\mathbf{q}'')\hat{h}_{-\zeta}(\mathbf{q} - \mathbf{q}' - \mathbf{q}'')\,\mathrm{d}\mathbf{q}'\,\mathrm{d}\mathbf{q}'', \tag{277}$$

In particular, for $\sigma = 0$ we have

$$\beta^d I^{(0)}_{\bar{n},\zeta,\zeta,-\zeta}(\zeta\mathbf{k}_* + \beta\mathbf{q}, \tau) = \frac{1}{\varrho}\int_0^\tau \mathrm{d}\tau_1 \int_{\mathbb{R}^{2d}} \exp\left\{\mathrm{i}\Phi^{(\nu)}(\vec{\zeta}_0, \beta\vec{q}^{\,0})\frac{\tau_1}{\varrho}\right\} \check{Q}_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_*)$$
$$\times\psi^3(\tau_1)\hat{h}_\zeta(\mathbf{q}')\hat{h}_\zeta(\mathbf{q}'')\hat{h}_{-\zeta}(\mathbf{q} - \mathbf{q}' - \mathbf{q}'')\,\mathrm{d}\mathbf{q}'\mathrm{d}\mathbf{q}'' \tag{278}$$

Let us show now that $I^{(\sigma)}_{\bar{n},\zeta,\zeta,-\zeta}(\zeta\mathbf{k}_* + \boldsymbol{\eta}, \tau)$ provides a good approximation to $I_{\bar{n},\zeta,\zeta,-\zeta}(\zeta\mathbf{k}_* + \boldsymbol{\eta}, \tau)$ in the weakly dispersive case when (30) and (269) hold (strongly dispersive case is discussed in section 4.3).

Comparing (275) with (277) and using (276) we obtain the following estimate

$$\beta^d I_{\bar{n},\zeta,\zeta,-\zeta}(\zeta\mathbf{k}_* + \beta\mathbf{q}, \tau) = \beta^d I^{(\sigma)}_{\bar{n},\zeta,\zeta,-\zeta}(\zeta\mathbf{k}_* + \beta\mathbf{q}, \tau) + O\left(\frac{\beta^{\nu+1}}{\varrho^2}\right) + O\left(\frac{\beta^{\sigma+1}}{\varrho}\right). \tag{279}$$

Now we use the above results to estimate the leading term $\tilde{u}_{\bar{n}}^{(1,0)}$ as in (210) of the expansion (175). Applying (210), (214), (248) and (279) we get the following formula for $\sigma \leq \nu - 2$

$$\begin{aligned} &\beta^d \tilde{u}_{\bar{n}}^{(1,0)}(\zeta \mathbf{k}_* + \beta \mathbf{q}, t) + \beta^d \tilde{u}_{\bar{n}}^{(1)}(\mathbf{J}_1; \zeta \mathbf{k}_* + \boldsymbol{\eta}, t) \\ &\quad = -3\beta^d I_{\bar{n},\zeta,\zeta,-\zeta}^{(\sigma)}(\zeta \mathbf{k}_* + \beta \mathbf{q}, \tau) + O\left(\frac{\beta^{\nu+1}}{\varrho^2}\right) + O\left(\frac{\beta^{\sigma+1}}{\varrho}\right) + O(1) \end{aligned} \tag{280}$$

(Note that leading expressions in (280) are of order ϱ^{-1}.)

4.3 *The strongly dispersive case: rectifying change of variables and critical points of the interaction phase*

In the strongly dispersive case when (28) holds, that is $\theta = \varrho/\beta^2 \ll 1$, we cannot apply the elementary approach of section 4.2. Still the NLM–NLS approximation is valid with the same NLS or ENLS, but the mathematical treatment is different. To provide applicability on very long time intervals of order $1/\varrho \gg \beta^{-2}$ where linear dispersive effects are much stronger, we have to match the linear Maxwell equation and the linear Schrödinger equation with a higher precision. This is made by applying a rectifying change of variables.

The *rectifying change of variables* Y is a one-to-one mapping of a small vicinity of a point $\mathbf{k}_*$ onto a similar vicinity, i.e.

$$\boldsymbol{\xi} = Y^{-1}(\boldsymbol{\eta}), \ \boldsymbol{\eta} = Y(\boldsymbol{\xi}) \text{ if } |\boldsymbol{\eta}| \leq 2\pi_0, \ \boldsymbol{\eta} = Y(\boldsymbol{\xi}) \text{ if } |\boldsymbol{\xi}| \leq 2\pi_0 \tag{281}$$

where π_0 is a small constant. *It converts the dispersion relation* $\omega_{n_0}(\mathbf{k}_* + \boldsymbol{\eta})$ *into its Taylor polynomial* $\gamma_{(\nu)}(\mathbf{k}_*; \boldsymbol{\eta})$ *at* k_* *of the degree* ν, i.e.

$$\omega_{n_0}(\mathbf{k}_* + Y(\boldsymbol{\xi})) = \gamma_{(\nu)}(\mathbf{k}_*; \boldsymbol{\xi}) = \gamma_{(\nu)}(\boldsymbol{\xi}) \tag{282}$$

We will refer to coordinates $\boldsymbol{\xi}$ as *rectifying coordinates*. This rectifying change of variables exactly reduces the linear part of the NLM to the linear part of the NLS for single doublet excitations localized around k_*. The rectifying change of variables $Y_\zeta^{-1}(\boldsymbol{\eta})$ allows us to establish an *exact equivalence* between the dynamics of the NLM and the NLS in the linear approximation ($\alpha = 0$) *for arbitrary long times*.

The rectifying change of variables $Y(\boldsymbol{\xi})$ satisfying (282) exists by the implicit function theorem and its power series expansions can be explicitly found. The rectifying change of variables $Y(\boldsymbol{\xi})$ is close to the identity, and if

$$\omega'_{n_0}(\mathbf{k}_*) \neq 0 \tag{283}$$

then

$$Y(\boldsymbol{\xi}) = \boldsymbol{\xi} + O(|\boldsymbol{\xi}|^{\nu+1}), \ Y^{-1}(\boldsymbol{\eta}) = \boldsymbol{\eta} + O(|\boldsymbol{\eta}|^{\nu+1}), \ \text{if } |\boldsymbol{\eta}| \leq 2\pi_0, \ |\boldsymbol{\xi}| \leq 2\pi_0 \tag{284}$$

If (283) does not hold, but instead we have

$$\omega'_n(\mathbf{k}_*) = 0, \ \det \omega''_{\bar{n}}(\mathbf{k}_*) \neq 0 \tag{285}$$

then $Y(\boldsymbol{\xi})$ exists by the Morse lemma [38], section 2.3.2, and

$$Y(\boldsymbol{\xi}) = \boldsymbol{\xi} + O(|\boldsymbol{\xi}|^{\nu}), \ Y^{-1}(\boldsymbol{\eta}) = \boldsymbol{\eta} + O(|\boldsymbol{\eta}|^{\nu}) \text{ if } |\boldsymbol{\eta}| \leq 2\pi_0, \ |\boldsymbol{\xi}| \leq 2\pi_0 \tag{286}$$

In this paper we assume that (283) holds. The case (285) in many respects is similar, but requires somewhat different treatment of higher order terms of asymptotic expansions. In the strongly dispersive case (28) we assume that

$$\omega_{n_0}(\pm\mathbf{k}_*) \neq 0, \ \det \omega''_{n_0}(\pm\mathbf{k}_*) \neq 0 \tag{287}$$

We use the stationary phase method along the lines of [17–19] to find an asymptotic expansion for the interaction integral (238) with respect to the small parameter θ. According to the method, we need to find the critical points of the phase (178) under the restriction (233) with respect to the variables $\mathbf{q}'$, $\mathbf{q}''$. The critical points are the solutions to the following system of equations

$$\nabla_{\mathbf{q}'}\phi_{\vec{n}}(\vec{\zeta}\mathbf{k}_* + \vec{q}) = 0, \ \nabla_{\mathbf{q}''}\phi_{\vec{n}}(\vec{\zeta}\mathbf{k}_* + \vec{q}) = 0 \tag{288}$$

Taking into account the inversion symmetry identities (140) we find that all small solutions to the system (288) together with (233) and (253) are exhausted by the following vectors

$$\vec{q}^{\,\flat} = (\mathbf{q}, \mathbf{q}, \mathbf{q}, -\mathbf{q}), \text{ i.e. } \mathbf{q}' = \mathbf{q}'' = \mathbf{q}, \ \mathbf{q}''' = -\mathbf{q} \tag{289}$$

with $\mathbf{q}$ being arbitrary (but small). *Remarkably, introduction of the rectifying change of variables does not affect the critical points. For any fixed and sufficiently small* $\mathbf{q}$ *the interaction phase* $\phi_{\vec{n}}(\vec{\zeta}\mathbf{k}_* + \vec{q})$ *has a unique critical point* $\vec{q}^{\flat}$ *described by (289).*

Using this fact and (286) we can prove that the estimates proven for the weakly dispersive case hold in the strongly dispersive case too. Note that the magnitude of the nonlinear corrections, in particular FNLR $\mathbf{U}^{(1)}$ on time intervals which satisfy (15) is estimated as follows:

$$\begin{aligned} &O\left(\left|\mathbf{U}^{(1)}\right|\right) = O(\varrho^{d-1})O(|\mathbf{J}|^3) \text{ in the dispersive case } \theta^{-1} \gg \|\omega''_{n_0}(\mathbf{k}_*)^{-1}\|, \\ &O\left(\left|\mathbf{U}^{(1)}\right|\right) = O(\varrho^{-1})O(|\mathbf{J}|^3) \text{ in the weakly dispersive case } \theta^{-1} \ll \|\omega''_{n_0}(\mathbf{k}_*)\|^{-1} \end{aligned} \tag{290}$$

When $\mathbf{U} = \mathbf{U}(\mathbf{r}, t)$ is a function of $\mathbf{r}, t$ we write $O(|\mathbf{U}|)$ for a function of $\mathbf{r}, t$ such that it is bounded in some sense when $\mathbf{U}$ is bounded, assuming that $O(|\mathbf{U}|)$ is homogeneous in $\mathbf{U}$ (in particular $O(|\beta^q\mathbf{U}|) = \beta^q O(|\mathbf{U}|)$). We do not want to elaborate and get more specific on the definition of $O(|\mathbf{U}|)$ since a mathematically rigorous discussion of this subject would require us to introduce concepts and technicalities that, though important for a mathematical justification, are not essential for presenting the results of our analysis. According to (290) and (16) in the weakly dispersive case

$$O\left(\alpha\left|\mathbf{U}^{(1)}\right|\right) = O(1) \tag{291}$$

whereas in the strongly dispersive case $O(\alpha|\mathbf{U}^{(1)}|) = O(\varrho^d)$ is much smaller. Therefore, the error terms in the right-hand sides in estimates (46), (49), (59), (65), (94), (106) and (112) now should include the extra factor ϱ^d. For example, in the strongly dispersive case (112) takes the form

$$\mathbf{U} - \mathbf{U}_Z = O(\varrho^d)[O(\beta^2) + O(\varrho)] \tag{292}$$

The definition of excitation currents now includes the rectifying change of variables

$$Y_\zeta(\boldsymbol{\xi}) = \zeta Y_\zeta(\zeta\xi), \ \zeta = \pm$$

namely in (188)

$$\tilde{j}^{(0)}_{\zeta,n_0}(\mathbf{k}, \tau) = -\varrho\beta^{-d}\psi_0(\tau)\Psi(\mathbf{k} - \mathbf{k}_*)\hat{h}_\zeta\left(\frac{1}{\beta}Y_\zeta^{-1}(\mathbf{k} - \zeta\mathbf{k}_*)\right) \tag{293}$$

and in (185)

$$\tilde{j}^{(1)}_{\zeta,n_0}(\zeta\mathbf{k}_* + Y_\zeta(\beta\mathbf{q}), \tau) = \hat{J}^{(1)}_{Z,\zeta}(\beta\mathbf{q}, \tau) \tag{294}$$

The relation between NLS and NLM, instead of (58), takes the form

$$\mathbf{U}^{\mathrm{dir}}_{Z,n_0}(\mathbf{r},t) = \frac{1}{(2\pi)^d}\int_{[-\pi,\pi]^d} \Psi(\eta)[\hat{Z}_+(Y_+^{-1}(\boldsymbol{\eta}),t)\tilde{\mathbf{G}}_{+,n_0}(\mathbf{r},\mathbf{k}_*+\eta) + \hat{Z}_-(Y_-^{-1}(\boldsymbol{\eta}),t)\tilde{\mathbf{G}}_{-,n_0}(\mathbf{r},-\mathbf{k}_*-\boldsymbol{\eta})]\,\mathrm{d}\boldsymbol{\eta} \tag{295}$$

5. Tailoring the NLS to approximate the NLM

In this section we introduce an NLS that is tailored to approximate the NLM. The NLSs we are interested in are the two equations (36) and (37) or their d-dimensional analogues (100) and (101). The two equations correspond to the two values of $\zeta = \pm 1 = \pm$.

The NLS (36) involves some constants that are to be related to the NLM. To do that let us construct first a linear part of NLS related to the linear part of NLM. We begin with picking a $\nu = 2, 3$ or 4 and then we introduce the Taylor polynomial $\gamma_{(\nu)}(\eta)$ of the order ν of the function $\omega_{n_0}(\mathbf{k})$ at the point $\mathbf{k}_*$ as defined by (200). The case $\nu = 2$ corresponds to the classical NLS, and $\nu = 3, 4$ correspond to an ENLS. For these three cases we have respectively

$$\begin{aligned}
\gamma_{(2)}(\boldsymbol{\eta}) &= \omega_{n_0}(\mathbf{k}_*) + \omega'_{n_0}(\mathbf{k}_*)(\eta) + \frac{1}{2}\omega''_{n_0}(\mathbf{k}_*)(\boldsymbol{\eta}^2)\\
\gamma_{(3)}(\boldsymbol{\eta}) &= \omega_{n_0}(\mathbf{k}_*) + \omega'_{n_0}(\mathbf{k}_*)(\eta) + \frac{1}{2}\omega''_{n_0}(\mathbf{k}_*)(\boldsymbol{\eta}^2) + \frac{1}{6}\omega'''_{n_0}(\mathbf{k}_*)(\boldsymbol{\eta})^3\\
\gamma_{(4)}(\boldsymbol{\eta}) &= \omega_{n_0}(\mathbf{k}_*) + \omega'_{n_0}(\mathbf{k}_*)(\eta) + \frac{1}{2}\omega''_{n_0}(\mathbf{k}_*)(\boldsymbol{\eta}^2) + \frac{1}{6}\omega'''_{n_0}(\mathbf{k}_*)(\boldsymbol{\eta})^3 + \frac{1}{24}\omega''''_{n_0}(\mathbf{k}_*)(\boldsymbol{\eta})^4
\end{aligned} \tag{296}$$

Substituting $\eta_j = -\mathrm{i}\partial_j$ into the polynomial $\gamma_{(\nu)}(\boldsymbol{\eta})$ we obtain the differential operator $\gamma_{(\nu)}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]$, in particular:

$$\gamma_{(3)}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]V = \omega_{n_0}(\mathbf{k}_*)V - \mathrm{i}\sum_m \gamma_m\partial_m V - \frac{1}{2}\sum_{m,l}\gamma_{ml}\partial_m\partial_l V + \frac{\mathrm{i}}{6}\sum_{m,l,j}\gamma_{mlj}\partial_m\partial_l\partial_j V \tag{297}$$

where γ_m, γ_{ml} and γ_{mlj} are the real-valued coefficients of the linear form $\omega'_{n_0}(\mathbf{k}_*)$, the quadratic form $\omega''_{n_0}(\mathbf{k}_*)$ and the cubic form $\omega'''_{n_0}(\mathbf{k}_*)$, respectively. In the simplest classical case when $\nu = 2$ and the problem is one-dimensional, i.e. $d = 1$, as in (37), the operator $\gamma_{(2)}(-\mathrm{i}\partial_x)$ has the form

$$\gamma_{(2)}(-\mathrm{i}\partial_x)V = \omega_{n_0}(\mathbf{k}_*)V - \mathrm{i}\omega'_{n_0}(\mathbf{k}_*)\partial_x V - \frac{1}{2}\omega''_{n_0}(\mathbf{k}_*)\partial_x^2 V \tag{298}$$

with the symbol $\gamma_{(2)}(\boldsymbol{\eta})$ being defined by (296). The polynomial $\gamma_{(\nu)}(\boldsymbol{\eta})$ is called the *symbol* of the operator $\gamma_{(\nu)}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]$.

Let us introduce a general *linear Schrödinger equation* of the form

$$\partial_t Z(\mathbf{r},t) = -\mathrm{i}\gamma_{(\nu)}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]Z(\mathbf{r},t),\ Z(\mathbf{r},t)|_{t=0} = h_\beta(\mathbf{r}),\ h_\beta(\mathbf{r}) = h(\beta\mathbf{r}) \tag{299}$$

This can be solved exactly in terms of the Fourier transform, namely

$$\hat{Z}(\boldsymbol{\eta},t) = \hat{h}_\beta(\boldsymbol{\eta})\exp\{-\mathrm{i}\gamma_{(\nu)}(\eta)t\} \tag{300}$$

The properties of the Fourier transform are discussed in section 8.2 (see equation (425) for its definition). Let us also consider the classical *nonlinear Schrödinger equation*

$$\partial_t Z_+ = -\mathrm{i}\gamma_{(2)}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]Z_+ + \alpha_\pi Q_+|Z_+|^2 Z_+,\ Z_+(\mathbf{r},t)|_{t=0} = h_+(\beta\mathbf{r}) = h_{+,\beta}(\mathbf{r}) \tag{301}$$

where Q_+ is a complex constant and the factor $\alpha_\pi = 3\alpha(2\pi)^{2d}$ is introduced for notational consistency with the related NLM (we have used (38) to simplify (36) and (37)).

The simplest *extended nonlinear Schrödinger equations* are given by (98) and (99); they have the form

$$\begin{aligned}\partial_t Z_\zeta &= -\mathrm{i}\zeta\gamma_{(\nu)}[-\mathrm{i}\zeta\vec{\nabla}_{\mathbf{r}}]Z_\zeta + \alpha_\pi p_\zeta^{[\nu-2]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]\big(Z_\zeta^2 Z_{-\zeta}\big)\\ Z_\zeta(\mathbf{r},t)|_{t=0} &= h_\zeta(\beta\mathbf{r}) = h_{\zeta,\beta}(\mathbf{r}),\ \zeta=\pm\end{aligned} \tag{302}$$

where $p_\zeta^{[\sigma]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}](Z_\zeta^2 Z_{-\zeta})$ is a linear differential operator with constant coefficients of the order $\sigma = \nu - 2$ with $p_\pm^{[\sigma]}(\vec{q}^\star)$ being defined by (262). This operator acts on the factors of the product $Z_\zeta^2 Z_{-\zeta}$. The action of such an operator is defined by (434). Note that this operator acts on all factors of the product $Z_\zeta^2 Z_{-\zeta} = Z_\zeta Z_\zeta Z_{-\zeta}$, and that the variables $\mathbf{q}', \mathbf{q}'', \mathbf{q}'''$ of the symbol are replaced, respectively, by differentiations of the first, second and third factor. In particular, according to (264)

$$p_\zeta^{[0]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]\big(Z_\zeta^2 Z_{-\zeta}\big) = Q_\zeta Z_\zeta^2 Z_{-\zeta},\ \zeta=\pm \tag{303}$$

where the coefficient Q_ζ is given in (264). The first-order operator $p_\zeta^{[1]} = p_\zeta^{[0]} + p_{1,\zeta}$ where $p_\zeta^{[0]}$ is given above and the symbol of $p_{1,\zeta}$ is defined by (266), i.e.

$$p_{1,\zeta}^{[1]}(\vec{q}^{\,\star}) = a_{11,\zeta}\cdot\mathbf{q}' + a_{12,\zeta}\cdot\mathbf{q}'' + a_{13,\zeta}\cdot\mathbf{q}''' \tag{304}$$

The corresponding operator acts as follows

$$\begin{aligned}p_{1,\zeta}^{[1]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]\big(Z_\zeta^2 Z_{-\zeta}\big) &= p_{1,\zeta}^{[1]}[\vec{\nabla}_{\mathbf{r}}](Z_\zeta Z_\zeta Z_{-\zeta})\\ &= Z_\zeta Z_{-\zeta}(a_{11,\zeta}+a_{12,\zeta})\cdot\nabla_{\mathbf{r}}Z_\zeta + Z_\zeta^2 a_{13,\zeta}\cdot\nabla_{\mathbf{r}}Z_{-\zeta}\end{aligned} \tag{305}$$

where the vectors $a_{11,\zeta}$, $a_{12,\zeta}$ and $a_{13,\zeta}$ are defined in (267); the components of the vectors are complex. Observe that the extended NLS (302) turns into the classical one (301) if we set $\nu = 2$, $\sigma = 0$ and use (38).

It is useful to consider the following scaled version of the function $Z_\pm(\mathbf{r}, t)$:

$$Z_{\beta,\zeta}(\mathbf{r},t) = Z_\zeta\left(\frac{\mathbf{r}}{\beta}, t\right),\ \zeta=\pm \tag{306}$$

The relation (306) between $Z_\pm(\mathbf{r}, t)$ and its scaled version $Z_{\beta,\pm}(\mathbf{r}, t)$ implies the following relation between their Fourier transforms as defined by (425):

$$\hat{Z}_\zeta(\boldsymbol{\eta},t) = \int_{\mathbf{R}^d} \mathrm{e}^{-\mathrm{i}\beta\mathbf{r}\cdot\frac{1}{\beta}\boldsymbol{\eta}} Z_{\beta,\zeta}(\beta\mathbf{r},t)\,\mathrm{d}\mathbf{r} = \beta^{-d}\hat{Z}_{\beta,\zeta}\left(\frac{\boldsymbol{\eta}}{\beta},t\right) \tag{307}$$

It is convenient to recast the general ENLSs (302) as a system for the rescaled quantities $Z_{\beta,\zeta}(\mathbf{r}, t)$, $\zeta = \pm$, namely

$$\begin{aligned}\partial_t Z_{\beta,\zeta} &= -\mathrm{i}\gamma_{(\nu)}[-\mathrm{i}\zeta\beta\nabla_{\mathbf{r}}]Z_{\beta,\zeta} + \alpha_\pi p_\zeta^{[\sigma]}[-\mathrm{i}\beta\nabla_{\mathbf{r}}]\big(Z_{\beta,+}^2 Z_{\beta,-\zeta}\big),\\ Z_{\beta,\zeta}(\mathbf{r},t)|_{t=0} &= h_\zeta(\mathbf{r})\end{aligned} \tag{308}$$

Obviously, the initial data for the rescaled equation do not depend on β, but the coefficients explicitly depend on β.

5.1 *Total error of the approximation of the NLM by an NLS*

In this section we outline how we estimate the total error of the approximation of the NLM with an NLS. For simplicity we discuss the case when the NLM contains purely cubic nonlinearity and we use a second-order NLS (that is with the order of the linear part $\nu = 2$) for the approximation.

An exact solution $\mathbf{U}(\mathbf{r}, t)$ of the NLM corresponding to an excitation current composed from a doublet of modes, as in (186) and (188), splits naturally into two parts corresponding to the directly and indirectly excited modes. The first part involves the directly excited modes (excited through the linear medium response) with modal amplitudes $U_{\zeta,n}(\mathbf{k}_*, t)$ with $n = n_0$, $|\mathbf{k} - \zeta \mathbf{k}_*| \leq \pi_0$ as in (219). Their magnitude is $O(1)$. The second part consists of the indirectly excited modes, which are excited solely through the nonlinear medium response, and this part involves modes with either $n \neq n_0$ or $|\mathbf{k} - \zeta \mathbf{k}_*| > \pi_0$. As was explained in section 3.2 the magnitude of the indirectly excited modes is estimated by $O(\varrho)$. The approximate solution includes both directly and indirectly excited modes. We take a solution $Z_\zeta(\mathbf{r}, t)$ of the NLS (100) and (101) and consider an exact solution $\mathbf{U}(\mathbf{r}, t)$ related to $Z_\zeta(\mathbf{r}, t)$ through properly chosen excitation currents. The currents are based on the initial data of the NLS (see section 5.2). We define the approximate solution $\mathbf{U}_Z(\mathbf{r}, t)$ of the NLM by (58), (68) and (69).

The nonlinear interactions of the directly excited modes with themselves are of order $O(1)$ as in the case of classical NLS scaling (26) and they, of course, are taken into account. We approximate the directly excited modes by appropriate solutions of the NLS. *To match/correspond the NLM and the NLS we use in concert the following two options: (i) setting up the excitation currents; (ii) choosing the coefficients of the NLS.* The linear part of the NLS is obtained based on the Taylor expansion of the dispersion relation, see (297). We choose the coefficients at the nonlinear terms of the NLS so that the FNLR of the NLM exactly matches the FNLR of the NLS.

The following sources of approximation error exist. First, we replace the causal integral nonlinear operators which enter the operator $\mathcal{F}_{\mathrm{NL}}$ by the instantaneous operators described by the nonlinear susceptibilities, see section 6 and [39] for details. This is necessary since the nonlinearity in the NLS is instantaneous. Second, we neglect the impact of indirectly excited modes onto directly excited. More precisely, we throw away all terms in (181) with $n' \neq n_0$, $n'' \neq n_0$, $n''' \neq n_0$. This is necessary if we consider dynamics of modal amplitudes of only one band $n = n_0$ independently of all other bands. Third, we replace exact dispersion relation $\omega_{n_0}(\mathbf{k})$ by its Taylor polynomial at $\mathbf{k} = \mathbf{k}_*$. Fourth, we replace frequency-dependent susceptibilities by their values at $\omega = \omega_{n_0}(\mathbf{k}_*)$; this is necessary since the nonlinearity in the NLS is not frequency-dependent.

Note that all mentioned replacements and modifications affect the FNLR in exactly the same way as the exact solution. The only difference is that for the FNLR the operators we mentioned above are applied to the linear approximation $\mathbf{U}^{(0)}$ whereas for the exact solution they act on $\mathbf{U}$ itself. This explains why the choice of the coefficients of the NLS based on matching the FNLR of the NLS and NLM gives a good approximation of exact solutions of NLM even in the case of classical NLS scaling $\varrho \sim \alpha \sim \beta^2$ for times $t \sim 1/\varrho$. Since we match only the zero and the first-order terms in α (the linear response and the FNLR), the higher order terms of order α^2 could create an additional discrepancy of order one, but the effect of the higher order terms is effectively eliminated since *we approximate the solution of the NLM by the exact solution of the NLS* rather than by the principal terms of the expansion in α of the solution of the NLS, see section 7 for details.

The total approximation error of the approximation of the exact solution $\mathbf{U}(\mathbf{r}, t)$ by the approximate solution $\mathbf{U}_Z(\mathbf{r}, t)$ defined by (58), (68) and (69) on the time interval (97) in the case $\nu = 2$ consists of the following components:

(i) the error of the nonlinear Schrödinger approximation of the directly excited modes on the interval $0 \leq t \leq \tau_*/\varrho$ is estimated by $O(\beta)$;

(ii) the error of the zero-order time-harmonic approximation (395) to the causal integral is estimated by $O(\varrho)$;

(iii) the error of the FNLR approximation ((68) and (69)) of the indirectly excited modes, in particular through non-FM interactions, is estimated by $O(\varrho)$;

(iv) the error from the impact of indirectly excited modes onto the directly excited modes (interband interactions) is estimated by $O(\varrho)$;

(v) the error of the polynomial approximation of the dispersion relation (271) in the weakly dispersive case is $O(\beta^3/\varrho)$, which gives $O(\beta)$ in the case of classical nonlinear Schrödinger scaling.

The inequality $t \leq \tau_*/\varrho$ in (97) ensures that our analysis is applicable, see (418). Consequently the total error of approximation of a solution to the NLM by a solution to the classical NLS when $\nu = 2$ is of order

$$O(\beta) + O(\varrho) \tag{309}$$

The error estimate in item (i) in the above list is addressed below in this section. The error estimate in item (ii) is discussed in section 6.2. The error estimate in item (iii) was discussed in sections 1.2 and 3.2. The error in item (iv) caused by interband interactions includes higher order terms of power expansions (11) and will be considered in a separate paper. The error in item (v) was discussed in section 4.1. There are also the negligible errors $O(\theta^{N_3+1})$ in the case (28) when $\theta = \varrho/\beta^2 < 1$ with arbitrary large N_3 and the error from the contribution of the cut-off function Ψ of the order $O(\beta^{N_\Psi - d})$ (see Remark in section 4.1.2) where N_Ψ is arbitrarily large; these errors are technical by nature and are negligible at any order of accuracy.

Reduction of the errors by means of using an ENLS instead of a classical NLS was discussed in section 1.2.

5.2 *The linear response and the FNLR for an NLS*

To provide a basis for relating the NLM and an NLS using their linear and first nonlinear responses, we need to construct for the general NLS (302) the linear and the first nonlinear responses along the same lines as we did for the NLM. For that we: (i) single out the linear part of the general NLS (302) and carry out its spectral analysis; (ii) introduce the source term in the NLS which replaces the initial condition; (iii) study the corresponding solution using the framework we developed for the NLM. The source term is introduced based on the initial data of the NLS so that it: (i) generates the same solution as the initial data; (ii) has the form of an almost time-harmonic function consistent with (158), (163) and (188). The importance of the relation between the excitation current for the NLM and initial data for the NLS can be seen from the following simple observation. When we compare solutions of the two differential equations the difference of the two solutions originates from two sources: the difference between the equations and the difference between the initial data. Even when the equation is the same, the difference in the solutions is proportional to the difference of the initial data. Since we study approximations of the solutions of the NLM by solutions of the NLS with high precision, and study effects of additional terms in the ENLS on the accuracy of approximation, we want to completely eliminate the source of differences that come from the initial data. This is not trivial since the initial data $h(\mathbf{r})$ for the NLS are instantaneously prescribed at $t = 0$ and their counterpart (the excitation currents $\mathbf{J}(\mathbf{r}, t)$ for the NLM) are defined on a time interval $0 \leq t \leq \tau_0/\varrho$. That requires consideration of technical issues, but the bottom line is that exact matching is possible for an arbitrary choice of h. We recall also that the constructed excitation current $\mathbf{J}(\mathbf{r}, t)$ vanishes for $t \geq \tau_0/\varrho$ and only after that time can we compare solutions of the NLM and the NLS.

In the subsequent treatment of the NLS we use the modal decomposition for its analysis. Since the linear part of the NLS (302) is the differential operator $-\mathrm{i}\gamma_{(\nu)}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]$ with constant coefficients, the corresponding eigenmodes are just plane waves. Consequently, here we use plane waves and the standard Fourier transform (425) instead of the Bloch modes and the Floquet–Bloch transform.

5.2.1 Source term for the NLS. First, let us show how the solution to the initial value problem (302) can be obtained as a solution to a similar differential equation with zero initial data and a source term $J_{Z,\zeta}(\mathbf{r},t)$ based on $h_\zeta(\beta\mathbf{r})$, $\zeta=\pm$. This form of the solution would be consistent with the form of the NLM (3). The general form of such a nonlinear equation with a source is provided by (56). Hence, in the case of (302) the relevant evolution equation with a source is

$$\partial_t V_\zeta = -\mathrm{i}\zeta\gamma_{(\nu)}[-\mathrm{i}\zeta\vec{\nabla}_{\mathbf{r}}]V_\zeta + \alpha_\pi p_\zeta^{[\sigma]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]\big(V_\zeta^2 V_{-\zeta}\big) - J_{Z,\zeta}(\mathbf{r},t),\ V_\zeta = 0 \text{ for } t\le 0 \quad (310)$$

and we want to find the source $J_{Z,\zeta}(\mathbf{r},t)$ so that the solution $V_\zeta(\mathbf{r},t)$ to (310) would be equal to $Z_\zeta(\mathbf{r},t)$ for $t\ge\tau_0/\varrho$. The final form of the desired source $J_{Z,\zeta}(\mathbf{r},t)$ is provided by the formula (313), and it is constructed as follows. We begin with picking up a smooth real-valued function $\psi(\tau)$ having the same properties as the function defined by (195) and (191), namely

$$0\le\psi(\tau)\le 1,\ \psi(\tau)=0,\ \tau\le 0;\ \psi(\tau)=1,\ \tau\ge\tau_0>0 \quad (311)$$

Then taking the functions $Z_\zeta(\mathbf{r},t)$ which solve problem (302) we introduce

$$V_\zeta(\mathbf{r},t)=\psi(\varrho t)Z_\zeta(\mathbf{r},t) \quad (312)$$

We can readily verify that V_ζ is a solution of equation (310) with

$$J_{Z,\zeta} = -\varrho\psi'(\varrho t)Z_\zeta - \alpha_\pi(\psi-\psi^3)p_\zeta^{[\sigma]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]\big(Z_\zeta^2 Z_{-\zeta}\big) \quad (313)$$

Evidently, in view of (311) and (312) we have

$$J_{Z,\zeta}(\mathbf{r},t)=0 \text{ when } t\ge\tau_0/\varrho \text{ or } t\le 0,\ \zeta=\pm \quad (314)$$

$$V_\zeta(\mathbf{r},t)=Z_\zeta(\mathbf{r},t) \text{ when } t\ge\tau_0/\varrho \quad (315)$$

Notice that the equalities (312), (315) and (314) establish the relation between the NLS as the initial value problem (302) and the NLS (310) with a source term. According to (315) for $t\ge\tau_0/\varrho$ the definition (58) can be rewritten in the form

$$\begin{gathered}\mathbf{U}_{Z,n_0}^{\mathrm{dir}}(\mathbf{r},t)=\tfrac{1}{(2\pi)^d}\int_{[-\pi,\pi]^d}\Psi(\eta)\\ \left[\hat{V}_+\Big(\boldsymbol{\eta},t\Big)\tilde{\mathbf{G}}_{+,n_0}(\mathbf{r},\mathbf{k}_*+\boldsymbol{\eta})+\hat{V}_-\Big(\boldsymbol{\eta},t\Big)\tilde{\mathbf{G}}_{-,n_0}(\mathbf{r},-\mathbf{k}_*-\boldsymbol{\eta})\right]\mathrm{d}\boldsymbol{\eta}\end{gathered} \quad (316)$$

where V_ζ is the solution of (310).

Notice that, similarly to solutions of the NLM, the solution Z_ζ, $\zeta=\pm$ of the NLS (302) admits the following expansion

$$Z_\zeta = Z_\zeta^{(0)}+\alpha Z_\zeta^{(1)}+\alpha^2 Z_\zeta^{(2)}+\cdots,\ \zeta=\pm \quad (317)$$

In the expansion (317) the term $Z_\zeta^{(0)}(\mathbf{r},t)$ is evidently a solution of equation (302) with $\alpha=0$, namely

$$\partial_t Z_\zeta^{(0)} = -\mathrm{i}\zeta\gamma_{(\nu)}[-\mathrm{i}\zeta\nabla_{\mathbf{r}}]Z_\zeta^{(0)},\ Z_\zeta^{(0)}(0)=h_{\zeta,\beta}(\mathbf{r}),\ h_{\zeta,\beta}(\mathbf{r})=h_\zeta(\beta\mathbf{r}) \quad (318)$$

and can be interpreted as the *linear response* corresponding to (302). Its Fourier transform (425) satisfies

$$\hat{Z}_{\zeta}^{(0)}(\boldsymbol{\eta}, t) = \exp(-\mathrm{i}\zeta\gamma_{(\nu)}(\zeta\eta)t)\hat{h}_{\zeta,\beta}(\boldsymbol{\eta}),\ \hat{h}_{\zeta,\beta}(\boldsymbol{\eta}) = \beta^{-d}\hat{h}_{\zeta}\left(\frac{1}{\beta}\boldsymbol{\eta}\right),\ \zeta = \pm \tag{319}$$

with the symbol $\gamma_{(\nu)}(\boldsymbol{\eta})$ being defined by (200).

The function $Z_{\zeta}^{(1)}$ in the expansion (317) is the FNLR of (302) and, based on (302) and (317), one can verify that the FNLR $Z_{\zeta}^{(1)}$, $\zeta = \pm$, solves the following initial value problem

$$\partial_t Z_{\zeta}^{(1)} = -\zeta\mathrm{i}\gamma_{(\nu)}[-\mathrm{i}\zeta\vec{\nabla}_{\mathbf{r}}]Z_{\zeta}^{(1)} + \alpha_{\pi} p_{\zeta}^{[\sigma]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]\big(Z_{\zeta}^{(0)2}Z_{-\zeta}^{(0)}\big),\ Z_{\zeta}^{(1)}(0) = 0 \tag{320}$$

Using (315) and (317) we then obtain the following expansion for the function $V_{\zeta}(\mathbf{r}, t)$

$$V_{\zeta}(\mathbf{r}, t) = \psi(\varrho t)Z_{\zeta}(\mathbf{r}, t) = \psi(\varrho t)Z_{\zeta}^{(0)}(\mathbf{r}, t) + \alpha\psi(\varrho t)Z_{\zeta}^{(1)}(\mathbf{r}, t) + \alpha^2\psi(\varrho t)Z_{\zeta}^{(2)}(\mathbf{r}, t) + \cdots \tag{321}$$

or, in other words

$$V_{\zeta}(\mathbf{r}, t) = V_{\zeta}^{(0)}(\mathbf{r}, t) + \alpha V_{\zeta}^{(1)}(\mathbf{r}, t) + \alpha^2 V_{\zeta}^{(2)}(\mathbf{r}, t) + \cdots \tag{322}$$

$$V_{\zeta}^{(m)}(\mathbf{r}, t) = \psi(\varrho t)Z_{\zeta}^{(m)}(\mathbf{r}, t),\ m = 1, 2\ldots \tag{323}$$

where $V_{\zeta}^{(0)}(\mathbf{r}, t)$ and $V_{\zeta}^{(1)}(\mathbf{r}, t)$ are, respectively, the linear and the first nonlinear responses corresponding to the exact solution $V_{\zeta}(\mathbf{r}, t)$ of the NLS (310) with the source (313).

To find the modal representation $\hat{V}_{\zeta}^{(0)}$ for the linear response $V_{\zeta}^{(0)}$ as defined by (322) and (323) we use (319) which implies

$$\hat{V}_{\zeta}^{(0)}(\boldsymbol{\eta}, t) = \hat{v}_{\zeta}^{(0)}(\boldsymbol{\eta}, \tau)\mathrm{e}^{-\mathrm{i}\zeta\gamma_{(\nu)}(\zeta\eta)t} = \psi(\tau)\hat{h}_{\zeta,\beta}(\boldsymbol{\eta})\mathrm{e}^{-\mathrm{i}\zeta\gamma_{(\nu)}(\zeta\eta)t},\ \tau = \varrho t \tag{324}$$

Comparing the above expression for $\hat{V}_{\zeta}^{(0)}(\boldsymbol{\eta}, t)$ and the coefficient $\tilde{U}_{\bar{n}}^{(0)}(\zeta\mathbf{k}_* + \boldsymbol{\eta}, t)$ determined by (207) and (189) we observe that

$$\Psi(\boldsymbol{\eta})\hat{V}_{\zeta}^{(0)}(\boldsymbol{\eta}, t) = \tilde{U}_{\bar{n}}^{(0)}(\zeta\mathbf{k}_* + \boldsymbol{\eta}, t) \tag{325}$$

and for $\beta \to 0$ according to (209)

$$\hat{V}_{\zeta}^{(0)}(\boldsymbol{\eta}, t) = \tilde{U}_{\bar{n}}^{(0)}(\zeta\mathbf{k}_* + \boldsymbol{\eta}, t) + O(\beta^{N_\Psi}) = \tilde{u}_{n_0,\zeta}^{(0)}(\zeta\mathbf{k}_* + \boldsymbol{\eta}, \tau)\mathrm{e}^{-\mathrm{i}\zeta\gamma_{(\nu)}(\zeta\eta)t} + O(\beta^{N_\Psi}) \tag{326}$$

where N_Ψ can be taken as large as we please, and the term $O(\beta^{N_\Psi})$ comes from the cut-off function Ψ (see the remark in section 4.1.2).

The first nonlinear response (FNLR) $V_{\zeta}^{(1)}$ is the term (323) with $m = 1$ in the expansion (322). Though the FNLR $V_{\zeta}^{(1)}$ is already described by (323), it is useful to derive a differential equation with a source for $V_{\zeta}^{(1)}$ based on (310) and (313). Notice that (313) and (317) imply

$$J_{Z,\zeta} = -\varrho\psi'(\varrho t)\big[Z_{\zeta}^{(0)}(\mathbf{r}, t) + \alpha Z_{\zeta}^{(1)}\big] - \alpha_{\pi}(\psi - \psi^3)p_{\zeta}^{[\sigma]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]\big(\big(Z_{\zeta}^{(0)}\big)^2 Z_{-\zeta}^{(0)}\big) + O(\alpha^2) \tag{327}$$

The equation for $V_{\zeta}^{(1)}$ can be obtained from (310) and (327) by collecting terms proportional to α, yielding

$$\partial_t V_{\zeta}^{(1)} = -\mathrm{i}\zeta\gamma_{(\nu)}[-\mathrm{i}\zeta\vec{\nabla}_{\mathbf{r}}]V_{\zeta}^{(1)} + \frac{\alpha_{\pi}}{\alpha}p_{\zeta}^{[\sigma]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]\big(V_{\zeta}^{(0)2}V_{-\zeta}^{(0)}\big) - J_{Z,\zeta}^{(1)} \tag{328}$$

with the source

$$J_{Z,\zeta}^{(1)}(\mathbf{r}, t) = -\varrho\psi'(\varrho t)Z_{\zeta}^{(1)} - \frac{\alpha_{\pi}}{\alpha}(\psi(\varrho t) + \psi^3(\varrho t))p_{\zeta}^{[\sigma]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]\big(Z_{\zeta}^{(0)2}Z_{-\zeta}^{(0)}\big) \tag{329}$$

Consequently, the Fourier transform of $J_{Z,\zeta}^{(1)}$ is given by

$$\hat{J}_{Z,\zeta}^{(1)}(\boldsymbol{\eta},t) = -\varrho\psi'(\varrho t)\hat{Z}_{\zeta}^{(1)}(\boldsymbol{\eta},t) - \frac{\alpha_{\pi}}{\alpha}(\psi-\psi^3)p_{\zeta}^{[\sigma]}\big(\widehat{Z_{\zeta}^{(0)2}Z_{-\zeta}^{(0)}}\big)(\boldsymbol{\eta},t) \tag{330}$$

where $Z_{\zeta}^{(0)}(\boldsymbol{\eta},t)$ and $\hat{Z}_{\zeta}^{(1)}(\boldsymbol{\eta},t)$ are defined respectively by (319) and (320).

We can use now the equation (328) together with (323) to find that

$$\hat{V}_{\zeta}^{(1)}(\boldsymbol{\eta},t) = \frac{\alpha_{\pi}}{\alpha}\int_0^t \exp\{-\mathrm{i}\zeta\gamma_{(\nu)}(\zeta\boldsymbol{\eta})(t-t_1)\}\psi^3(\varrho t)p_{\zeta}^{[\sigma]}\widehat{Z_{\zeta}^{(0)2}Z_{-\zeta}^{(0)}}(\boldsymbol{\eta},t_1)\,\mathrm{d}t_1$$
$$-\int_0^t \exp\{-\mathrm{i}\zeta\gamma_{(\nu)}(\zeta\boldsymbol{\eta})(t-t_1)\}\hat{J}_{Z,\zeta}^{(1)}(\boldsymbol{\eta},t_1)\,\mathrm{d}t_1 \tag{331}$$

Analogously to (150) it is convenient to single out a slow time factor $\hat{v}_{\zeta}(\boldsymbol{\eta},\tau)$ of the modal amplitude $\hat{V}_{\zeta}(\boldsymbol{\eta},t)$ defined by

$$\hat{V}_{\zeta}(\boldsymbol{\eta},t) = \hat{v}_{\zeta}(\boldsymbol{\eta},\tau)\mathrm{e}^{-\mathrm{i}\zeta\gamma_{(\nu)}(\zeta\boldsymbol{\eta})t},\ \tau=\varrho t,\ \hat{v}_{\zeta}(\boldsymbol{\eta},\tau) = \hat{v}_{\zeta}^{(0)}(\boldsymbol{\eta},\tau)+\alpha\hat{v}_{\zeta}^{(1)}(\boldsymbol{\eta},\tau)+\cdots \tag{332}$$

Then (331) implies, after the substitution $\boldsymbol{\eta}=\beta\mathbf{q}$,

$$\hat{v}_{\zeta}^{(1)}(\beta\mathbf{q},\tau) = \frac{3\beta^{-d}}{\varrho}\int_0^{\tau}\int_{\mathbf{q}'''+\mathbf{q}''+\mathbf{q}'=\mathbf{q}} \exp\left\{\mathrm{i}\Phi^{(\nu)}(\vec{\zeta}_0,\beta\vec{q})\frac{\tau_1}{\varrho}\right\}\psi^3(\tau_1)$$
$$\times p_{\zeta}^{[\sigma]}(\beta\vec{q}^{\,\star})\hat{h}_{\zeta}(\mathbf{q}')\hat{h}_{\zeta}(\mathbf{q}'')\hat{h}_{-\zeta}(\mathbf{q}''')\,\mathrm{d}\mathbf{q}'\mathrm{d}\mathbf{q}''\mathrm{d}\tau_1 - \hat{v}_{\zeta}^{(1)}\big(J_Z^{(1)},\beta\mathbf{q},\tau\big) \tag{333}$$

where

$$\hat{v}_{\zeta}^{(1)}\big(J_Z^{(1)},\beta\mathbf{q},\tau\big) = \frac{1}{\varrho}\int_0^{\tau}\exp\left\{\mathrm{i}\zeta\gamma_{(\nu)}(\zeta\beta\mathbf{q})\frac{\tau_1}{\varrho}\right\}\hat{J}_{Z,\zeta}^{(1)}\left(\beta\mathbf{q},\frac{\tau_1}{\varrho}\right)\mathrm{d}\tau_1,\ \zeta=\pm \tag{334}$$

with $\vec{\zeta}_0$ and $\Phi^{(\nu)}(\vec{\zeta}_0,\beta\vec{q})$ being defined by (253) and (254), respectively. Comparing with (180) we see that

$$\hat{v}_{\zeta}^{(1)}\big(J_Z^{(1)},\beta\mathbf{q},\tau\big) = \tilde{u}_{\zeta,n_0}^{(1)}(\mathbf{J}_1;\mathbf{k},\tau)+O\left(\frac{\beta^{\nu+1}}{\varrho^2}\right)$$

Comparing the equality (333) for the NLS with the interaction integral (277) for the NLM we establish the following relation between them:

$$\hat{v}_{\zeta}^{(1)}(\boldsymbol{\eta},\tau) = -3I_{\bar{n},\zeta,\zeta,-\zeta}^{(\sigma)}(\zeta\mathbf{k}_*+\boldsymbol{\eta},\tau) - \hat{v}_{\zeta}^{(1)}\big(J_Z^{(1)},\boldsymbol{\eta},\tau\big) \tag{335}$$

5.3 *Relating the NLS and the NLM*

Using (180) and (185) we obtain that the part of the linear response $\tilde{u}_{\zeta,n_0}^{(1,0)}$ of the NLM originating from $\mathbf{J}^{(1)}$ is

$$\tilde{u}_{\zeta,n_0}^{(1)}(\mathbf{J}^{(1)};\zeta\mathbf{k}_*+\boldsymbol{\eta},\tau) = \frac{1}{\varrho}\int_0^{\tau}\exp\left\{-\mathrm{i}\zeta\gamma_{(\nu)}(\zeta\boldsymbol{\eta})\frac{\tau_1}{\varrho}\right\}\hat{J}_{Z,\zeta}^{(1)}\left(\boldsymbol{\eta},\frac{\tau_1}{\varrho}\right)\mathrm{d}\tau_1 \tag{336}$$

Therefore, from (335), (334) and (280) we readily obtain

$$\tilde{u}_{\bar{n}_0}^{(1,0)}(\zeta\mathbf{k}_*+\boldsymbol{\eta},t) = \hat{v}_{\zeta}^{(1)}(\eta,\tau)+O\left(\frac{\beta^{\nu+1}}{\varrho^2}\right)+O\left(\frac{\beta^{\nu-1}}{\varrho}\right)+O(\varrho)O(|\mathbf{U}^{(1)}|) \tag{337}$$

Using (326) we obtain that

$$\hat{V}_{\zeta}^{(0)}(\boldsymbol{\eta},t)+\alpha\hat{V}_{\zeta}^{(1)}(\boldsymbol{\eta},t) = \tilde{U}_{\zeta,n_0}^{(0)}(\mathbf{k}_*+\boldsymbol{\eta},t)+\alpha\tilde{U}_{\zeta,n_0}^{(1)}(\mathbf{k}_*+\boldsymbol{\eta},t)+O(\beta^{\nu-1})+O(\varrho) \tag{338}$$

Notice that in the equality (338) and below we omit the error terms which include $O(\beta^{N_\Psi})$ with a large N_Ψ since such terms are absorbed by larger terms in the relevant expressions. Note that (338) and (30) imply the fulfilment of (173) with $N_1 = 0$, $N_2 = \sigma = \nu - 2$, $N_3 = 0$ and also imply (62) and (63).

In the following section we explain the origin of the additional terms in the ENLS with $\nu = 2$ and $\nu = 4$. After matching these terms with the FNLR of the NLM, better error estimates are derived similarly to the above derivation for the classical NLS with $\nu = 2$.

5.4 *Bidirectional waves and four-mode coupling*

In this section we assume that the condition (6) for the electric permittivity $\varepsilon(\mathbf{r})$ holds, implying, in particular, the property of complex conjugation (145) for the corresponding eigenmodes and allowed using (38). Recall now that for the excitation current $\mathbf{J}$ to be real-valued its modal coefficients $\tilde{j}_{\zeta,n_0}(\mathbf{k}_*, t)$ must satisfy the relations (164). In other words, if a mode $(\zeta, n_0, \zeta\mathbf{k}_*)$ is in the modal composition of $\mathbf{J}$ with an amplitude $j_{\zeta,n_0}(\mathbf{k}_*, t)$ then the mode $(-\zeta, n_0, -\zeta\mathbf{k}_*)$ is there as well with amplitude $[\tilde{j}_{\zeta,n_0}(\mathbf{k}_*, t)]^*$. Evidently, for real-valued currents the modes in their modal compositions are always presented in pairs

$$\uparrow n_0, \mathbf{k}_* \downarrow = \{(1, n_0, \mathbf{k}_*), (-1, n_0, -\mathbf{k}_*)\} = \cup_{\zeta=\pm 1}(\zeta, n_0, \zeta\mathbf{k}_*) \tag{339}$$

and, in view of (2) and (140), every such a pair involves modes $\tilde{\mathbf{G}}_{1,n}(\mathbf{r}, \mathbf{k}_*)$ and $\tilde{\mathbf{G}}_{-1,n}(\mathbf{r}, -\mathbf{k}_*)$ having the same frequency $\omega_{n_0}(\mathbf{k}_*) = \omega_{n_0}(-\mathbf{k}_*)$, the same group velocity $\omega'_{n_0}(\mathbf{k}_*)$ and complex conjugate amplitudes. We refer to the modal pairs (339) as *doublets*.

Observe that (2), (140) and (145) imply also that modes involved in another modal pair $\{(1, n_0, -\mathbf{k}_*), (-1, n_0, \mathbf{k}_*)\}$ are

$$\tilde{\mathbf{G}}_{1,n_0}(\mathbf{r}, -\mathbf{k}_*) = [\tilde{\mathbf{G}}_{-1,n_0}(\mathbf{r}, \mathbf{k}_*)]^* \quad \text{and} \quad \tilde{\mathbf{G}}_{-1,n_0}(\mathbf{r}, \mathbf{k}_*) = [\tilde{\mathbf{G}}_{1,n_0}(\mathbf{r}, -\mathbf{k}_*)]^* \tag{340}$$

and that they have the frequency $\omega_{n_0}(\mathbf{k}_*)$ and the group velocity $\omega'_{n_0}(-\mathbf{k}_*) = -\omega'_{n_0}(\mathbf{k}_*)$. Hence, evidently the two doublets $\uparrow n_0, \mathbf{k}_* \downarrow$ and $\uparrow n_0, -\mathbf{k}_* \downarrow$ involve the complex conjugate eigenmodes (340) of the same frequency $\omega_{n_0}(\mathbf{k}_*)$ and they have opposite group velocities $\pm\omega'_{n_0}(\mathbf{k}_*)$. Consequently, a wave composed of the modal quadruplet

$$\begin{aligned} \Uparrow n_0, \mathbf{k}_* \Downarrow &= \uparrow n_0, \mathbf{k}_* \downarrow \cup \uparrow n_0, -\mathbf{k}_* \downarrow \\ &= \{(1, n_0, \mathbf{k}_*), (-1, n_0, -\mathbf{k}_*), (1, n_0, -\mathbf{k}_*), (-1, n_0, \mathbf{k}_*)\} \end{aligned} \tag{341}$$

is *bidirectional* since wavepackets corresponding to its two constitutive doublets $\uparrow n_0, \mathbf{k}_* \downarrow$ and $\uparrow n_0, -\mathbf{k}_* \downarrow$ propagate with opposite group velocities $\pm\omega'_{n_0}(\mathbf{k}_*)$. Such a bidirectional wave can be directly excited, i.e. excited through the linear mechanism, by a current composed of the quadruplet $\Uparrow n_0, \mathbf{k}_* \Downarrow$ in (341). The four modes in the quadruplet $\Uparrow n_0, \mathbf{k}_* \Downarrow$ are coupled through relatively strong nonlinear interactions and this is the subject of this section. It turns out that the quadruplet $\Uparrow n_0, \mathbf{k}_* \Downarrow$ in (341) is the only modal quadruple which is generic with respect to the excitation frequency variations and which has relatively strong (of the order $O(\varrho)$) nonlinear interaction between its doublets; we name and will refer to it as a *bidirectional quadruplet*. Note that two doublets $\uparrow n_0, \mathbf{k}_* \downarrow$ and $\uparrow n_1, \mathbf{k}_* \downarrow$ corresponding to different bands, if directly excited, also have the same order of interaction $O(\varrho)$ between them, but one has to use a special pair of exactly matched excitation carrier frequencies $\omega_1 = \omega_{n_0}(\mathbf{k}_*)$ and $\omega_2 = \omega_{n_1}(\mathbf{k}_*)$ to directly excite them, whereas only one frequency $\omega_{n_0}(\mathbf{k}_*)$ is required to excite the bidirectional quadruplet.

In the preceding sections we have considered the case where only one doublet $\uparrow n_0, \mathbf{k}_* \downarrow$ was excited. This case describes the situation when the directly excited wave propagates only in one direction. When only the excitation frequencies $\omega = \omega_{n_0}(\mathbf{k}_*)$ and the wave numbers $\pm\mathbf{k}_*$

are fixed, a general time-harmonic excitation produces waves propagating in both directions. As we have shown in section 3.2 the magnitude of their non-FM interaction is $O(\varrho)$, and when we take into account corrections of the order $O(\varrho)$ they have to be considered. For a discussion of multimode interactions see section 1.2.

Now let us introduce a bidirectional excitation current of the form (158) where the amplitude $\tilde{j}^{(0)}_{\zeta,n_0}(\mathbf{k}, t)$ is defined by the following (slightly more general than (188)) formula

$$\tilde{j}^{(0)}_{\zeta,n_0}(\mathbf{k}, t) = -\varrho\psi_0(\varrho t)\beta^{-d}\left[\Psi(\mathbf{k}-\zeta\mathbf{k}_*)\hat{h}^{+}_{\zeta}\left(\frac{\mathbf{k}-\zeta\mathbf{k}_*}{\beta}\right) + \Psi(\mathbf{k}+\zeta\mathbf{k}_*)\hat{h}^{-}_{\zeta}\left(\frac{\mathbf{k}+\zeta\mathbf{k}_*}{\beta}\right)\right] \quad (342)$$

with functions $\hat{h}^{\pm}_{\zeta}(\mathbf{q})$ having the same properties as $\hat{h}_{\zeta}(\mathbf{q})$ in (188). The excitation current defined by (342) directly excites the four modes for the bidirectional quadruplet $\Uparrow n_0, \mathbf{k}_*\Downarrow$ (341) with the linear response modal amplitudes $\tilde{U}^{(0)}_{\zeta,n_0}(\pm\mathbf{k}_* + \boldsymbol{\eta}, t)$, $\zeta = \pm 1$. If we introduce

$$\mathbf{k}^{\vartheta}_{*} = \vartheta\mathbf{k}_*, \ \vartheta = \pm \quad (343)$$

then the bidirectional quadruplet $\Uparrow n_0, \mathbf{k}_*\Downarrow$ is composed of the two doublets $\uparrow n_0, \mathbf{k}^{\vartheta}_{*}\downarrow$, $\vartheta = \pm$, with the group velocities $\vartheta\omega'_{n_0}(\mathbf{k}_*)$, and, importantly, the nonlinear interactions between those doublets are non-FM interactions satisfying relation (243). Using relation (243), the phase matching relation (71), with $\mathbf{k}' = \zeta'\mathbf{k}^{\vartheta}_{*}$, $\mathbf{k}'' = \zeta''\mathbf{k}^{\vartheta}_{*}$, $\mathbf{k}''' = \zeta'''\mathbf{k}^{\vartheta}_{*}$, and arguments similar to those used to derive (225), (227) and (243), we conclude that the triad of modes from the quadruplet $\Uparrow n_0, \mathbf{k}_*\Downarrow$ affects only the mode from this quadruplet with the quasimomentum

$$(\zeta' + \zeta'' + \zeta''')\mathbf{k}^{\vartheta}_{*} = -\zeta\mathbf{k}^{\vartheta}_{*} = \zeta\mathbf{k}^{-\vartheta}_{*} \quad (344)$$

This condition selects, from formally possible 4^4 interactions, only a few significant ones. Namely, condition (344) implies that a pair of three numbers $\zeta', \zeta'', \zeta'''$ has the same sign as $-\zeta$, we set

$$\zeta' = \zeta'' = -\zeta, \ \zeta''' = \zeta, \ \vec{\zeta}_{0,\times} = (\zeta, -\zeta, -\zeta, \zeta) \quad (345)$$

and two more cases are similar to the above. The corresponding interaction wavevectors $\vec{k}$ are given by the formula:

$$\vec{k}_{*,\times,-} = (-\mathbf{k}_*, \mathbf{k}_*, \mathbf{k}_*, \mathbf{k}_*), \ \vec{k}_{*,\times,+} = (\mathbf{k}_*, -\mathbf{k}_*, -\mathbf{k}_*, -\mathbf{k}_*) \quad (346)$$

The interaction integral (211) with $\vec{\zeta}$, $\vec{k}$ satisfying (345) and (346) describes the impact of the triad of waves from the forward-propagating doublet onto the modal coefficient of the backward-propagating doublet, note that this interaction has three modes (two of them forward-propagating) with the same ζ (that is, in the same band) and one mode (forward-propagating) with the opposite ζ (that is, in the opposite band).

The analysis of the interaction integral (211) with $\vec{\zeta}$ given by (345) and $\vec{k}$ in a vicinity of $\vec{k}_{*,\times}$ determined by (346) is similar to the analysis of the integral (237) in the FM case where $\vec{k}$ is in a vicinity of $\vec{k}_*$ determined by (232). We obtain that similarly to (277) and (278) the principal part of the non-FM interaction integral is given by the formula

$$\beta^d I^{(0)}_{\bar{n},-\zeta,-\zeta,\zeta}(-\zeta\mathbf{k}_* + \beta\mathbf{q}, \tau) = \frac{1}{\varrho}\int_0^{\tau}\int_{\mathbf{q}'''+\mathbf{q}''+\mathbf{q}'=\mathbf{q}} \exp\left\{\mathrm{i}\Phi^{(\nu)}(\vec{\zeta}, \beta\vec{q}^{\,0})\frac{\tau_1}{\varrho}\right\}\breve{Q}_{\vec{n}_0}(\vec{\zeta}_{0,\times}\vec{k}_{*,\times,-})$$
$$\times\psi^3(\tau_1)\hat{h}^{+}_{-\zeta}(\mathbf{q}')\hat{h}^{+}_{-\zeta}(\mathbf{q}'')\hat{h}^{+}_{\zeta}(\mathbf{q}'''(\vec{q}))\,\mathrm{d}\mathbf{q}'\mathrm{d}\mathbf{q}''\mathrm{d}\tau_1 \quad (347)$$

where we take $\sigma = 0$ since we do not need to take into account the higher approximations of this integral which is already ϱ times smaller than the FM interactions. Similarly to (277) the error of approximation of the integral (237) is given by the formula

$$\beta^d[I_{\bar{n},-\zeta,-\zeta,\zeta}(-\zeta\mathbf{k}_* + \beta\mathbf{q}, \tau) - I^{(0)}_{\bar{n},-\zeta,-\zeta,\zeta}(-\zeta\mathbf{k}_* + \beta\mathbf{q}, \tau)] = O(\beta) \quad (348)$$

The four modal amplitudes $\tilde{U}_{\zeta,n_0}(\zeta\mathbf{k}_*^{\pm}+\boldsymbol{\eta},t)$ are approximated by the Fourier transforms of the four NLS solutions $Z_\zeta^\vartheta(\mathbf{r},t)$. We write the corresponding system in the most general case of complex-valued excitation currents and then we will discuss the reduction of the system when the excitation is real. To approximate the non-FM terms we include the coupling terms $\delta_{\times,\zeta}^{+}((Z_{-\zeta}^{-})^2 Z_\zeta^{-})$ in the equation for Z_ζ^{+} and $\delta_{\times,\zeta}^{-}((Z_{-\zeta}^{+})^2 Z_\zeta^{+})$ in the equation for Z_ζ^{-} where

$$\delta_{\times,\zeta}^{-}=3\breve{Q}_{\vec{n}_0}(\vec{\zeta}_{0,\times}\vec{k}_{*,\times,-}),\ \delta_{\times,\zeta}^{+}=3\breve{Q}_{\vec{n}_0}(\vec{\zeta}_{0,\times}\vec{k}_{*,\times,+}) \tag{349}$$

We take in (98) and (99) $\nu=4$ to take into account terms of order $O(\beta^2)$ which are comparable with the terms coming from (347) in the case of classical nonlinear Schrödinger scaling. We obtain two pairs of coupled equations for Z_ζ^{+} and Z_ζ^{-}, $\zeta=\pm$:

$$\{\partial_t+\zeta\mathrm{i}\gamma_{(4)}(-\mathrm{i}\vartheta\zeta\nabla_{\mathbf{r}})\}Z_\zeta^\vartheta+\alpha\delta_{\times,\zeta}^{\vartheta}\left(\left(Z_{-\zeta}^{-\vartheta}\right)^2 Z_\zeta^{-\vartheta}\right)=+\alpha_\pi p_\zeta^{\vartheta,[2]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]\left(Z_{-\zeta}^\vartheta\left(Z_\zeta^\vartheta\right)^2\right) \tag{350}$$

with the initial data

$$Z_\zeta^\vartheta(\mathbf{r},t)|_{t=0}=h_\zeta^\vartheta(\mathbf{r}),\quad \zeta=\pm,\quad \vartheta=\pm \tag{351}$$

We have

$$\tilde{U}_{\zeta,n_0}(\zeta\mathbf{k}_*^\vartheta+\eta,t)=\hat{Z}_\zeta^\vartheta(\boldsymbol{\eta},t)+O(\beta^3)+O(\varrho)+O(\alpha) \tag{352}$$

If we take into account in (353) the terms that originate from the first-order correction to the susceptibility we obtain the system

$$\begin{aligned}&\{\partial_t+\zeta\mathrm{i}\gamma_{(4)}(-\mathrm{i}\vartheta\zeta\nabla_{\mathbf{r}})\}Z_\zeta^\vartheta+\alpha\delta_{\times,\zeta}^{\vartheta}\left(\left(Z_{-\zeta}^{-\vartheta}\right)^2 Z_\zeta^{-\vartheta}\right)+\alpha_\pi\delta_{1,\zeta}^\vartheta Z_\zeta^\vartheta Z_{-\zeta}^\vartheta\{\partial_t+\zeta\mathrm{i}\gamma_{(4)}(-\mathrm{i}\vartheta\zeta\nabla_{\mathbf{r}})\}Z_\zeta^\vartheta\\&=-\alpha_\pi\delta_{2,\zeta}^\vartheta\left(Z_\zeta^\vartheta\right)^2\{\partial_t-\zeta\mathrm{i}\gamma_{(3)}(\mathrm{i}\vartheta\zeta\nabla_{\mathbf{r}})\}Z_{-\zeta}^\vartheta+\alpha_\pi p_\zeta^{\vartheta,[2]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]\left(Z_{-\zeta}^\vartheta\left(Z_\zeta^\vartheta\right)^2\right)\end{aligned} \tag{353}$$

These additional terms with coefficients $\delta_{1,\zeta}^{\pm}$ are discussed in section 6. Addition of these terms and taking into account interband coupling (see section 1.4.4) improves the error term in (352) replacing $O(\varrho)$ by $O(\varrho\beta)$. The term $O(\alpha)$ in the approximation error term in (352) can be replaced by $O(\alpha\beta)$ if the fifth-order terms of the nonlinearity are taken into account (see section 1.4.6 and the 7.1.1. Note that if the initial data are real and the nonlinearity is real, then we have

$$h_{-\zeta}^\vartheta(\mathbf{r})=h_\zeta^\vartheta(\mathbf{r})^*,\quad Z_{-\zeta}^\vartheta(\mathbf{r},t)=\left[Z_\zeta^\vartheta(\mathbf{r},t)\right]^*,\quad \vartheta=\pm,\quad \zeta=\pm \tag{354}$$

and, consequently, we can apply (38). Namely, we exclude $\zeta=-1$ and (353) is reduced to the system (125) of two scalar equations (similar equations are known as *coupled mode equations*) for Z_+^ϑ, $\vartheta=\pm$.

5.5 *Representation of solutions in the space domain*

Here we derive formula (42) providing a representation for $\mathbf{U}_Z^{\mathrm{dir}}(\mathbf{r},t)$ in the space domain in terms of $Z_\zeta(\mathbf{r},t)$. The principal part of the approximate solution $\mathbf{U}_Z(\mathbf{r},t)$ is determined by (58) in terms of its Bloch modal coefficients and the Fourier transform $\hat{Z}_\zeta(\boldsymbol{\eta},t)$ of the solution of the NLS as follows:

$$\mathbf{U}_Z^{\mathrm{dir}}(\mathbf{r},t)=\mathbf{U}_{Z_+}(\mathbf{r},t)+\mathbf{U}_{Z_-}(\mathbf{r},t) \tag{355}$$

where

$$\mathbf{U}_{Z_\zeta}(\mathbf{r},t)=\frac{\beta^d}{(2\pi)^d}\int_{\mathbb{R}^d}\Psi(\beta\mathbf{q})\hat{Z}_\zeta(\beta\mathbf{q},t)\tilde{\mathbf{G}}_{\zeta,n_0}(\mathbf{r},\zeta\mathbf{k}_*+\beta\mathbf{q})\mathrm{d}\mathbf{q},\ \zeta=\pm \tag{356}$$

where $\hat{Z}_\zeta(\boldsymbol{\eta}, t)$ is the Fourier transform of the solution $Z_\zeta(\mathbf{r}, t)$ of (36), (37) or (47). According to (307) $\beta^d \hat{Z}_\zeta(\beta\mathbf{q}, t) = \hat{Z}_{\beta,\zeta}(\mathbf{q}, t)$ where $Z_{\beta,\zeta}$ is a solution of (308) which regularly depends on β. According to (142)

$$\tilde{\mathbf{G}}_{\zeta,n_0}(\mathbf{r}, \zeta\mathbf{k}_* + \beta\mathbf{q}) = \hat{\mathbf{G}}_{\zeta,n_0}(\mathbf{r}, \zeta\mathbf{k}_* + \beta\mathbf{q})\mathrm{e}^{\mathrm{i}(\zeta\mathbf{k}_* + \beta\mathbf{q})\cdot\mathbf{r}} \tag{357}$$

where $\hat{\mathbf{G}}_{\zeta,n_0}(\mathbf{r}, \mathbf{k})$ is a 1-periodic function of $\mathbf{r}$. We then approximate $\hat{\mathbf{G}}_{\zeta,n_0}(\mathbf{r}, \zeta\mathbf{k}_* + \beta\mathbf{q})$ by its Taylor polynomial of degree σ

$$\hat{\mathbf{G}}_{\zeta,n_0}(\mathbf{r}, \zeta\mathbf{k}_* + \beta\mathbf{q}) = \mathbf{p}_{\zeta,g,\sigma}(\mathbf{r}, \beta\mathbf{q}) + O(\beta^{\sigma+1}), \ \sigma + 1 \le \nu \tag{358}$$

For $\sigma = 2$, the polynomials have the form

$$\mathbf{p}_{\zeta,g,2}(\mathbf{r}, \beta\mathbf{q}) = \hat{\mathbf{G}}_{\zeta,n_0}(\mathbf{r}, \zeta\mathbf{k}_*) + \beta\hat{\mathbf{G}}'_{\zeta,n_0}(\mathbf{r})(\mathbf{q}) + \frac{1}{2}\beta^2\hat{\mathbf{G}}''_{\zeta,n_0}(\mathbf{r})(\mathbf{q}^2) \tag{359}$$

with coefficients defined in terms of the tensors:

$$\hat{\mathbf{G}}'_{\zeta,n_0}(\mathbf{q}) = \nabla_{\mathbf{k}}\hat{\mathbf{G}}_{\zeta,n_0}(\mathbf{r}, \zeta\mathbf{k}_*) \cdot \mathbf{q}, \quad \hat{G}''_{\zeta,n_0}(\mathbf{q}^2) = \nabla_{\mathbf{k}}^2\hat{\mathbf{G}}_{+,n_0}(\mathbf{r}, \zeta\mathbf{k}_*)\mathbf{q}\cdot\mathbf{q} \tag{360}$$

Hence,

$$\mathbf{U}_{Z_\zeta}(\mathbf{r}, t) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \beta^d\Psi(\beta\mathbf{q})\hat{Z}_\zeta(\beta\mathbf{q}, t)\mathbf{p}_{\zeta,g,\sigma}(\mathbf{r}, \beta\mathbf{q})\mathrm{e}^{\mathrm{i}(\zeta\mathbf{k}_* + \beta\mathbf{q})\cdot\mathbf{r}}\,\mathrm{d}\mathbf{q} + O(\beta^{\sigma+1}) \tag{361}$$

Assuming that $\hat{Z}_\zeta(\beta\mathbf{q}, t)$ decays sufficiently fast as $|\mathbf{q}| \to \infty$, i.e.

$$\beta^d|\hat{Z}_\zeta(\beta\mathbf{q}, t)| \le C_N(1 + |\mathbf{q}|)^{-N_\Psi} \tag{362}$$

with large enough N_Ψ (this assumption follows from the regularity of $Z_{\beta,\zeta}$ according to (307)), we obtain

$$\mathbf{U}_{Z_\zeta}(\mathbf{r}, t) = \frac{\mathrm{e}^{\mathrm{i}\zeta\mathbf{k}_*\cdot\mathbf{r}}}{(2\pi)^d}\int_{\mathbb{R}^d} \hat{Z}_\zeta(\mathbf{q}, t)\mathbf{p}_{\zeta,g,\sigma}(\mathbf{r}, \mathbf{q})\mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{r}}\,\mathrm{d}\mathbf{q} + O(\beta^{\sigma+1}) \tag{363}$$

Note that

$$\mathbf{p}_{\zeta,g,\sigma}(\mathbf{r}, \mathbf{q})\hat{Z}_\zeta(\mathbf{q}, t) = \mathbf{p}_{\zeta,g,\sigma}(\widehat{\mathbf{r}, -\mathrm{i}\nabla_{\mathbf{r}}})Z_\zeta(\mathbf{r}, t) \tag{364}$$

where $\mathbf{p}^{[\sigma]}(\mathbf{r}, -\mathrm{i}\nabla_{\mathbf{r}})$ is a differential operator with the polynomial symbol $\mathbf{p}^{[\sigma]}(\mathbf{r}, \mathbf{q})$ with coefficients that depend on $\mathbf{r}$, see (430). Hence, since the integral in (363) is the inverse Fourier transform, we obtain that

$$\mathbf{U}_{Z_\zeta}(\mathbf{r}, t) = \mathrm{e}^{\mathrm{i}\zeta\mathbf{k}_*\cdot\mathbf{r}}\mathbf{p}_{\zeta,g,\sigma}(\mathbf{r}, -\mathrm{i}\nabla_{\mathbf{r}})Z_\zeta(\mathbf{r}, t) + O(\beta^{\sigma+1}) \tag{365}$$

For $\sigma = 2$ we obtain

$$\mathbf{U}_{Z_\zeta}(\mathbf{r}, t) = \mathbf{U}^0_{Z_\zeta}(\mathbf{r}, t) + \mathbf{U}^1_{Z_\zeta}(\mathbf{r}, t) + \mathbf{U}^2_{Z_\zeta}(\mathbf{r}, t) + O(\beta^3) \tag{366}$$

According to (359) and (360) the dominant term is

$$\mathbf{U}^0_{Z_\zeta}(\mathbf{r}, t) = \mathrm{e}^{\mathrm{i}\zeta\mathbf{k}_*\cdot\mathbf{r}}\hat{\mathbf{G}}_{\zeta,n_0}(\mathbf{r}, \zeta\mathbf{k}_*)Z_\zeta(\mathbf{r}, t) = \tilde{\mathbf{G}}_{\zeta,n_0}(\mathbf{r}, \zeta\mathbf{k}_*)Z_\zeta(\mathbf{r}, t) \tag{367}$$

Note that this term has the form which is used as an ansatz for the solution of the NLM in [12]. The first-order correction takes the form

$$\begin{aligned}\mathbf{U}^1_{Z_\zeta}(\mathbf{r}, t) &= -\mathrm{i}\mathrm{e}^{\mathrm{i}\zeta\mathbf{k}_*\cdot\mathbf{r}}\nabla_{\mathbf{k}}\hat{\mathbf{G}}_{\zeta,n_0}(\mathbf{r}, \zeta\mathbf{k}_*)\cdot\nabla_{\mathbf{r}}Z_\zeta(\mathbf{r}, t)\\ &= -\mathrm{i}\mathrm{e}^{\mathrm{i}\zeta\mathbf{k}_*\cdot\mathbf{r}}\left[\partial_{r_1}Z_\zeta(\mathbf{r}, t)\partial_{k_1}\hat{\mathbf{G}}_{\zeta,n_0}(\mathbf{r}, \zeta\mathbf{k}_*) + \cdots + \partial_{r_d}Z_\zeta(\mathbf{r}, t)\partial_{k_d}\hat{\mathbf{G}}_{\zeta,n_0}(\mathbf{r}, \zeta\mathbf{k}_*)\right]\end{aligned} \tag{368}$$

The second-order correction is

$$\mathbf{U}^2_{Z_\zeta}(\mathbf{r}, t) = -\mathrm{e}^{\mathrm{i}\zeta\mathbf{k}_*\cdot\mathbf{r}}\frac{1}{2}\sum_{j,l=1}^{d}\partial_{k_j}\partial_{k_l}\hat{\mathbf{G}}_{\zeta,n_0}(\mathbf{r}, \zeta\mathbf{k}_*)\partial_{r_j}\partial_{r_l}Z_\zeta(\mathbf{r}, t) \tag{369}$$

Similarly, we can write higher terms of the expansion in β. Note that since $Z_\zeta(\mathbf{r}, 0) = h_\zeta(\beta\mathbf{r})$ we have

$$\mathbf{U}^1_{Z_\zeta}(\mathbf{r}, t) = O(\beta), \mathbf{U}^2_{Z_\zeta}(\mathbf{r}, t) = O(\beta^2) \tag{370}$$

6. The first nonlinear response and its time-harmonic approximation

To improve the accuracy of the NLM–NLS approximation we have to improve the term $O(\varrho)$ in (337) originating from the time-harmonic approximation. If we use a certain modification of the ENLS it modifies $\hat{v}^{(1)}_\zeta(\mathbf{q}, \tau)$ so that the term $O(\varrho)$ can be replaced by $O(\varrho\beta)$. To this end we consider in this section an approximation of the FNLR $\mathbf{U}^{(1)}$ with $\tilde{\mathbf{u}}^{(1)}_{\bar{n}}$ determined by (158), (160), (162) and (167) in terms of the susceptibility $\chi^{(3)}_D(\mathbf{r}; \omega_1, \omega_2, \omega_3)$ defined by (156). This approximation is based on the following asymptotic formulas for $\varrho \to 0$

$$\tilde{\mathbf{u}}^{(1)}_{\bar{n}}(\mathbf{r}, \mathbf{k}, \tau) = \tilde{\mathbf{u}}^{(1,0)}_{\bar{n}}(\mathbf{r}, \mathbf{k}, \tau) + O(\varrho)O(|\tilde{\mathbf{u}}^{(1)}|) \tag{371}$$

$$\tilde{\mathbf{u}}^{(1,0)}_{\bar{n}}(\mathbf{r}, \mathbf{k}, \tau) = \tilde{u}^{(1,0)}_{\bar{n}}(\mathbf{k}, \tau)\tilde{\mathbf{G}}_{\bar{n}}(\mathbf{r}, \mathbf{k})$$

$$\tilde{\mathbf{u}}^{(1)}_{\bar{n}}(\mathbf{r}, \mathbf{k}, \tau) = \tilde{\mathbf{u}}^{(1,0)}_{\bar{n}}(\mathbf{r}, \mathbf{k}, \tau) + \tilde{\mathbf{u}}^{(1,1)}_{\bar{n}}(\mathbf{r}, \mathbf{k}, \tau) + O(\varrho^2)O(|\tilde{\mathbf{u}}^{(1)}|) \tag{372}$$

as well as higher order expansions which are derived below.

6.1 *The first nonlinear response as a causal integral*

Here we recast the FNLR for an almost time-harmonic excitation as a causal convolution integral. In the next subsections we derive asymptotic expansions of the causal nonlinearity which do not involve explicit time convolution integration. Notice first that the solution of (3) can be written in the form

$$\mathbf{U}(t) = \int_0^t \mathrm{e}^{-\mathrm{i}\mathbf{M}(t-t')}[\alpha\mathcal{F}_{\mathrm{NL}}(\mathbf{U}) - \mathbf{J}]\,\mathrm{d}t' \tag{373}$$

which, after a change of variables $\tau = \varrho t$, yields the following expression for the $\bar{n}$th mode:

$$\tilde{U}_{\bar{n}}\left(\mathbf{k}, \frac{\tau}{\varrho}\right) = \frac{1}{\varrho}\int_0^{\tau/\varrho} \mathrm{e}^{-\mathrm{i}\omega_{\bar{n}}(\mathbf{k})\frac{(\tau-\tau_1)}{\varrho}}[\widetilde{\mathcal{F}_{\mathrm{NL}}(\mathbf{U})}_{\bar{n}}(\mathbf{k}, \tau_1) - \tilde{J}_{\bar{n}}(\mathbf{k}, \tau_1)]\,\mathrm{d}\tau_1 \tag{374}$$

According to (152) and (153)

$$\mathcal{F}_{\mathrm{NL}}(\mathbf{U}) = \mathcal{F}^{(3)}_{\mathrm{NL}}(\mathbf{U}) + \alpha\mathcal{F}^{(5)}_{\mathrm{NL}}(\mathbf{U}) + \alpha^2\mathcal{F}^{(7)}_{\mathrm{NL}}(\mathbf{U}) + \cdots \tag{375}$$

and, as follows from (152) and (143), the modal coefficient $\widetilde{\mathcal{F}_{\mathrm{NL}}(\mathbf{U})}_{\bar{n}}(\mathbf{k}, \tau_1)$ is given by

$$\begin{aligned}\widetilde{\mathcal{F}_{\mathrm{NL}}(\mathbf{U})}_{\bar{n}}(\mathbf{k}, \tau_1) &= (\widetilde{\mathcal{F}_{\mathrm{NL}}(\mathbf{U})}(\cdot, \tau_1), \tilde{\mathbf{G}}_{\bar{n}}(\cdot, \mathbf{k}))_{\mathcal{H}} \\ &= \int_{[0,1]^d} \widetilde{\mathcal{F}_{\mathrm{NL}}(\mathbf{U})}(\mathbf{r}, \tau_1, \mathbf{k}) \cdot \sigma_\varepsilon(\mathbf{r})\tilde{\mathbf{G}}^*_{\bar{n}}(\mathbf{r}, \mathbf{k})\,\mathrm{d}\mathbf{r}\end{aligned} \tag{376}$$

In particular, using formula (217) of [17] this coefficient can be rewritten as

$$\begin{aligned}\widetilde{\mathcal{F}_{\mathrm{NL}}(\mathbf{U})}_{\bar{n}}(\mathbf{k}, \tau_1) &= \int_{[0,1]^d} \mathbf{S}_D\widetilde{(\mathbf{r}, \tau_1; \mathbf{D})} \cdot \nabla \times \tilde{\mathbf{G}}_{B,\bar{n}}(\mathbf{r}, \mathbf{k})^*\,\mathrm{d}\mathbf{r} \\ &= -\mathrm{i}\zeta\omega_{\bar{n}}(\mathbf{k}) \int_{[0,1]^d} \mathbf{S}_D\widetilde{(\mathbf{r}, t; \mathbf{D})} \cdot \tilde{\mathbf{G}}_{D,\bar{n}}(\mathbf{r}, \mathbf{k})^*\,\mathrm{d}\mathbf{r}\end{aligned} \tag{377}$$

This form of coefficients may be useful in computations since it uses only the $\mathbf{D}$-component of $\tilde{\mathbf{G}}_{\bar{n}}$. In fact, this specific form is not important in our analysis. The cubic part $\mathbf{S}_D^{(3)}$ of $\mathbf{S}_D(\mathbf{r}, t; \mathbf{D})$ can be written in the form of the causal integral (154) where the tensors $\mathbf{R}_D^{(3)}$ are smooth for $t_1, t_2, t_3 \geq 0$ and satisfy the inequality

$$\left|\mathbf{R}_D^{(3)}(\mathbf{r}; t_1, t_2, t_3)\right| \leq C \exp[-c_0(t_1 + t_2 + t_3)] \tag{378}$$

for some $c_0 > 0$. It is convenient to introduce, similarly to (183), operators $\mathbf{S}_{D,B}^{(m)}$ and $\mathbf{R}_{D,B}^{(m)}$ that act in six-dimensional $(\mathbf{D}, \mathbf{B})$-space; they act on the $\mathbf{D}$-components of $\tilde{\mathbf{u}}_{\bar{n}'}^{(0)}$ and take values in the $\mathbf{B}$-component. For example, when $m = 3$

$$\mathbf{R}_{D,B}^{(3)} \vdots \mathbf{U}_1\mathbf{U}_2\mathbf{U}_3 = \begin{bmatrix} \mathbf{0} \\ \mathbf{R}_D^{(3)} \vdots \mathbf{D}_1\mathbf{D}_2\mathbf{D}_3 \end{bmatrix}, \quad \mathbf{U}_j = \begin{bmatrix} \mathbf{D}_j \\ \mathbf{B}_j \end{bmatrix} \tag{379}$$

Using the above notation we get

$$\mathbf{S}_{D,B}^{(3)}(\mathbf{r}, t; \mathbf{U}) = \int_{-\infty}^{t}\int_{-\infty}^{t}\int_{-\infty}^{t} \mathbf{R}_{D,B}^{(3)}(\mathbf{r}; t - t_1, t - t_2, t - t_3) \vdots \mathbf{U}(\mathbf{r}, t_1) \times \mathbf{U}(\mathbf{r}, t_2)\mathbf{U}(\mathbf{r}, t_3)\, \mathrm{d}t_1\, \mathrm{d}t_2\, \mathrm{d}t_3 \tag{380}$$

and from (375) together with (153) we obtain the expansion

$$\mathcal{F}_{\mathrm{NL}}(\mathbf{U})(\mathbf{r}, t) = \sum_{i=0}^{\infty} \alpha^i \nabla \times \mathbf{S}_{D,B}^{(2i+3)}(\mathbf{r}, t; \mathbf{U}) \tag{381}$$

To evaluate the integral in (374) for $\varrho \ll 1$ we, as is commonly done in nonlinear optics, represent the term $\widetilde{\mathcal{F}_{\mathrm{NL}}(\mathbf{U})}(\mathbf{k}, \tau_1)$ using the frequency-dependent susceptibilities. The FNLR has a form similar to (373)

$$\tilde{U}_{\bar{n}}^{(1)}(\mathbf{k}, t') = \int_0^{t'} \mathrm{e}^{\mathrm{i}\omega_{\bar{n}}(\mathbf{k})(t'-t)} \left[\widetilde{\mathcal{F}_{\mathrm{NL}}(\mathbf{U}^{(0)})}_{\bar{n}}^{(3)}(\mathbf{k}, t) - \tilde{j}_{\bar{n}}^{(1)}(\mathbf{k}, t)\right] \mathrm{d}t \tag{382}$$

By (375), (376) and (154) the modal coefficient of the FNLR is given by

$$\tilde{\mathbf{u}}_{\bar{n}}^{(1)}(\mathbf{r}, \mathbf{k}, \varrho t') = \mathrm{e}^{\mathrm{i}\omega_{\bar{n}}(\mathbf{k})t'} \widetilde{\mathbf{S}_{D,B}^{(3)}(\cdot; \mathbf{D}^{(0)})}_{\bar{n}}(\mathbf{r}, \mathbf{k}, t') - \tilde{\mathbf{u}}_{\bar{n}}^{(1)}(\mathbf{J}_1; \mathbf{k}, \varrho t') \tag{383}$$

where $\tilde{\mathbf{u}}_{\bar{n}}^{(1)}(\mathbf{J}_1; \mathbf{k}, \tau)$ is defined by (180). Using expansion (52) and the convolution formula (see [17]) we write the Floquet-Bloch transform of (380) in terms of slowly varying coefficients $\tilde{u}_{\bar{n}}$ defined by (150) and obtain

$$\begin{aligned}\widetilde{\mathbf{S}_{D,B}^{(3)}(\cdot; \mathbf{U}^{(0)})}(\mathbf{r}, \mathbf{k}, t') = {} & \sum_{\bar{n}', \bar{n}'', \bar{n}'''} \frac{1}{(2\pi)^{2d}} \int_0^{t'} \mathrm{e}^{\{-\mathrm{i}\sum_{j=1}^{3} \omega_{\bar{n}^{(j)}}(\mathbf{k}^{(j)})t\}} \int_{\substack{[-\pi,\pi]^{2d} \\ \mathbf{k}'+\mathbf{k}''+\mathbf{k}'''=\mathbf{k}}} \int_{-\infty}^{t}\int_{-\infty}^{t}\int_{-\infty}^{t} \\ & \times \mathrm{e}^{\{-\mathrm{i}\sum_{j=1}^{3} \omega_{\bar{n}^{(j)}}(\mathbf{k}^{(j)})(t_j - t)\}} \mathbf{R}_{D,B}^{(3)}(\mathbf{r}; t - t_1, t - t_2, t - t_3) \vdots \prod_{j=1}^{3} \tilde{\mathbf{G}}_{\bar{n}^{(j)}}(\mathbf{k}^{(j)}, \mathbf{r}) \\ & \times \Psi^3(\mathbf{k})\psi(\varrho t_j)\hat{h}_{\zeta^{(j)}}\left(\frac{\mathbf{k}^{(j)} - \zeta^{(j)}\mathbf{k}_*}{\beta}\right) \mathrm{d}t_1\, \mathrm{d}t_2\, \mathrm{d}t_3\, \mathrm{d}\mathbf{k}'\, \mathrm{d}\mathbf{k}''\, \mathrm{d}t\end{aligned} \tag{384}$$

This integral is simplified in section 6.2.

6.2 *Time-harmonic approximation*

In this section we introduce an expansion yielding powers ϱ^{l_1} in the structured power series (169). In particular, we obtain (371) and (372). If $\mathbf{U}_n^{(0)}(\mathbf{r},t)$ has the form (198) of a slowly modulated wavepacket with $\varrho \ll 1$ then the time-harmonic approximation can be applied. It effectively substitutes the integration with respect to time in the causal integral (384) with expressions involving frequency dependent susceptibilities. This approximation is based on the Fourier transform $\chi_D^{(3)}$ of $\mathbf{R}^{(3)}$ swith respect to the time variables as in (156), and it is constructed as follows. Below we approximate $\tilde{u}_{\bar{n}}^{(1)}(\mathbf{k},\tau)$ in (383) by $\tilde{u}_{\bar{n}}^{(1,0)}(\mathbf{k},t)$, which is defined by (181) and (184), and then estimate the error providing the higher order terms as well. Using the Taylor approximation of $\psi(\varrho t_j)$ in (384) we get:

$$\psi(\varrho(t_j-t)+\varrho t)=\sum_{l=0}^{N_1}\frac{(-1)^l\varrho^l}{l!}\psi^{(l)}(\varrho t)(t-t_j)^l+\psi_{(N_1)} \tag{385}$$

with

$$|\psi_{(N_1)}|\le C_{N_1}\varrho^{N_1+1}|t_j-t|^{N_1+1} \tag{386}$$

Substituting the Taylor polynomial approximation for $\psi(\varrho t_j)$ into (384) we obtain a sum of terms

$$\int_{-\infty}^{t}\int_{-\infty}^{t}\int_{-\infty}^{t}\mathbf{R}_{D,B}^{(3)}(\mathbf{r};t-t_1,t-t_2,t-t_3)\,\mathrm{e}^{\{-\mathrm{i}\sum_{j=1}^{3}\omega_{\bar{n}^{(j)}}(\mathbf{k}^{(j)})(t_j-t)\}}$$
$$\times\prod_{j=1}^{3}\frac{(-1)^{l_j}\varrho^{l_j}}{l_j!}(t-t_j)^{l_j}\psi^{(l_j)}(\varrho t)\,\mathrm{d}t_1\,\mathrm{d}t_2\,\mathrm{d}t_3=\chi_{D,B,\bar{l}}^{(3)}(\mathbf{r};\omega_{\bar{n}'}(\mathbf{k}'),\omega_{\bar{n}''}(\mathbf{k}''),\omega_{\bar{n}'''}(\mathbf{k}''')) \tag{387}$$

where $\bar{l}=(l_1,l_2,l_3)$ and we use the following notation:

$$\chi_{D,B,\bar{l}}^{(3)}(\mathbf{r};\omega_1,\omega_2,\omega_3)=\int_0^\infty\int_0^\infty\int_0^\infty\mathbf{R}_{D,B}^{(3)}(\mathbf{r};t_1,t_2,t_3)\mathrm{e}^{\mathrm{i}(\omega_1t_1+\omega_2t_2+\omega_3t_3)}\prod_{j=1}^{3}\frac{(-1)^{l_j}}{l_j!}t_j^{l_j}\,\mathrm{d}t_1\,\mathrm{d}t_2\,\mathrm{d}t_3 \tag{388}$$

Evidently the susceptibility $\chi_{D,B}^{(3)}$ defined by (183) and (156) equals $\chi_{D,B,0}^{(3)}$. A straightforward computation shows that the quantity $\chi_{D,B,\bar{l}}^{(3)}(\mathbf{r};\omega_1,\omega_2,\omega_3)$ defined by (388) equals the partial derivative of $\chi_{D,B}^{(3)}$, defined by (183) and (156), with respect to its frequency arguments, namely

$$\chi_{D,B,\bar{l}}^{(3)}(\mathbf{r};\omega_1,\omega_2,\omega_3)=\frac{\mathrm{i}^{|\bar{l}|}}{l_1!l_2!l_3!}\frac{\partial^{|\bar{l}|}\chi_{D,B}^{(3)}(\mathbf{r};\omega_1,\omega_2,\omega_3)}{\partial\omega_1^{l_1}\partial\omega_2^{l_2}\partial\omega_3^{l_3}} \tag{389}$$

Substituting (387) into (384) we obtain

$$\nabla\times\widetilde{\mathbf{S}_{D,B}^{(3)}(\cdot;\mathbf{U}^{(0)})}_{\bar{n}}(\mathbf{k},t)=\frac{1}{(2\pi)^{2d}}\sum_{|\bar{l}|=0}^{N_1}\varrho^{|\bar{l}|}\sum_{\bar{n}',\bar{n}'',\bar{n}'''}\mathbf{Q}(\widetilde{(\mathbf{u}^{(0)})^3})(\vec{n},\bar{l},\mathbf{k},\varrho t)+O(\varrho^{N_1+1}) \tag{390}$$

where

$$\mathbf{Q}(\widetilde{(\mathbf{u}^{(0)})^3})(\vec{n},\bar{l},\mathbf{k},\varrho t)=\int_0^t\mathrm{e}^{\mathrm{i}\phi_{\bar{n}}(\vec{k})t}\int_{\substack{[-\pi,\pi]^{2d}\\ \mathbf{k}'+\mathbf{k}''+\mathbf{k}'''=\mathbf{k}}}\int_{[0,1]^d}\nabla$$
$$\times\left[\chi_{D,B,\bar{l}}^{(3)}(\mathbf{r};\omega_{\bar{n}'}(\mathbf{k}'),\omega_{\bar{n}''}(\mathbf{k}''),\omega_{\bar{n}'''}(\mathbf{k}''')):\prod_{j=1}^{3}\tilde{\mathbf{G}}_{\bar{n}^{(j)}}(\mathbf{k}^{(j)},\mathbf{r})\right]$$
$$\cdot\,\tilde{\mathbf{G}}_{\bar{n}}(\mathbf{k},\mathbf{r})\mathrm{d}\mathbf{r}\,\Psi^3(\vec{q})\prod_{j=1}^{3}\psi^{(l_j)}(\varrho t')\hat{h}_{\zeta^{(j)}}\left(\frac{\mathbf{k}^{(j)}-\zeta^{(j)}\mathbf{k}_*}{\beta}\right)\mathrm{d}\mathbf{k}'\mathrm{d}\mathbf{k}''\mathrm{d}t' \tag{391}$$

Hence, we obtain from (384) the formula for the modal coefficients:

$$u_{\vec{n}}^{(1,\bar{l})}(\mathbf{k},\tau)=\frac{1}{\varrho}\sum_{\vec{n}',\vec{n}'',\vec{n}'''}\int_0^{\tau}\mathrm{e}^{\mathrm{i}\phi_{\vec{n}}(\vec{k})\frac{\tau_1}{\varrho}}\int_{\substack{[-\pi,\pi]^{2d}\\ \mathbf{k}'+\mathbf{k}''+\mathbf{k}'''=\mathbf{k}}}$$

$$\breve{Q}_{\vec{n},\bar{l}}(\vec{k})\Psi^3(\vec{q})\prod_{j=1}^{3}\psi^{(l_j)}(\tau_1)\beta^{-d}\hat{h}_{\zeta^{(j)}}((\mathbf{k}^{(j)}-\zeta^{(j)}\mathbf{k}_*)/\beta)\,\mathrm{d}\mathbf{k}'\mathrm{d}\mathbf{k}''\mathrm{d}\tau_1 \quad (392)$$

where $\breve{Q}_{\vec{n},\bar{l}}(\vec{k})$ is given by the following formula (similar to (179) with $\chi_D^{(3)}$ being replaced with its frequency derivative $\chi_{D,B,\bar{l}}^{(3)}$)

$$\breve{Q}_{\vec{n},\bar{l}}(\vec{k})=\frac{1}{(2\pi)^{2d}}\left(\begin{bmatrix}\mathbf{0}\\ \nabla\times\chi_{G,D,B,\bar{l}}^{(3)}\end{bmatrix},\tilde{\mathbf{G}}_{\vec{n}}(\mathbf{r},\mathbf{k})\right)_{\mathcal{H}}$$

$$\chi_{G,D,B,\bar{l}}^{(3)}=\chi_{D,B,\bar{l}}^{(3)}(\omega_{\bar{n}'}(\mathbf{k}'),\omega_{\bar{n}''}(\mathbf{k}''),\omega_{\bar{n}'''}(\mathbf{k}'''))\tilde{\mathbf{G}}_{D,\bar{n}'}(\mathbf{r},\mathbf{k}')\tilde{\mathbf{G}}_{D,\bar{n}''}(\mathbf{r},\mathbf{k}'')\tilde{\mathbf{G}}_{D,\bar{n}'''}(\mathbf{r},\mathbf{k}''') \quad (393)$$

Observe that the modal susceptibility $\breve{Q}_{\vec{n}}(\vec{k})=\breve{Q}_{\vec{n},0}(\vec{k})$ defined by (179) is symmetric with respect to permutations of $\zeta^{(j)}$, $\mathbf{k}^{(j)}$ whereas $\breve{Q}_{\vec{n},\bar{l}}(\vec{k})$ defined by (393) with nonsymmetric $\bar{l}$ is not.

Hence, taking the term at α in (381) and using (390) we obtain that (175) holds, i.e.

$$\tilde{u}_{\vec{n}}^{(1)}(\mathbf{k},\tau)=\tilde{u}_{\vec{n}}^{(1,0)}(\mathbf{k},\tau)+\sum_{l=1}^{N_1}\varrho^l\tilde{u}_{\vec{n}}^{(1,l)}(\mathbf{k},\tau)+O(\varrho^{N_1+1})O(|\tilde{\mathbf{u}}^{(1)}|) \quad (394)$$

The dominant term $\tilde{u}_{\vec{n}}^{(1,0)}(\mathbf{k},t\varrho)$ is given by (176). In particular, we obtain the formula

$$\beta^d\tilde{u}_{\vec{n}}^{(1)}(\zeta\mathbf{k}_*+\beta\mathbf{q},\tau)=\beta^d\tilde{u}_{\vec{n}}^{(1,0)}(\zeta\mathbf{k}_*+\beta\mathbf{q},\tau)+O(\varrho)O(|\mathbf{u}^{(1)}|) \quad (395)$$

Remark Though here we consider the time-harmonic approximation of the third-order term $\nabla_B\times\mathbf{S}_{D,B}^{(3)}(\mathbf{r},t;\mathbf{U}^{(0)})$ in (381), similar time-harmonic approximations are applicable to terms $\nabla\times\mathbf{S}_{D,B}^{(m)}(\mathbf{r},t;\mathbf{U}^{(0)})$ of an arbitrary order m of homogeneity.

6.3 *The first-order correction to the susceptibility*

For $|\bar{l}|=1$, the term $\tilde{u}_{\vec{n}}^{(1,1)}$ involves three expressions of the form (392) with $l_1+l_2+l_3=1$. Hence

$$\tilde{u}_{\vec{n}}^{(1,1)}(\mathbf{r},\mathbf{k},\tau)=\sum_{l_1+l_2+l_3=1}\tilde{u}_{\vec{n}}^{(1,\bar{l})}=\frac{1}{\varrho}\sum_{l_1+l_2+l_3=1}\int_0^{\tau}\mathrm{e}^{\{-\mathrm{i}\sum_{j=1}^{3}\omega_{\bar{n}^{(j)}}(\mathbf{k}^{(j)})\frac{\tau_1}{\varrho}\}}\int_{\substack{[-\pi,\pi]^{2d}\\ \mathbf{k}'+\mathbf{k}''+\mathbf{k}'''=\mathbf{k}}}\breve{Q}_{\vec{n},\bar{l}}(\vec{k})$$

$$\times\Psi^3(\vec{q})\beta^{-d}\psi^2(\tau_1)\psi'(\tau_1)\prod_{j=1}^{3}\hat{h}_{\zeta^{(j)}}\left(\frac{1}{\beta}(\mathbf{k}^{(j)}-\zeta^{(j)}\mathbf{k}_*)\right)\mathrm{d}\mathbf{k}'\,\mathrm{d}\mathbf{k}''\,\mathrm{d}\tau_1 \quad (396)$$

The integral (392) is similar to (211). Consequently, the principal order contribution there is given by the FM terms such that $\vec{n}$ satisfies (213) and (224). The integral in (392) with such $\vec{n}$ takes a form similar to (238), i.e.

$$\beta^d\tilde{u}_{\vec{n}}^{(1,\bar{l})}(\zeta\mathbf{k}_*+\beta\mathbf{q},\tau)=\frac{1}{\varrho}\int_0^{\tau}\int_{\mathbf{q}'+\mathbf{q}''+\mathbf{q}'''=\mathbf{q}}\exp\left\{\mathrm{i}\phi_{\vec{n}_0}(\vec{\zeta}\vec{k}_*+\beta\vec{q})\frac{\tau_1}{\varrho}\right\}\partial_{\tau_1}\psi(\tau_1)\psi^2(\tau_1)$$

$$\times\Psi^3((\beta\vec{q}))\breve{Q}_{\vec{n},\bar{l}}(\vec{\zeta}_0\mathbf{k}_*\hat{h}_{\zeta}(\mathbf{q}'')\hat{h}_{-\zeta}(\mathbf{q}''')\mathrm{d}\mathbf{q}'\,\mathrm{d}\mathbf{q}''\,\mathrm{d}\tau_1 \quad (397)$$

with the only difference that $\partial_{\tau_1}\psi(\tau_1)\psi^2(\tau_1)$ replaces $\psi^3(\tau_1)$ (see section 8.3 where similar terms are derived from the ENLS). The integral in (397) can be treated similarly to (238).

The next, first-order, approximation is given by the formula

$$\tilde{u}_{\bar{n}}^{(1)}(\zeta\mathbf{k}_* + \beta\mathbf{q}, \tau) = \tilde{u}_{\bar{n}}^{(1,0)}(\zeta\mathbf{k}_* + \beta\mathbf{q}, \tau) + \varrho\tilde{u}_{\bar{n}}^{(1,1)}(\zeta\mathbf{k}_* + \beta\mathbf{q}, \tau) + O(\varrho^2)O(|\tilde{\mathbf{u}}^{(1)}|)$$

where $\tilde{u}_{\bar{n}}^{(1,1)}$is given by (396) and (397), which yields (372).

Remark Note that the above expansions for the FNLR integral (382) can be applied to the integral (374) for the exact solution. The formula (384) holds with $\tilde{\mathbf{U}}_{\bar{n}^{(j)}}^{(0)}$ replaced by $\tilde{\mathbf{U}}_{\bar{n}^{(j)}}$ and $\psi(\varrho t_j)\hat{h} = \tilde{u}_{\bar{n}^{(j)}}^{(0)}(\varrho t_j)$ replaced by $\tilde{u}_{\bar{n}^{(j)}}(\varrho t_j)$. We have, similarly to (394),

$$\tilde{\mathbf{u}}_{\bar{n}}(\mathbf{r}, \mathbf{k}, \tau) = \tilde{\mathbf{u}}_{\bar{n},0}(\mathbf{r}, \mathbf{k}, \tau) + \sum_{l=1}^{N_1} \varrho^l \tilde{\mathbf{u}}_{\bar{n},l}(\mathbf{r}, \mathbf{k}, \tau) + O(\varrho^{N_1+1})O(|\tilde{\mathbf{u}}|) \tag{398}$$

For the modal coefficients we get, similarly to (396),

$$\begin{aligned}\tilde{u}_{\bar{n}}(\mathbf{k}, \tau) = \frac{1}{\varrho} \sum_{l=0}^{N_1} \sum_{l_1+l_2+l_3=l} \sum_{\bar{n}',\bar{n}'',\bar{n}'''} \varrho^l \int_0^\tau \int_{\substack{[-\pi,\pi]^{2d} \\ \mathbf{k}'+\mathbf{k}''+\mathbf{k}'''=\mathbf{k}}} \mathrm{e}^{\mathrm{i}\phi_{\bar{n}}(\vec{k})\frac{\tau_1}{\varrho}} \breve{Q}_{\bar{n},\bar{l}}(\vec{k}) \partial_{\tau_1}^l [\tilde{u}_{\bar{n}'}(\mathbf{k}', \tau_1)\tilde{u}_{\bar{n}''}(\mathbf{k}'', \tau_1) \\ \times \tilde{u}_{\bar{n}'''}(\mathbf{k}''', \tau_1)]\, \mathrm{d}\mathbf{k}'\mathrm{d}\mathbf{k}''\mathrm{d}\tau_1 - \frac{1}{\varrho}\int_0^\tau \tilde{j}_{\bar{n}}(\mathbf{k}, \tau_1)\, \mathrm{d}\tau_1 + O(\varrho^{N_1})\end{aligned} \tag{399}$$

where $\breve{Q}_{\bar{n},\bar{l}}(\vec{k})$ are the same as in (393), $\tau \leq \tau_*$ with a fixed τ_*. In the above formula we assumed that the nonlinearity involves only the cubic term, but in the general case a similar formula involves series with respect to powers α^m with coefficients which are $(2m+1)$-linear tensors. ♦

7. Beyond the FNLR

In this section we answer two questions (i) Why do we use in representation (58) the exact solution $Z(\mathbf{r}, t)$ of the NLS rather than its FNLR approximation? (ii) Why do we impose a restriction $\frac{\tau_*}{\varrho} \leq \frac{\alpha_0}{\alpha}$ in (15), and what can be expected on longer time intervals? These two subjects happen to be related.

7.1 *Advantages of using an exact solution of the NLS*

So, why use in (58) the exact solution $Z(\mathbf{r}, t)$ of the NLS rather than its FNLR approximation? There are at least two advantages. First, the FNLR based on a solution of a linear nonhomogeneous equation of the form (320) may lead to functions that grow linearly as $t \to \infty$, whereas an exact soliton-type solution of the NLS is bounded for all t. Note though that in the time interval $t \leq \tau_*/\varrho \leq \alpha_0/\alpha$ both functions are bounded, therefore to see the difference one has to consider longer time intervals. The second and more important advantage is that using in (58) the exact solution $Z(\mathbf{r}, t)$ of the NLS rather than its FNLR approximation produces a smaller approximation error. Now let us take a look at the above arguments trying to avoid technical details.

We need to use some information on solutions of the NLS. In some cases the NLS admits explicit solutions, which are regular. In many cases information on the regularity of general solutions $Z(\mathbf{r}, t)$ of NLS and ENLS is available (see [40], [41], and [7] in particular, p. 64 of [7] and references cited there concerning nonelliptic NLS and ENLS). So, providing the error estimates we assume the solutions of the NLS or ENLS to be sufficiently regular.

Namely, we assume that, in addition to (206), we have the following estimate for solutions of the NLS

$$|\hat{Z}_{\beta,\zeta}(\mathbf{q},t)|+\sum_{|\bar{l}|\le m_0}\left|\partial_{\mathbf{q}}^{\bar{l}}\hat{Z}_{\beta,\zeta}(\mathbf{q},t)\right|+\sum_{l\le N_0}\left|\partial_t^{l}\hat{Z}_{\beta,\zeta}(\mathbf{q},t)\right|\le C_{N_\Psi}(1+|\mathbf{q}|)^{-N_\Psi},\ 0\le t\le\frac{\tau_*}{\varrho} \quad (400)$$

with a large enough N_Ψ. Note that the above condition includes rescaled $\hat{Z}_{\beta,\zeta}$ which is a solution of (308). Regular dependence of $\hat{Z}_{\beta,\zeta}$ on β is consistent with the β-independent form of the above estimate. In fact, we choose the value of N_Ψ depending on the chosen order of approximation, and if we take only a few lower order terms of the approximation, the value of N_Ψ does not have to be very large. The value of N_Ψ can be recovered from the Remark in section 4.1.2. Here, for simplicity, we primarily consider the case of the classical NLS with $\nu=2$, $\sigma=0$ and the simplest ENLS with $\nu=3$, $\sigma=1$; at in section 7.1.1 when we discuss fifth-order corrections we take $\nu=4$, $\sigma=2$.

Below we show that in fact formula (58) gives a better approximation than can be seen from the FNLR. To see that we first consider a simpler case when higher order terms in expansion (153) satisfy the estimate

$$\mathbf{S}_D(\mathbf{r},t;\mathbf{D})=\mathbf{S}_D^{(3)}(\mathbf{r},t;\mathbf{D})+O(\alpha_5\alpha) \quad (401)$$

where the constant $\alpha_5\ll 1$ controls the magnitude of the next, the fifth-orders term in the expansion of the nonlinearity (in particular, if the nonlinearity in (3) is purely cubic, $\alpha_5=0$). Let

$$Z_\zeta^{[1]}=Z_\zeta^{(0)}+\alpha Z_\zeta^{(1)} \quad (402)$$

be the first-order approximation based on the linear and the first nonlinear responses to the exact solution Z_ζ of the NLS. If in (58) we replaced Z_ζ by $Z_\zeta^{[1]}$ we would obtain an approximate solution $\mathbf{U}_{Z^{[1]}}(\mathbf{r},t)$ of the NLM which satisfies the equation:

$$\begin{aligned}\partial_t\mathbf{U}_{Z^{[1]}}&=-\mathrm{i}\mathbf{M}\mathbf{U}_{Z^{[1]}}+\alpha\mathcal{F}_{\mathrm{NL}}(\mathbf{U}_{Z^{[1]}})-\mathbf{J}\\&=O(\alpha_5\alpha)+O\left(\frac{\alpha^2}{\varrho}\right)O(|\mathbf{U}^{(1)}|)+O(\beta^{\nu-1})+O(\varrho)\end{aligned} \quad (403)$$

To see the origin of the leading term $O(\frac{\alpha^2}{\varrho})O(|\mathbf{U}^{(1)}|)$ of the discrepancy in (403) let us look at the second-order term in the expansion (158). The next after the FNLR correction term $\alpha^2\mathbf{U}^{(2)}$ in the series (158) for $\alpha_5=0$ can be found by solving the equation

$$\partial_t\mathbf{U}^{(2)}=-\mathrm{i}\mathbf{M}\mathbf{U}^{(2)}+3\mathcal{F}_{\mathrm{NL}}^{(1)}(\mathbf{U}^{(0)},\mathbf{U}^{(0)},\mathbf{U}^{(1)})-\mathbf{J}^{(2)};\ \mathbf{U}^{(2)}(t)=0\quad\text{for } t\le 0 \quad (404)$$

where the form of the expression $\mathcal{F}_{\mathrm{NL}}^{(1)}(\mathbf{U}^{(0)},\mathbf{U}^{(0)},\mathbf{U}^{(1)})$ is based on the fact that $\mathcal{F}_{\mathrm{NL}}^{(1)}$ is a trilinear operator. Note that since the expression for $\mathbf{U}^{(1)}$ is itself a cubic with respect to $\mathbf{U}^{(0)}$, the next term $\mathbf{U}^{(2)}$ is quintic. Note also that $\mathbf{U}^{(1)}$ is of order $1/\varrho$ and the time interval is of the same order $1/\varrho$. Since $\mathcal{F}_{\mathrm{NL}}^{(1)}(\mathbf{U}^{(0)},\mathbf{U}^{(0)},\mathbf{U}^{(1)})$ involves frequency matched terms, we can conclude that $\mathbf{U}^{(2)}$ is of order $1/\varrho^2$, or, equivalently, $\alpha^2\mathbf{U}^{(2)}$ is of order $O(\alpha^2/\varrho)O(|\mathbf{U}^{(1)}|)$. The deciding advantage of using the exact solution Z_ζ of the NLS, as we do in (58), is making the discrepancy much smaller than (403):

$$\partial_t\mathbf{U}_Z=-\mathrm{i}\mathbf{M}\mathbf{U}_Z+\alpha\mathcal{F}_{\mathrm{NL}}(\mathbf{U}_{Z^{[1]}})-\mathbf{J}=\mathbf{J}_Z \quad (405)$$

$$\mathbf{J}_Z=O(\alpha_5\alpha^2)+[O(\alpha\beta^{\nu-1})+O(\alpha\varrho)]O(|\mathbf{U}^{(1)}|) \quad (406)$$

The terms of order $O(\alpha^2/\varrho^2)$, $O(\alpha^3/\varrho^3)$ and similar to them in the right-hand side of (404) and its higher analogues disappear in (405). The reason for this is that those terms have exactly the same form in the Floquet–Bloch expansion of solution of the NLM as the corresponding terms in the Fourier expansion of the solution of the NLS. Since Z_ζ satisfy the NLS exactly, these terms completely cancel in the expansion of the solution of the NLS, consequently corresponding terms completely cancel in the expansion of the solution of the NLM.

Now we provide some more details for the above considerations. We still use (40), (58) and (67) to define the approximate solution $\mathbf{U}_Z$. To make sure that the discrepancy does not include terms coming from the mismatch in the initial data in all orders of accuracy, the excitation current $\mathbf{J}$ in (3) is given by (35) where $\tilde{\mathbf{J}}_{\bar{n}}$ has the form

$$\tilde{\mathbf{J}}_{\zeta,n}(\zeta\mathbf{k}_* + \zeta\mathbf{q}, t) = \Psi(\mathbf{q})\hat{J}_{Z,\zeta}(\mathbf{q}, t)\tilde{\mathbf{G}}_{\bar{n}}(\zeta\mathbf{k}_* + \zeta\mathbf{q}), n = n_0,\ \zeta = \pm 1 \tag{407}$$

where $J_{Z,\zeta}(\mathbf{r}, t)$ is given in (313). Note that the difference between $\tilde{\mathbf{J}}_{\bar{n}}^{(1)}(\mathbf{q}, t)$ defined by (185), (163) and found by subtracting $\mathbf{J}^{(0)}$ from (407) is of order α^2 and does not affect the FNLR approximation.

Note that the term $O(\alpha\varrho)O(|\mathbf{U}^{(1)}|)$ in (405) comes from the almost time-harmonic approximation of the nonlinearity in the NLM. We consider then the NLS equations in the form (100) and (101) with the initial conditions (102). From formulas (40), (58) and (67) we define the modal coefficients $\tilde{U}_{Z,\zeta,n}(\zeta\mathbf{k}_* + \zeta\boldsymbol{\eta}, t)$ of the approximate solution $\mathbf{U}_Z(\mathbf{r}, t)$. To show that $\mathbf{U}_Z(\mathbf{r}, t)$ satisfies the NLM with a small discrepancy $\mathbf{J}_Z$ we consider equations for the Floquet–Bloch modal coefficients $\tilde{U}_{\zeta,n}(\zeta\mathbf{k}_* + \zeta\boldsymbol{\eta}, t)$ of the exact solution $\mathbf{U}(\mathbf{r}, t)$. We expand operators that enter the equations with respect to ϱ and β as we did for the FNLR. The leading part of the expansion of the equation which includes $n = n' = n'' = n''' = n_0$ of the exact solution $\mathbf{U}(\mathbf{r}, t)$ of the NLM has exactly the same form as the NLS written in terms of Fourier transform. Therefore $\tilde{u}_{Z,\zeta,n_0}(\zeta\mathbf{k}_* + \zeta\boldsymbol{\eta}, t)$ exactly satisfies this part of the equations. All remaining terms of the expansion contribute to the discrepancy. The estimates of these terms are completely similar to estimates for the FNLR. The only difference is that instead of explicitly given $\tilde{u}_{\zeta,n}^{(0)}(\zeta\mathbf{k}_* + \zeta\boldsymbol{\eta}, t)$ which was in the FNLR we have have to consider the same formulas with $\tilde{u}_{Z,\zeta,n}(\zeta\mathbf{k}_* + \zeta\boldsymbol{\eta}, t)$. The analysis is the same, but now we have to use (400) instead of (206). The analysis implies that the discrepancy is small, namely $\mathbf{J}_Z$ satisfies (406). From the estimate of the discrepancy of the equations we derive the estimate for the difference of solutions

$$\mathbf{U}(\mathbf{r}, t) - \mathbf{U}_Z(\mathbf{r}, t) = [O(\alpha_5\alpha^2) + O(\alpha\beta^{\nu-1}) + O(\alpha\varrho)]O(|\mathbf{U}^{(1)}|) \tag{408}$$

in the interval $\tau_0/\varrho \le t < \tau_*/\varrho$ (in the final part of this subsection we discuss the relevance of the restriction on the length of the time interval). Estimate (408) implies (46) and estimates in section 1.4. Note that when $\tilde{u}_Z$ is defined by the FNLR, as in (403), we would have much larger term $O(\alpha^2/\varrho)$ in addition to $O(\alpha_5\alpha^2)$. This is the main and very important advantage of using the exact solution.

7.1.1 Fifth-order corrections. If the coefficient α_5 in (401) is not small, in order to get the approximation by $\mathbf{U}_Z$ with error term $O(\alpha_5\alpha^2)$ replaced by $O(\beta\alpha^2)$, one has to take into account the fifth-order terms of $\mathcal{F}_{\mathrm{NL}}$, and include into the NLS (302) a term similar to $\alpha_\pi^2 Q_{5,\pm}|Z_\pm|^4 Z_\pm$ as in (118):

$$\partial_t Z_\zeta = -\mathrm{i}\zeta\gamma_{(4)}[-\mathrm{i}\zeta\vec{\nabla}_{\mathbf{r}}]Z_\zeta + \alpha_\pi p_\zeta^{[2]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]\big(Z_\zeta^2 Z_{-\zeta}\big) + \alpha_\pi^2 Q_{5,\zeta} Z_\zeta^3 Z_{-\zeta}^2 \tag{409}$$

The coefficient

$$Q_{5,\zeta} = 10\check{Q}_{\bar{n},5}(\vec{\zeta}_0\vec{k}_*) \tag{410}$$

is determined by the modal susceptibility of fifth-order similar to (179):

$$\nabla \times \chi_D^{(5)}(\omega_{\bar{n}'}(\mathbf{k}'), \omega_{\bar{n}''}(\mathbf{k}''), \ldots, \omega_{\bar{n}^{(5)}}(\mathbf{k}^{(5)}))\dot{:}\tilde{\mathbf{G}}_{D,\bar{n}'}(\cdot, \mathbf{k}') \cdots \tilde{\mathbf{G}}_{D,\bar{n}^{(5)}}(\cdot, \mathbf{k}^{(5)}) \tag{411}$$

$$\breve{Q}_{\bar{n},5}(\vec{k}) = \frac{1}{(2\pi)^{4d}} \times \left(\begin{bmatrix} 0 \\ \nabla \times \chi_D^{(5)} \end{bmatrix}, \tilde{\mathbf{G}}_{\bar{n}}(\cdot, \mathbf{k}) \right)_{\mathcal{H}}$$

$$\chi_D^{(5)} = \chi_D^{(5)}(\omega_{\bar{n}'}(\mathbf{k}'), \omega_{\bar{n}''}(\mathbf{k}''), \ldots, \omega_{\bar{n}^{(5)}}(\mathbf{k}^{(5)})) \dot{:} \tilde{\mathbf{G}}_{D,\bar{n}'}(\cdot, \mathbf{k}') \cdots \tilde{G}_{D,\bar{n}^{(5)}}(\cdot, \mathbf{k}^{(5)}) \tag{412}$$

with $n = n_0, \vec{\zeta}_0 = (\zeta, \zeta, \zeta, \zeta, -\zeta, -\zeta)$. The tensor $\chi_D^{(5)}$ is defined by a formula similar to (156) based on the kernel $\mathbf{R}_D^{(5)}$ that corresponds to $\mathbf{S}_D^{(5)}$ in (153).

Note that to obtain high-precision matching of initial data for the NLS and the source term for the NLM one has to use there instead of (313) the following modified source

$$J_{Z,\zeta} = -\varrho\psi'(\varrho t)Z_\zeta - \alpha_\pi(\psi - \psi^3)p_\zeta^{[\sigma]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]\big(Z_\zeta^2 Z_{-\zeta}\big) - \alpha_\pi^2 Q_{5,\zeta}(\psi - \psi^5)Z_\zeta^3 Z_{-\zeta}^2 \tag{413}$$

After the inclusion of the term $\alpha_\pi^2 Q_{5,\zeta} Z_\zeta^3 Z_{-\zeta}^2$ the approximation error of the NLS–NLM approximation, which stems from the truncation of $\mathcal{F}_{\mathrm{NL}}$, becomes $O(\beta\alpha^2)$ instead of $O(\alpha_5\alpha^2)$ and formula (408) with $\sigma = 2$, $\nu = 4$ takes the form

$$\mathbf{U}(\mathbf{r}, t) - \mathbf{U}_Z(\mathbf{r}, t) = [\alpha^2\beta + O(\alpha\beta^3) + O(\alpha\varrho)]O(|\mathbf{U}^{(1)}|) \tag{414}$$

Similarly, a more elaborate analysis shows that if we take in the ENLS $\nu = 4, \sigma = 2$, and take into account the first-order susceptibility correction as in (114) or (118) with $Q_{5,\pm}$ defined by (120) we obtain the following improved error estimate

$$\mathbf{U}(\mathbf{r}, t) - \mathbf{U}_Z(\mathbf{r}, t) = [\alpha^2\beta + O(\alpha\beta^3) + O(\alpha\varrho\beta)]O(|\mathbf{U}^{(1)}|) \tag{415}$$

Note that in the above error estimates when $\nu = 4, \sigma = 2$ we assumed that the ENLSs are constructed so that they take into account the effects of interband interactions.

7.2 *Longer time intervals*

Here we consider the case when (15) does not hold, namely $1/\varrho \gg 1/\alpha$, that is for time scales large compared with the time scale $1/\alpha$ related with the magnitude of the nonlinearity.

Still the approximate solution $\mathbf{U}_Z(\mathbf{r}, t)$, which is constructed based on the ENLS (now we take $\nu = 4, \sigma = 2$), satisfies the Maxwell equation with a high precision on a long time interval, namely

$$\partial_t \mathbf{U}_Z(\mathbf{r}, t) = -\mathrm{i}\mathbf{M}\mathbf{U}_Z(\mathbf{r}, t) + \alpha\mathcal{F}_{\mathrm{NL}}(\mathbf{U}_Z(\mathbf{r}, t)) - \mathbf{J} + \mathbf{J}_Z, \ t \le \frac{\tau_*}{\varrho}$$

$$\mathbf{J}_Z = O(\alpha^2\beta) + O(\alpha\beta^3) + O(\alpha\varrho\beta) \tag{416}$$

even when $1/\varrho \gg 1/\alpha$. The only difference between equation (3) and equation (416) is the discrepancy term $\mathbf{J}_Z$. The discrepancy is small if

$$[O(\alpha^2\beta) + O(\alpha\beta^3) + O(\alpha\varrho\beta)] \ll 1 \tag{417}$$

In this case equation (3) is satisfied by $\mathbf{U}_Z$ with a small error.

Smallness of the discrepancy $\mathbf{J}_Z$, generally speaking, implies smallness of the approximation error only on time intervals of order $1/\alpha$ or shorter. Without assumptions on the stability of the the exact solution $\mathbf{U}$ of (3) and the approximate solution $\mathbf{U}_Z$ the difference between $\mathbf{U}$ and $\mathbf{U}_Z$ can be estimated as follows

$$\mathbf{U}(\mathbf{r}, t) - \mathbf{U}_Z(\mathbf{r}, t) = O\left(\frac{1}{\alpha}\big[\mathrm{e}^{O(\frac{\alpha\tau_*}{\varrho})} - 1\big]\right)[O(\alpha^2\beta) + O(\alpha\beta^3) + O(\alpha\varrho\beta)] \tag{418}$$

Clearly, this estimate implies smallness of the difference between the solutions of equations (3) and equation (416) if α/ϱ is bounded (or if it grows at most at a logarithmic rate). If the discrepancy $\mathbf{J}_Z$ in (416) is small and we want the exact solution $\mathbf{U}(\mathbf{r}, t)$ to be close to the approximate solution $\mathbf{U}_Z(\mathbf{r}, t)$ for times much greater than α/ϱ then we have to impose some kind of a stability condition on the nonlinearity $\mathcal{F}_{\mathrm{NL}}$. More detailed analysis shows that it is sufficient to impose a stability condition on the solution Z of the NLS which serves as a basis for $\mathbf{U}_Z$. For stability results for solutions of NLS see [7], section II.4 and [42–44]. A detailed, mathematically rigorous analysis of the validity of the approximation by a stable solution of the NLS on a long time interval is done for some particular cases in [45].

8. Some technical topics

In this section for our readers' convenience we introduce some notations and discuss some technical topics instrumental to the description of solutions to the NLM.

8.1 *The Taylor formula*

Let us introduce notations related to the Taylor formula. For a function h of variables $x_1, \ldots, x_L = \mathbf{x}$ we write the Taylor formula as

$$h(\mathbf{x}+\mathbf{y}) = h(x_1 + y_1, \ldots, x_L + y_L) = h(x) + \sum_{|l|=1}^{\nu} \frac{1}{\bar{l}!} h^{[\bar{l}]}(x) y^{\bar{l}} + O(|y|^{\nu+1}) \tag{419}$$

where

$$\bar{l} = (l_1, \ldots, l_L), \quad |\bar{l}| = l_1 + \cdots + l_L, \quad y^{\bar{l}} = y_1^{l_1} \ldots y_L^{l_L}$$
$$\frac{1}{\bar{l}!} = \frac{1}{l_1! \cdots l_L!}, \quad h^{[\bar{l}]}(x) = \frac{\partial^{|\bar{l}|} h(x)}{\partial x_1^{l_1} \cdots \partial x_L^{l_L}} \tag{420}$$

We often use a shorter notation

$$\sum_{|\bar{l}|=l_0} \frac{1}{\bar{l}!} h^{[\bar{l}]}(x) y^{\bar{l}} = \frac{1}{l_0!} H^{(l_0)} \vdots (\mathbf{y}^{l_0}) \tag{421}$$

where l_0 is an integer (not a integer vector) and $H^{(l_0)}(\mathbf{y}^{l_0})$ is an l_0-linear symmetric form. For example, a symmetric cubic form can be written as:

$$H^{(3)} \vdots \mathbf{uvw} = \sum_{j_1, j_2, j_3 = 1}^{L} H^{(3)}_{j_1, j_2, j_3} u_{j_1} v_{j_2} w_{j_3} \tag{422}$$

with the following symmetry property satisfied by the coefficients:

$$H^{(3)}_{j_1, j_2, j_3} = H^{(3)}_{j_2, j_1, j_3} = H^{(3)}_{j_1, j_3, j_2} \tag{423}$$

Using this notation we can rewrite (419) as

$$h(\mathbf{x}+\mathbf{y}) = h(\mathbf{x}) + h'(\mathbf{x})(\mathbf{y}) + \frac{1}{2} h''(\mathbf{x})(\mathbf{y}^2) + \cdots + \frac{1}{\nu!} h^{(\nu)}(\mathbf{x})(\mathbf{y}^{\nu}) + O(|y|^{\nu+1}) \tag{424}$$

8.2 *Fourier transform and linear differential operators*

The Fourier transform $\hat{u}(\boldsymbol{\eta})$, $\boldsymbol{\eta} \in \mathbf{R}^d$, of a function $u(\mathbf{r})$, $\mathbf{r} \in \mathbf{R}^d$, and its inverse are defined by

$$\hat{u}(\boldsymbol{\eta}) = \int_{\mathbf{R}^d} u(\mathbf{r})\mathrm{e}^{-\mathrm{i}\mathbf{r}\cdot\boldsymbol{\eta}}\,\mathrm{d}\mathbf{r}$$

$$u(\mathbf{r}) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \hat{u}(\boldsymbol{\eta})\mathrm{e}^{\mathrm{i}\mathbf{r}\cdot\boldsymbol{\eta}}\,\mathrm{d}\boldsymbol{\eta},\ \ \mathrm{d}\boldsymbol{\eta} = \mathrm{d}\eta_1\cdots\mathrm{d}\eta_d \tag{425}$$

Evidently

$$[\hat{u}(-\boldsymbol{\eta})]^* = \widehat{[u(\boldsymbol{\eta})]^*} \tag{426}$$

Let $\gamma(\boldsymbol{\eta})$ be a polynomial of the variable $\boldsymbol{\eta}$ written in a form of a Taylor polynomial similar to (424)

$$\gamma(\boldsymbol{\eta}) = \sum_{l=0}^{\nu}\frac{1}{|l|!}\gamma^{(l)}\vdots(\boldsymbol{\eta}^l) \tag{427}$$

where $\gamma^{(l)}\vdots(\boldsymbol{\eta}^l)$ is an l-linear form similar to (422). Then the differential operator $\gamma[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]$ is defined by formally replacing variables η_j by the differential operators $-\mathrm{i}\partial_j = -\mathrm{i}\partial/\partial r_j$ in the polynomial $\gamma(\boldsymbol{\eta})$. The polynomial $\gamma(\boldsymbol{\eta})$ is called the *symbol* of the operator $\gamma[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]$. In particular, a general polynomial of third degree with $\nu = 3$ takes the form

$$\gamma_{(3)}(\boldsymbol{\eta}) = \gamma_0 + \sum_m \gamma_{1m}\eta_m + \frac{1}{2}\sum_{m,l}\gamma_{ml}\eta_m\eta_l + \frac{1}{6}\sum_{m,l,j}\gamma_{mlj}\eta_m\eta_l\eta_j \tag{428}$$

Consequently, the operator $\gamma_{(3)}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]$ with this symbol takes the form

$$\gamma_{(3)}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]V = \omega_{n_0}(\mathbf{k}_*)V - \mathrm{i}\sum_m \gamma_{1m}\partial_m V - \frac{1}{2}\sum_{m,l}\gamma_{ml}\partial_m\partial_l V + \frac{\mathrm{i}}{6}\sum_{m,l,j}\gamma_{mlj}\partial_m\partial_l\partial_j V \tag{429}$$

More exactly, the operator $\gamma[-\mathrm{i}\beta\nabla_{\mathbf{r}}]$ is defined based on the Fourier transform and the symbol $\gamma(\beta\boldsymbol{\eta})$ is

$$\widehat{\gamma[-\mathrm{i}\beta\nabla_{\mathbf{r}}]u}(\boldsymbol{\eta}) = \gamma(\beta\boldsymbol{\eta})\hat{u}(\boldsymbol{\eta}) \tag{430}$$

8.3 *Nonlinearity in ENLS*

Nonlinear terms in the ENLS involve spatial and time derivatives. We show that their Fourier transforms have the same form as the convolution integrals in (277) or (397).

8.3.1 Nonlinearity involving spatial derivatives. Here we briefly describe the Fourier transform of expressions that involve a product of spatial derivatives. From (425) we obtain, for $\bar{l} = (l_1, \ldots, l_d)$, that

$$\nabla_{\mathbf{r}}^{\bar{l}}V(\mathbf{r}) = \frac{\partial^{|\bar{l}|}V}{\partial r_1^{l_1}\ldots\partial r_d^{l_d}} = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\mathrm{i}^{|\bar{l}|}\boldsymbol{\eta}^{\bar{l}}\hat{V}(\boldsymbol{\eta})\mathrm{e}^{\mathrm{i}\mathbf{r}\cdot\eta}\,\mathrm{d}\boldsymbol{\eta} \tag{431}$$

Consequently,

$$\widehat{(-\mathrm{i})^{|\bar{l}|}\nabla^{\bar{l}}V}(\boldsymbol{\eta}) = \boldsymbol{\eta}^{\bar{l}}\hat{V}(\boldsymbol{\eta}) \tag{432}$$

Now we introduce *linear operators acting in a nonsymmetric way on the three factors of a product of three functions of d variables as in (111).* First, let us introduce a symbol of such an operator. If $p^{[\sigma]}(\vec{q}^{\,\star})$ is a polynomial of $\vec{q}^{\,\star}=(\mathbf{q}',\mathbf{q}'',\mathbf{q}''')$ of degree σ it can be written as a sum of monomials in the form

$$p^{[\sigma]}(\vec{q}^{\,\star})=\sum_{|\bar{l}'|+|\bar{l}''|+|\bar{l}'''|\leq\sigma}a_{\bar{l}',\bar{l}'',\bar{l}'''}(\mathbf{q}')^{\bar{l}'}(\mathbf{q}'')^{\bar{l}''}(\mathbf{q}''')^{\bar{l}'''} \tag{433}$$

where $a_{\bar{l}',\bar{l}'',\bar{l}'''}$ are the coefficients of the polynomial and multi-indices $\bar{l}',\bar{l}'',\bar{l}'''$ have the form $\bar{l}'=(l_1',\ldots,l_d')$ etc.

We define the action of the differential operator $p_m[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]$ on the product of three functions V',V'',V''' by the following formula

$$p^{[\sigma]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}](V'V''V''')=\sum_{|\bar{l}'|+|\bar{l}''|+|\bar{l}'''|\leq\sigma}a_{\bar{l}',\bar{l}'',\bar{l}'''}([-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]^{\bar{l}'}V')([-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]^{\bar{l}''}V'')([-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]^{\bar{l}'''}V''') \tag{434}$$

Notice that the order of factors in the product $V'V''V'''$ matters for the action of $p^{[\sigma]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}]$ and, generically,

$$p^{[\sigma]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}](V'V''V''')\neq p^{[\sigma]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}](V'V'''V'')\text{if } V''\neq V''' \tag{435}$$

For the Fourier transform we obtain the convolution formula

$$\widehat{p^{[\sigma]}[-\mathrm{i}\vec{\nabla}_{\mathbf{r}}](V'V''V''')}(\boldsymbol{\eta})=\frac{1}{(2\pi)^{2d}}\int_{\boldsymbol{\eta}'+\eta''+\boldsymbol{\eta}'''=\boldsymbol{\eta}}p^{[\sigma]}(\vec{\eta}^{\,\star})\hat{V}'(\boldsymbol{\eta}')\hat{V}''(\boldsymbol{\eta}'')\hat{V}'''(\boldsymbol{\eta}''')\,\mathrm{d}\eta'\,\mathrm{d}\boldsymbol{\eta}'' \tag{436}$$

Multiplying this expression by an oscillating exponent we obtain expressions completely similar to the integrands in (277) and (333).

8.3.2 Nonlinearity involving time derivatives. Equations (114) and (125) have time derivatives in the nonlinear terms, for example $Z_\zeta Z_{-\zeta}(\partial_t+\mathrm{i}\mathcal{L}_\zeta^{[4]})Z_\zeta$ in (114). We show here that the FNLR that corresponds to these terms has the same form as the FNLR coming from the first-order correction to the susceptibility which is given by (396). The FM terms of the form (396) lead to the FNLR of the following form similar to (397):

$$\begin{aligned}&\beta^d\tilde{u}_{\bar{n}}^{(1,\bar{l})}\left(\zeta\mathbf{k}_*+\beta\mathbf{q},\frac{\tau}{\varrho}\right)\\&\quad=\frac{1}{\varrho}\int_0^\tau\int_{\mathbf{q}'+\mathbf{q}''+\mathbf{q}'''=\mathbf{q}}\exp\left\{\mathrm{i}\Phi^{(\nu)}(\vec{\zeta}_0,\beta\vec{q})\frac{\tau_1}{\varrho}\right\}\partial_{\tau_1}\psi(\tau_1)\psi^2(\tau_1)\\&\quad\Psi^3(\beta\vec{q}^{\,0})\breve{Q}_{\bar{n},\bar{l}}(\vec{\zeta}_0\mathbf{k}_*+\vec{q}^{\,0})\hat{h}_\zeta(\mathbf{q}')\hat{h}_\zeta(\mathbf{q}'')\hat{h}_{-\zeta}(\mathbf{q}'''(\vec{q}))\mathrm{d}\mathbf{q}'\,\mathrm{d}\mathbf{q}''\,\mathrm{d}\tau_1\end{aligned} \tag{437}$$

Using (258) and (262) we obtain for the terms of the expansion (394)

$$\begin{aligned}&\varrho^{|\bar{l}|}\tilde{u}_{\bar{n}}^{(1,\bar{l})}\left(\zeta\mathbf{k}_*+\beta\mathbf{q},\frac{\tau}{\varrho}\right)\\&\quad=\frac{\varrho^{|l|}\beta^{2d}}{\varrho}\int_0^\tau\int_{\mathbb{R}^{2d}}\exp\left\{\mathrm{i}\Phi^{(\nu)}(\vec{\zeta}_0,\beta\vec{q})\frac{\tau_1}{\varrho}\right\}\partial_{\tau_1}\psi(\tau_1)\psi^2(\tau_1)\left[p_{\zeta,\bar{l}}^{[\sigma]}(\beta\vec{q})+O(\beta^{\sigma+1})\right]\\&\quad\hat{h}_\zeta(\mathbf{q}')\hat{h}_\zeta(\mathbf{q}'')\hat{h}_{-\zeta}(\mathbf{q}'''(\vec{q}))(1+O(\beta^\nu))\,\mathrm{d}\mathbf{q}'\,\mathrm{d}\mathbf{q}''\,\mathrm{d}\tau_1+O\left(\frac{\beta^{N_\Psi-2d}}{\varrho}\right)\end{aligned} \tag{438}$$

where $p_{\zeta,l}^{[\sigma]}(\beta\vec{q})$ is the Taylor approximation for $\breve{Q}_{\bar{n},\bar{l}}(\vec{k})$ calculated at $\vec{k}=\vec{\zeta}_0\mathbf{k}_*$ by a formula similar to (260). We take in this formula $\sigma=0$, $\nu=2$. Note that $p_{\zeta,\bar{l}}^{[0]}(\beta\vec{q})=\breve{Q}_{\bar{n},\bar{l}}(\vec{\zeta}_0\mathbf{k}_*)$ is the same for $\bar{l}=(1,0,0)$ and $\bar{l}=(0,1,0)$.

We introduce the additional terms in the NLS that approximate this integral. From (332) and (324) we obtain the identity

$$\varrho \exp\{-\mathrm{i}\zeta\gamma_{(\nu)}(\beta\zeta\mathbf{q}')t\}\partial_t\hat{v}_\zeta^{(0)}(\mathbf{q},\varrho t) = \partial_t\big[\hat{V}_\zeta^{(0)}(\mathbf{q}',t)\big] + \mathrm{i}\zeta\gamma_{(\nu)}(\beta\zeta\mathbf{q}')\hat{V}_\zeta^{(0)}(\mathbf{q}',t) \qquad (439)$$

Since $\varrho t = \tau$, $\partial_t = \varrho\partial_\tau$, (439) implies

$$\varrho \exp\left\{\mathrm{i}\Phi^{(\nu)}(\vec{\zeta}_0,\beta\vec{q})\frac{\tau_1}{\varrho}\right\}\partial_{\tau_1}\psi(\tau_1)\psi^2(\tau_1)\hat{h}_\zeta(\mathbf{q}')\hat{h}_\zeta(\mathbf{q}'')\hat{h}_{-\zeta}(\mathbf{q}'''(0))$$
$$= \mathrm{e}^{\mathrm{i}\zeta\gamma_{(\nu)}(\beta\zeta\mathbf{q})t}\big[\partial_t[\hat{V}^{(0)}(\mathbf{q}',t)] + \mathrm{i}\zeta\gamma_{(\nu)}(\beta\zeta\mathbf{q}')\hat{V}_\zeta^{(0)}(\mathbf{q}',t)\big]\hat{V}_\zeta^{(0)}(\zeta\mathbf{q}'',t)\hat{V}_{-\zeta}^{(0)}(\mathbf{q}'''(0),t)$$

where $\hat{V}_\zeta^{(0)}(\mathbf{q},t)$ is given by (324). Hence, for $\bar{l} = (1,0,0)$ the principal part of the integral (438) coincides with the slow time factor as in (332) of the Fourier transform of the term

$$\frac{(2\pi)^{2d}}{\varrho}\int_0^\tau \check{Q}_{\vec{n},\bar{l}}(\vec{\zeta}_0\mathbf{k}_*)\big(\hat{V}_\zeta^{(0)}\hat{V}_{-\zeta}^{(0)}\big(\partial_t V_\zeta^{(0)} + \mathrm{i}\zeta\gamma_{(\nu)}(-\mathrm{i}\zeta\beta\nabla_{\mathbf{r}})V_\zeta^{(0)}\big)\big)\,\mathrm{d}\tau_1 \qquad (440)$$

where $V^{(0)}$ is the linear response of the NLS given by (324). A similar formula holds for $\bar{l} = (0,1,0)$. For $\bar{l} = (0,0,1)$ the principal part of the integral (438) coincides with the slow time factor of the Fourier transform of the term

$$\frac{(2\pi)^{2d}}{\varrho}\int_0^\tau \check{Q}_{\vec{n},\bar{l}}(\vec{\zeta}_0\mathbf{k}_*)\left[V^{(0)2}\left(\partial_t V_{\zeta'''}^{(0)} + \frac{\mathrm{i}\zeta'''}{\varrho}\gamma_{(\nu)}(-\mathrm{i}\beta\zeta'''\nabla_{\mathbf{r}})V_{\zeta'''}^{(0)}\right)\right]\mathrm{d}\tau_1,\ \zeta''' = -\zeta \qquad (441)$$

Hence the part of the FNLR of the ENLS corresponding to the terms in (114) with

$$\delta_{1,\zeta} = 2\check{Q}_{\vec{n},1,0,0}(\vec{\zeta}_0\mathbf{k}_*), \quad \delta_{2,\zeta} = \check{Q}_{\vec{n},0,0,1}(\vec{\zeta}_0\mathbf{k}_*), \quad \zeta = \pm \qquad (442)$$

coincides with the principal part of (438), here we use notation (393) where $\bar{l} = (1,0,0)$ or $\bar{l} = (0,0,1)$.

9. Lattice nonlinear Schrödinger equation

We now show how the NLM can be approximated by a lattice NLS with the same precision as by the classical NLS in the entire space. We consider for simplicity the case of real-valued excitations and lower order approximations.

In the one-dimensional case the lattice NLSs replacing the NLSs (36) and (37) have the form similar to (50), i.e.

$$\partial_t Z_\zeta(m) = -\mathrm{i}\zeta(\gamma_0+\gamma_2)Z_\zeta(m) - \gamma_1\left(\frac{1}{2}[Z_\zeta(m+1) - Z_\zeta(m-1)]\right)$$
$$+\mathrm{i}\zeta\frac{\gamma_2}{2}[Z_\zeta(m+1) + Z_\zeta(m-1)] + \alpha_\pi Q_\zeta Z_{-\zeta}(m)Z_\zeta^2(m)$$
$$Z_\zeta(m)|_{t=0} = h_\zeta(\beta m), \quad \alpha_\pi = 3\alpha(2\pi)^2, \quad m = \cdots -1,0,1,2,\ldots, \quad \zeta = \pm \quad (443)$$

The equations do not involve the spatial derivatives and have the form of a sequence of ordinary differential equations describing coupled nonlinear oscillators.

The approximation of the NLM by the NLS is based on: (i) the approximation of $\omega_{n_0}(\mathbf{k}_*+\boldsymbol{\eta})$ by its Taylor polynomial $\gamma_{(\nu)}(\boldsymbol{\eta})$ in (200); and (ii) the approximation of the modal susceptibility $\check{Q}_{\vec{n}_0}(\vec{\zeta}_0\mathbf{k}_* + \vec{\eta})$ defined by (179) by its Taylor polynomial $p_{\mathrm{T},\zeta}^{[\sigma]}(\vec{\eta})$ (see (260)) in a vicinity of $\zeta\mathbf{k}_*$. Here we use trigonometric polynomials instead of algebraic and as a result we obtain (443). Here we consider the case $\nu = 2$, $\sigma = 0$, with the understanding that larger values of

ν and σ can be considered similarly. Using an orthogonal change of variables $\boldsymbol{\eta} = \boldsymbol{\Theta\xi}$ we reduce the quadratic form to the diagonal form $\omega''_{n_0}(\mathbf{k}_*)$ and obtain

$$\gamma_{(2)}(\boldsymbol{\Theta\xi}) = \gamma_0 + \sum_m \Gamma_{1,m}\xi_m + \frac{1}{2}\sum_m \Gamma_{2,m}\xi_m^2 \tag{444}$$

where

$$\omega_{n_0}(\mathbf{k}_* + \boldsymbol{\Theta\xi}) = \gamma_{(2)}(\boldsymbol{\Theta\xi}) + O(|\xi|^3) \tag{445}$$

Now we use trigonometric polynomials instead of algebraic ones. Obviously

$$\xi_m = \sin\xi_m + O(|\xi|^3) \tag{446}$$

$$\xi_m^2 = 2 - 2\cos\xi_m + O(|\xi|^4) \tag{447}$$

We set

$$\Gamma_{(2)}(\boldsymbol{\xi}) = \Gamma_0 + \sum_m \Gamma_{1,m}\sin\xi_m - \sum_m \Gamma_{2,m}\cos\xi_m \tag{448}$$

$$\Gamma_0 = \gamma_0 + \sum_m \Gamma_{2,m} \tag{449}$$

which together with (444) yield

$$\omega_{n_0}(\mathbf{k}_* + \boldsymbol{\Theta\xi}) = \Gamma_{(2)}(\boldsymbol{\xi}) + O(|\xi|^3) \tag{450}$$

In particular, for $d = 1$

$$\begin{aligned}\Gamma_{(2)}(\boldsymbol{\eta}) &= [\omega_{n_0}(k_*) + \omega''_{n_0}(k_*)] + \omega'_{n_0}(k_*)\sin\boldsymbol{\eta} - \omega''_{n_0}(k_*)\cos\boldsymbol{\eta}\\ &\times \omega_{n_0}(k_*) + \omega'_{n_0}(k_*)\boldsymbol{\eta} + \frac{1}{2}\omega''_{n_0}(k_*)\boldsymbol{\eta}^2\\ &= [\omega_{n_0}(k_*) + \omega''_{n_0}(k_*)] + \omega'_{n_0}(k_*)\sin\boldsymbol{\eta} - \omega''_{n_0}(k_*)\cos\boldsymbol{\eta} + O(\boldsymbol{\eta}^3)\end{aligned} \tag{451}$$

An advantage of this representation compared with its algebraic counterpart is that it involves a periodic function similar to $\omega_n(\mathbf{k})$, namely

$$\Gamma_{(2)}(\boldsymbol{\xi} + 2\pi\boldsymbol{\zeta}) = \Gamma_{(2)}(\boldsymbol{\xi}) \tag{452}$$

Remark In the case $\nu \geq 3$ we can approximate functions $\omega_n(\mathbf{k}_* + \boldsymbol{\eta})$ by $\sin^l(\boldsymbol{\eta}_j)$, $l = 1, \ldots, \nu$. Based on the Taylor polynomial

$$\gamma_{(\nu)}(\boldsymbol{\eta}) = \sum_{j=0}^{\nu} \frac{1}{j!}\omega_{n_0}^{(j)}(\mathbf{k}_*)(\boldsymbol{\eta}^j), \quad \boldsymbol{\eta} = \mathbf{k} - \mathbf{k}_* \tag{453}$$

we form a trigonometric polynomial

$$\Gamma_{(\nu)}(\boldsymbol{\eta}) = \sum_{j=0}^{\nu} \frac{1}{j!}\Gamma_{(\nu)}^{(j)}(\sin\boldsymbol{\eta})^j \tag{454}$$

where

$$\boldsymbol{\eta} = (\eta_1, \ldots, \eta_d), \quad \sin\eta = (\sin\eta_1, \ldots, \sin\eta_d) \tag{455}$$

The coefficients $\Gamma_{(\nu)}^{(j)}$ are uniquely determined by $\omega_{n_0}^{(i)}(\mathbf{k}_*)$, $i = 0, \ldots, j$ since the change of variables $\sin\eta_l \leftrightarrow \eta_l$ is invertible about the origin. ♦

9.1 *Functions on a lattice and the discrete Fourier transform*

We consider the lattice of vectors with integer components

$$\mathbf{m} = (m_1, m_2, \ldots, m_d) \in \mathbb{Z}^d \tag{456}$$

and functions $Z(\mathbf{m})$ on the lattice $\mathbb{Z}^d$. The shift operators are defined as

$$\dot{\partial}_{+,j} Z = Z(\ldots, m_j + 1, \ldots), \quad \dot{\partial}_{-,j} Z = Z(\ldots, m_j - 1, \ldots) \tag{457}$$

The elementary difference operators are then defined as

$$\begin{aligned}\dot{\Delta}_{+,j} Z &= \frac{1}{2}[Z(\ldots, m_j + 1, \ldots) + Z(\ldots, m_j - 1, \ldots)] = \frac{1}{2}[\dot{\partial}_{+,j} Z + \dot{\partial}_{-,j} Z] \\ \dot{\Delta}_{-,j} Z &= \frac{1}{2\mathrm{i}}[Z(\ldots, m_j + 1, \ldots) - Z(\ldots, m_j - 1, \ldots)] = \frac{1}{2\mathrm{i}}[\dot{\partial}_{+,j} Z - \dot{\partial}_{-,j} Z]\end{aligned} \tag{458}$$

For every lattice function $Z(\mathbf{m})$ which decays at infinity fast enough we define its Fourier transform

$$\bar{Z}(\boldsymbol{\xi}) = \sum_{\mathbf{m}} Z(\mathbf{m}) \mathrm{e}^{-\mathrm{i}\mathbf{m}\cdot\boldsymbol{\xi}} \tag{459}$$

with the inverse transform

$$Z(\mathbf{m}) = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \mathrm{e}^{\mathrm{i}\mathbf{m}\cdot\boldsymbol{\xi}} \bar{Z}(\boldsymbol{\xi})\, \mathrm{d}\boldsymbol{\xi} \tag{460}$$

Obviously $\bar{Z}(\boldsymbol{\xi})$ is a 2π -periodic function of $\boldsymbol{\xi} \in \mathbb{R}^d$. The Fourier transform of the difference operators $\dot{\Delta}_{+,i} Z$ is given by

$$\overline{[\dot{\Delta}_{+,j} Z]}(\boldsymbol{\xi}) = \cos \boldsymbol{\xi}_j \bar{Z}(\mathbf{k}), \ \overline{[\dot{\Delta}_{-,j} Z]}(\boldsymbol{\xi}) = \sin \boldsymbol{\xi}_j \bar{Z}(\boldsymbol{\xi}) \tag{461}$$

When $d = 1$ we omit j and set

$$\dot{\Delta}_+ Z = \frac{1}{2}[Z(m+1) + Z(m-1)], \ \dot{\Delta}_- Z = \frac{1}{2\mathrm{i}}[Z(m+1) - Z(m-1)] \tag{462}$$

Note that the Fourier transform of the product is given by the following convolution formula

$$\overline{XZ}(\boldsymbol{\xi}) = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \overline{X}(\mathbf{q}) \overline{Z}(\boldsymbol{\xi} - \mathrm{q})\, \mathrm{d}\mathbf{q} \tag{463}$$

as in the case of the continuous Fourier transform.

9.2 *Lattice NLS (LNLS)*

When $\Gamma_{(2)}(\boldsymbol{\xi})$ is given by (448), we define the difference operator on the lattice by the formula

$$\Gamma_{(2)}(\zeta \dot{\nabla}) Z = \Gamma_0 Z + \zeta \sum_m \Gamma_{1,m} \dot{\Delta}_{-,m} Z - \sum_m \Gamma_{2,m} \dot{\Delta}_{+,m} Z \tag{464}$$

Note that its Fourier transform is

$$\overline{\Gamma_{(2)}(\zeta \dot{\nabla}) Z}(\boldsymbol{\xi}) = \Gamma_{(2)}(\zeta \boldsymbol{\xi}) \overline{Z}(\boldsymbol{\xi}) \tag{465}$$

Let us introduce a *linear lattice Schrödinger equation (LLS)*

$$\partial_t Z(\mathbf{m}, t) = -\mathrm{i}\Gamma_{(2)}(\dot{\nabla}) Z(\mathbf{m}, t), \ Z(\mathbf{m}, t)|_{t=0} = h(\mathbf{m}), \quad \mathbf{m} \in \mathbb{Z}^d \tag{466}$$

It can be solved exactly in terms of its lattice Fourier transform (459), namely

$$\bar{Z}(\boldsymbol{\xi}, t) = \bar{h}(\boldsymbol{\xi}) \exp\{-\mathrm{i}\Gamma_{(2)}(\boldsymbol{\xi}) t\} \tag{467}$$

Let us now introduce a *lattice nonlinear Schrödinger equation (LNLS)*

$$\partial_t Z_\zeta = -\mathrm{i}\zeta \Gamma_{(2)}(\zeta \dot{\nabla}) Z_\zeta + \alpha_\pi Q_\zeta Z_{-\zeta} Z_\zeta^2, \quad Z_\zeta(\mathbf{m}, t)|_{t=0} = h_{\beta,\zeta}(\mathbf{m}) \tag{468}$$

where Q_ζ is a complex constant and the factor $\alpha_\pi = 3\alpha(2\pi)^{2d}$ is introduced for notational consistency with the related NLM. Here

$$h_{\beta,\zeta}(\mathbf{m}) = h_\zeta(\beta \mathbf{m}), \quad \zeta = \pm \tag{469}$$

where $h_\zeta(\mathbf{r})$, $\mathbf{r} \in \mathbb{R}^d$ is a given smooth function of continuous argument. Its lattice Fourier transform is given by

$$\bar{h}_{\beta,\zeta}(\boldsymbol{\xi}) = \sum_{\mathbf{m}} h_\zeta(\beta \mathbf{m}) \mathrm{e}^{-\mathrm{i}\mathbf{m}\cdot\boldsymbol{\xi}}, \quad \zeta = \pm \tag{470}$$

with the inverse formula

$$h_{\beta,\zeta}(\mathbf{m}) = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \mathrm{e}^{\mathrm{i}\mathbf{m}\cdot\boldsymbol{\xi}} \bar{h}_{\beta,\zeta}(\boldsymbol{\xi})\, \mathrm{d}\boldsymbol{\xi} \tag{471}$$

Note that this formula makes sense even for non-integer values of $\mathbf{m} = \mathbf{r}$ providing an interpolation to such values. We can replace in (58) $\hat{Z}_\zeta$ based on a solution of the NLS by by $\overline{Z}_\zeta$ based on the LNLS. Similarly to (59) we obtain that the modal coefficient $\tilde{U}_{+,n_0} = \tilde{U}_{+,n_0}$ of the solution of the NLM is well approximated in terms of the solution of the LNLS, namely

$$\tilde{U}_{\zeta,n_0}(\mathbf{k}_* + \Theta\boldsymbol{\xi}, t) = \bar{Z}_\zeta(\boldsymbol{\xi}, t) + O(\beta) + O(\varrho) \tag{472}$$

Note that in the one-dimensional case $\boldsymbol{\Theta}\boldsymbol{\xi} = \boldsymbol{\xi}$. One can see that LNLS gives the same order of accuracy as the NLS.

9.3 *Presentation in spatial domain*

Formula (356) takes the form

$$\mathbf{U}_Z(\mathbf{r}, t) = \frac{\beta^d}{(2\pi)^d} \int_{[-\pi/\beta,\pi/\beta]^d} \Psi(\beta\mathbf{q}) \\ [\bar{Z}_+(\beta\mathbf{q}, t)\tilde{\mathbf{G}}_{+,n_0}(\mathbf{r}, \mathbf{k}_* + \beta\mathbf{q}) + \bar{Z}_-(\beta\mathbf{q}, t)\tilde{\mathbf{G}}_{-,n_0}(\mathbf{r}, -\mathbf{k}_* + \beta\mathbf{q})]\mathrm{d}\mathbf{q} \tag{473}$$

According to (142)

$$\tilde{\mathbf{G}}_{\zeta,n_0}(\mathbf{r}, \mathbf{k}) = \hat{\mathbf{G}}_{\zeta,n_0}(\mathbf{r}, \mathbf{k}) \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}}, \quad \zeta = \pm \tag{474}$$

where $\hat{\mathbf{G}}_{\zeta,n_0}(\mathbf{r}, \mathbf{k})$ is a 1-periodic function of $\mathbf{r}$. Instead of (358) we have, similarly to (448), expansion into trigonometric functions

$$\hat{\mathbf{G}}_{\zeta,n_0}(\mathbf{r}, \mathbf{k}_* + \beta\mathbf{q}) = \dot{\mathbf{p}}_{\zeta,g,\sigma}(\mathbf{r}, \beta\mathbf{q}) + O(\beta^{\sigma+1}), \ \sigma + 1 \le \nu \tag{475}$$

where, for $\sigma = 2$,

$$\dot{\mathbf{p}}_{\zeta,g,\sigma}(\mathbf{r}, \beta\mathbf{q}) = [\hat{\mathbf{G}}_{\zeta,n_0}(\mathbf{r}, \mathbf{k}_*) + \hat{\mathbf{G}}''_{\zeta,n_0}(\mathbf{r}, \mathbf{k}_*)] + \hat{\mathbf{G}}'_{\zeta,n_0}(\mathbf{r}) \sin(\beta\mathbf{q}) - \hat{\mathbf{G}}''_{\zeta,n_0}(\mathbf{r}) \cos(\beta\mathbf{q}) \tag{476}$$

We obtain (355) where, similarly to (366)

$$\mathbf{U}_{Z_\zeta}(\mathbf{m}, t) = \mathbf{U}^0_{Z_\zeta}(\mathbf{m}, t) + \mathbf{U}^1_{Z_\zeta}(\mathbf{m}, t) + \mathbf{U}^2_{Z_\zeta}(\mathbf{m}, t) + O(\beta^3) \tag{477}$$

where $\mathbf{U}^0_{Z_+}$ is given by

$$\mathbf{U}^0_{Z_\zeta}(\mathbf{m}, t) = \mathrm{e}^{\mathrm{i}\mathbf{k}_*\cdot\mathbf{m}}[\hat{\mathbf{G}}_{\zeta,n_0}(\mathbf{m}, \mathbf{k}_*) + \hat{\mathbf{G}}''_{\zeta,n_0}(\mathbf{m}, \mathbf{k}_*)] Z_\zeta(\mathbf{m}, t) \tag{478}$$

and, similarly to (367),

$$\mathbf{U}^1_{Z_\zeta}(\mathbf{m},t) = \mathrm{e}^{\mathrm{i}\mathbf{k}_*\cdot\mathbf{m}}\hat{\mathbf{G}}'_{\zeta,n_0}(\mathbf{m},\mathbf{k}_*)\cdot\dot{\Delta}_- Z_\zeta(\mathbf{m},t) \tag{479}$$

$$\mathbf{U}^2_{Z_\zeta}(\mathbf{m},t) = -\mathrm{e}^{\mathrm{i}\mathbf{k}_*\cdot\mathbf{m}}\hat{\mathbf{G}}''_{\zeta,n_0}(\mathbf{m},\mathbf{k}_*)\cdot\dot{\Delta}_+ Z_\zeta(\mathbf{m},t) \tag{480}$$

Note that using (471) we can interpolate $\mathbf{U}_{Z_\pm}(\mathbf{m},t)$ to non-integer $\mathbf{m}=\mathbf{r}$.

9.4 *Comparison of the lattice NLS with the NLS*

The LNLS approximation, compared with the classical NLS has the following properties.

- The accuracy of approximation by the lattice NLS (468) is the same as by the NLS (100).
- The right-hand side of (468) is a bounded operator, which is an advantage over the NLS.
- The lattice system (468) is already in a spatially discretized form, which can be advantageous for numerical simulations.
- The form of LNLS (468) suggests that small-scale (compared with the cell size) features of the wave dynamics are effectively eliminated. Note that the derivation of the NLS also assumes the elimination of the small-scale, but the differential form of the NLS still allows small-scale perturbations to be of importance for large-scale wave dynamics.

There is extensive literature on coupled nonlinear oscillators on lattices, see for example [46] and [47] and references therein. For photonic crystals such equations were used in [48].

Remark In the case when we use (454) the difference differential equation (468) takes the form

$$\partial_t Z = -\mathrm{i}\Gamma_{(2)}(\dot{\Delta}_-)Z + \alpha_\pi Q_+|Z|^2 Z,\ \ Z(\mathbf{m},t)|_{t=0} = h_\beta(\mathbf{m}) \tag{481}$$

where the operator $\Gamma_{(2)}(\dot{\Delta}_-)$ is obtained by substituting $\dot{\Delta}_{-,j}$ in place of $\sin\eta_j$ in the polynomial (454). ♦

Remark Note that usually the terms with operators $\dot{\Delta}_{-,m}$ are not involved in the lattice Schrödinger equations considered in the literature. The reason is that the influence of these terms on solutions with the initial data $h(\beta\mathbf{m})$ for $\beta \ll 1$ can be taken into account by choosing a coordinate frame moving with the group velocity. To give a simple explanation, we use another approximation for $\omega_{n_0}(\zeta\mathbf{k}_* + \boldsymbol{\Theta}\boldsymbol{\xi})$, namely

$$\omega_{n_0}(\zeta\mathbf{k}_* + \boldsymbol{\Theta}\boldsymbol{\xi}) = \Gamma_{(2)}(\boldsymbol{\xi}) = \Gamma_0 + \sum_m \Gamma_{1,m}\xi_m - \sum_m \Gamma_{2,m}\cos\xi_m + O(|\xi|^3) \tag{482}$$

which combines linear functions with trigonometric. The corresponding difference differential equation in $\mathbb{R}^d$ has the form

$$\partial_t Z(\mathbf{x},t) = -\mathrm{i}\Gamma_0 Z(\mathbf{x},t) + \sum_m \Gamma_{1,m}\frac{\partial}{\partial x_m}Z(\mathbf{x},t) + \mathrm{i}\sum_m \Gamma_{2,m}\dot{\Delta}_{+,m}Z(\mathbf{x},t)$$
$$+\alpha_\pi Q_+|Z|^2 Z(\mathbf{x},t), \quad \mathbf{x}\in\mathbb{R}^d \tag{483}$$

involving both differential and finite difference operators. The standard change of variables

$$Z(\mathbf{x},t) = z((\mathbf{x}+\Gamma t, t)),\ \ \Gamma = (\Gamma_{1,1},\ldots,\Gamma_{1,d}) \tag{484}$$

reduces this equation to the following NLS difference equation

$$\partial_t z(\mathbf{x},t) = -\mathrm{i}\Gamma_0 z(\mathbf{x},t) + \mathrm{i}\sum_m \Gamma_{2,m}\dot{\Delta}_{+,m}z(\mathbf{x},t) + \alpha_\pi Q_+|z|^2 z(\mathbf{x},t) \tag{485}$$

Obviously this equation is equivalent to a family of independent equations on the lattice $\mathbb{Z}^d$ of the same form as (468) but without the terms $\dot{\Delta}_{-,m}$. ♦

10. Conclusions

The basic conditions on a periodic dielectric medium to support nonlinear Schrödinger regimes of electromagnetic wave propagation are: (i) the inversion symmetry $\omega_n(-\mathbf{k}) = \omega_n(\mathbf{k})$ of the dispersion relations; and (ii) the leading term in the nonlinearity is cubic. Nonlinear Schrödinger regimes are generated by almost time-harmonic excitation currents with localized quasimomenta and their most essential properties are as follows.

- The asymptotic nature of nonlinear Schrödinger regimes is determined by three small parameters α, ϱ, β. The parameter α scales the magnitude of the nonlinearity; it is proportional to the square of the amplitude of the excitation. The parameter $1/\varrho$ is proportional to time extension of the initial current excitation. The parameter β describes the range of quasimomenta $\mathbf{k}$ about a fixed $\mathbf{k}_*$ in the modal composition of the excitation current. The NLS regimes arise when $\alpha \sim \varrho$ and
$$\alpha \sim \varrho \quad \text{and} \quad \varrho \sim \beta^{\varkappa_1}, \quad \text{for some } \varkappa_1 > 0$$
In particular, the classical nonlinear Schrödinger regime is characterized by the following relations between the three small parameters
$$\alpha \sim \varrho \sim \beta^2$$
- The NLS and their extended versions describe approximately the evolution of the Floquet–Bloch modal coefficients $\tilde{U}_{\pm,n_0}(\mathbf{k}_* + \boldsymbol{\eta}, t)$ of the propagating wave.
- Multimodal excitation currents about several $\mathbf{k}_{*,j}$ generate nonlinear Schrödinger regimes that satisfy the principle of approximate superposition, that is with a very high accuracy $O(\beta^\infty)$ the modal components about different $\mathbf{k}_{*,j}$ evolve essentially independently according to NLSs or ENLSs.
- Higher accuracy approximations for longer time intervals are achieved by the analysis of the modal decomposition of the wave.

The accuracy of the NLS/ENLS approximation by developed methods can be characterized as follows.

- The classical NLS gives an approximation with error $O(\beta)$ over the time interval $O(1/\beta^2)$.
- To improve the accuracy to $O(\beta^2)$ over the time interval $O(1/\beta^2)$ it is sufficient to take into account the frequency dependence of the susceptibility tensor (in terms of first-order derivatives with respect to $\mathbf{k}$ of the tensor at $\mathbf{k} = \mathbf{k}_*$) and the third-order derivatives of the dispersion relation $\omega_{n_0}(\mathbf{k})$ at $\mathbf{k} = \mathbf{k}_*$, leading to the third-order ENLS (see (109)).
- To improve the accuracy further to $O(\beta^3)$ over the time interval $O(1/\beta^2)$ the following characteristics of the media have to be taken into account: (i) fourth-order derivatives of the dispersion relation $\omega_{n_0}(\mathbf{k})$ at $\mathbf{k} = \mathbf{k}_*$; (ii) second-order derivatives of the susceptibility tensor at $\mathbf{k} = \mathbf{k}_*$; (iii) nonlinear interactions between the forward- and backward-propagating waves; (iv) finer effects of the susceptibility approximation expressed in terms of the first-order frequency derivatives of the susceptibility; (v) fifth-order terms in the nonlinearity; (vi) nonlinear interactions between different spectral bands. The above effects are taken into account in the fourth-order ENLS (see section 1.4.3).
- The lattice NLS provides the same accuracy of approximation as the classical NLS with evident advantages for numerically efficient analysis.

Acknowledgment and disclaimer

Effort of A. Babin and A. Figotin is sponsored by the Air Force Office of Scientific Research, Air Force Materials Command, USAF, under grant number FA9550-04-1-0359. The US Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Office of Scientific Research or the US Government.

References

[1] Ablowitz, M. J. and Clarkson, P. A., 1991, *Solitons, Nonlinear Evolution Equations and Inverse Scattering*, London Mathematical Society Lecture Note Series, 149 (Cambridge: Cambridge University Press).

[2] Cai, D., McLaughlin, D. W., and McLaughlin, K. T. R., 2002, The nonlinear Schrödinger equation as both a PDE and a dynamical system. *Handbook of Dynamical Systems*, Vol. 2 (Amsterdam: North-Holland).

[3] Dodd, R. K., Eilbeck, J. C., Gibbon, J. D., and Morris, H. C., 1984, *Solitons and Nonlinear Wave Equations* (New York: Academic).

[4] Drazin, P. G., 1991, *Solitons: An Introduction* (Cambridge: Cambridge University Press).

[5] Lamb, G. L., 1980, *Elements of Soliton Theory* (New York: Wiley).

[6] *The nonlinear Schrödinger equation. Proceedings of the conference held in Chernogolovka*, July 25–August 3, 1994. Edited by A. V. Mikhaĭlov, E. A. Kuznetsov, A. C. Newell and V. E. Zakharov. Phys. D 87 (1995), no. 1–4. North-Holland Publishing Co., Amsterdam, 1995. pp. i–xiv and 1–379.

[7] Sulem, C. and Sulem, P.-L., 1999, *The Nonlinear Schrödinger Equation* (Berlin: Springer).

[8] Trullinger, S. E., Zakharov, V. E., and Pokrovsky, V. L., (Eds), 1986, *Solitons* (Amsterdam: North-Holland).

[9] Zakharov, V. E., L'vov, V. S., and Falkovich, C., 1992, *Kolmogorov Spectra of Turbulence I. Wave Turbulence* (New York: Springer).

[10] Newell, A. C. and Moloney, J. V., 1992, *Nonlinear Optics* (Reading, MA: Addison-Wesley).

[11] Whitham, G., 1974, *Linear and Nonlinear Waves* (New York: John Wiley).

[12] Bhat, N. R. E. and Sipe, J. E., 2001, Optical pulse propagation in nonlinear photonic crystals. *Physical Review E*, **64**, 056604.

[13] Bambusi, D., Carati, A., and Ponno, A., 2002, The nonlinear Schrödinger equation as a resonant normal form. *Discrete and Continuous Dynamical Systems*, **2**, 109–128.

[14] Colin, T., 2002, *Rigorous derivation of the nonlinear Schrödinger equation and Davey–Stewartson systems from quadratic hyperbolic systems*, *Asymptot. Anal.*, **31**, 69–91.

[15] Schneider, G., 1998, *Approximation of the Korteweg–de Vries equation by the nonlinear Schrödinger equation*. *Journal of Differential Equations*, **147**, 333–354.

[16] Schneider, G., 1998, *Justification of modulation equations for hyperbolic systems via normal forms*. *Nonlinear Differential Equations*, **5**, 69–82.

[17] Babin, A. and Figotin, A., 2001, *Nonlinear photonic crystals: I. Quadratic nonlinearity*. *Waves in Random Media*, **11**, R31–R102.

[18] Babin, A. and Figotin, A., 2002, *Nonlinear photonic crystals: II. Interaction classification for quadratic nonlinearities*. *Waves in Random Media*, **12**, R25–R52.

[19] Babin, A. and Figotin, A., 2003, *Nonlinear photonic crystals: III. Cubic nonlinearity*. *Waves in Random Media*, **13**, R41–R69.

[20] Babin, A. and Figotin, A., 2003, *Nonlinear Maxwell equations in inhomogenious media*. *Communication in Mathematics Physics*, **241**, 519–581.

[21] Butcher, P. and Cotter, D., 1990, *The Elements of Nonlinear Optics* (Cambridge: Cambridge University Press).

[22] Boyd, R., 1992, *Nonlinear Optics* (New York: Academic).

[23] Agrawal, G., 1995, *Nonlinear Fiber Optics* (New York: Academic).

[24] Dudley, J. M., Provino, L., Grossard, N., Maillotte, H., Windeler, R. S., Eggleton, B. J., and Coen, S., 2002, *Supercontinuum generation in air–silica microstructured fibers with nanosecond and femtosecond pulse pumping*. *Journal of the Optical Society of America B*, **19**, 765–771.

[25] Hong, W. P., 2002, *Modulational instability of optical waves in the high dispersive cubic–quintic nonlinear Schrödinger equation*. *Optics Communications*, **213**, 173–182.

[26] Hiroshi Kumagai, Sung-Hak Cho, Kenichi Ishikawa, Katsumi Midorikawa, Masatoshi Fujimoto, Shin-ichiro Aoshima, and Yutaka Tsuchiya, 2003, *Observation of the complex propagation of a femtosecond laser pulse in a dispersive transparent bulk material*. *J. Opt. Soc. Am. B*, **20**, 597–602.

[27] Hile, C. V., 1996, *Comparisons between Maxwell's equations and an extended nonlinear Schrödinger equation*. *Wave Motion*, **24**, 1–12.

[28] Karpman, V. I., 2000, *Radiation of solitons described by a high-order cubic nonlinear Schrödinger equation*. *Physical Review E*, **62**, 5678–5687.

[29] Nakkeeran, K., 2000, *On the integrability of the extended nonlinear Schrödinger equation and the coupled extended nonlinear Schrödinger equations*. *Journal of Physics A: Mathematical and General*, **33**, 3947–3949.
[30] Berge, L. and Couairon, A., 2000, Nonlinear propagation of self-guided ultra-short pulses in ionized gases. *Physics of Plasmas*, **7**, 210–230.
[31] Frantzeskakis, D. J., 1996, *Small-amplitude solitary structures for an extended nonlinear Schrödinger equation*. *Journal of Physics A: Mathematical and General*, **29**, 3631–3639.
[32] Cavalcanti, S. B., Cressoni, J. C., da Cruz, H. R., and Gouveia–Neto, A. S., 1991, *Modulation instability in the region of minimum group-velocity dispersion of single-mode optical fibers via an extended nonlinear Schrödinger equation*. *Physical Review A*, **43**, 6162–6165.
[33] Hayashi, N. and Naumkin, P., 2002, *On the reduction of the modified Benjamin–Ono equation to the cubic derivative nonlinear Schrödinger equation*. *Discrete and Continuous Dynamical Systems—Series A*, **8**, 237–255.
[34] Potasek, M. J., 1987, *Modulation instability in an extended nonlinear Schrödinger equation*. *Optics Letters*, **12**, 921–923.
[35] Kong, J. A., 1990, *Electromagnetic Wave Theory* (New York: Wiley).
[36] Goodman, R. H., Weinstein, M. I., and Holmes, P. J., 2001, Nonlinear propagation of light in one-dimensional periodic structures. *Journal of Nonlinear Science*, **11**, 123–168.
[37] Mills, D., 1991, *Nonlinear Optics* (New York: Springer).
[38] Stein, E., 1993, *Harmonic Analysis. Real-Variable Methods, Orthogonality, and Oscillatory Integrals* (Princeton, NJ: Princeton University Press).
[39] Babin, A. and Figotin, A., 2002, *Multilinear spectral decomposition for nonlinear Maxwell equations*. In: M. S. Agranovich and M. A. Shubin, *Partial Differential Equations, Advances in Mathematical Sciences, American Mathematical Society Translations-Series 2*, Vol. 206, pp. 1–28.
[40] Kenig, C. E., Ponce, G., and Vega, L., 1993, Small solutions to nonlinear Schrödinger equations. *Annales de l'Institut Henri Poincaré (c) Anal. Non Linéaire*, **10**, 255–288.
[41] Kenig, C. E., Ponce, G., and Vega, L., 1998, Smoothing effects and local existence theory for the generalized nonlinear Schrödinger equations. *Inventiones Mathematicae*, **134**, 489–545.
[42] Sandstede, B., 2002, Stability of travelling waves. *Handbook of Dynamical Systems*, Vol. 2 (Amsterdam: North-Holland).
[43] Weinstein, M. I., 1985, Modulational stability of ground states of nonlinear Schrödinger equations. *SIAM Journal of Mathematical Analysis*, **16**, 472–491.
[44] Weinstein, M. I., 1986, Lyapunov stability of ground states of nonlinear dispersive evolution equations. *Communications on Pure and Applied Mathematics* **39**, 51–67.
[45] Schneider, G. and Uecker, H., 2003, *Existence and stability of modulating pulse solutions in Maxwell's equations describing nonlinear optics*, *Zeitschrift für angewandte Mathematik und Physik*, **54**, 677–712.
[46] Localization in nonlinear lattices. *Physica D*, **119**, 1–238.
[47] Flach, S. and MacKay, R. S., 2000, Discrete breathers: classical and quantum. *Physica A*, **288**, 174–198.
[48] Eisenberg, H., Silberberg Y., *Discrete Solitons*, in [59] p. 23–350.