Adiabatic perturbation theory for the soliton of the nonlinear Dirac equation in 1D

Taras I. Lakoba*

Department of Mathematics and Statistics, University of Vermont, Burlington, VT 05405

February 28, 2026

Abstract

We derive equations for the slow changes of parameters of the soliton of the 1D nonlinear Dirac, or Gross–Neveu, equation under the action of a small perturbation. Our perturbation theory uses the neutral modes of the linearized operator of the nonlinear Dirac equation. In addition to the conventional soliton parameters such as its frequency (related to the amplitude and width), velocity, and shifts of the center and phase, we also account for an additional parameter, related the so-called Bogoliubov symmetry of the Dirac equation, which was first pointed out almost half a century ago and rediscovered in the last decade. This aspect of our theory allows one to explain both asymmetric changes of the soliton profile and large growth of the soliton amplitude, which was observed in previous studies via numerical simulations.

Keywords: Nonlinear Dirac equation, Gross–Neveu equation, soliton perturbation theory.

1 Introduction

This work is dedicated to the memory of my thesis advisor Prof. D.J. Kaup, who made seminal contributions to the perturbation theory of equations integrable by the Inverse Scattering Transform.

*tlakoba@uvm.edu

In this work, we develop a perturbation theory that allows one to find the slow evolution of the parameters of a solitary wave of the one-dimensional Dirac equation with cubic nonlinearity, also known as the massive Gross–Neveu (GN) model [1] or the 1D Soler model [2]:

$$i\left(\vec{\psi}_t+\sigma_1\vec{\psi}_x\right)+\sigma_3\,\left(|\psi_1|^2-|\psi_2|^2-1\right)\,\vec{\psi}=\epsilon\vec{r},\qquad \vec{\psi}\equiv\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix},\tag{1}$$

where subscripts stand for the respective partial derivatives, σ_j, $j=0,1,2,3$ are the standard Pauli matrices, and $\vec{r}$ denotes a perturbation, whose smallness is characterized by $\epsilon\ll 1$. The standing solitary wave (in what follows referred to as a GN soliton, even though the GN equation is not integrable by the Inverse Scattering Transform) of the unperturbed Eq. (1) is given by

$$\vec{\psi}=\overrightarrow{\psi_{(0)}}\,e^{-i\omega t}\;\equiv\;\begin{pmatrix}\psi_{1(0)}\\ \psi_{2(0)}\end{pmatrix}e^{-i\omega t},\qquad \omega\in(0,1)\tag{2}$$

where

$$\begin{aligned}\psi_{1(0)}(x)&=\frac{\sqrt{2(1-\omega)}\,\operatorname{sech}(\beta x)}{1-\mu^2\,\tanh^2(\beta x)}\;\equiv\;\sqrt{2}\beta\,\frac{\sqrt{1+\omega}\,\cosh(\beta x)}{1+\omega\cosh(2\beta x)},\\ \psi_{2(0)}(x)&=i\mu\,\tanh(\beta x)\,\psi_{1(0)},\end{aligned}\tag{3}$$

and $\beta=\sqrt{1-\omega^2}$, $\mu=\sqrt{(1-\omega)/(1+\omega)}$. One obtains a replica of this soliton moving with velocity $v\in(-1,1)$ using the Lorentz invariance of the unperturbed Eq. (1):

$$\left\{x,t,\;\vec{\psi}\right\}\to\left\{x',t',\;\vec{\psi'}\right\};\qquad\begin{pmatrix}x'\\ t'\end{pmatrix}=\gamma(\sigma_0-v\sigma_1)\begin{pmatrix}x\\ t\end{pmatrix},\qquad\vec{\psi'}=T\,\vec{\psi},\tag{4a}$$

where $\gamma=1/\sqrt{1-v^2}$ and

$$T=\cosh\big((\operatorname{arccosh}\gamma)/2\big)\sigma_0+\operatorname{sgn}(v)\sinh\big((\operatorname{arccosh}\gamma)/2\big)\sigma_1.\tag{4b}$$

The dynamics of the GN soliton under various perturbations has been considered in [3, 4, 5, 6, 7, 8] using the collective coordinates method. This method consists of obtaining the Euler–Lagrange equations for the soliton parameters from the Lagrangian density of the GN equation (or, as a variation, from certain conserved quantities).

The perturbation theory that we present here differs from the collective coordinates method in two main aspects. First, we relate the evolution of the soliton parameters to the modes corresponding to zero eigenvectors and generalized (or associate) eigenvectors of a linearized GN operator, denoted in what follows as $\mathbf{L}$. The same approach was used, e.g., in [9] to develop a version of the perturbation theory for the soliton of the Nonlinear Schrödinger equation. Second, and more importantly, our theory captures a (complex-valued) soliton parameter that was not included in the studies cited in the previous paragraph. Namely, it

was shown in [10], following an observation made in [11], that if (2) is a solution of (1) (with $\vec{r} = \vec{0}$), then so is

$$\vec{\psi} = a \overrightarrow{\psi_{(0)}} e^{-i\omega t} + b \sigma_1 \left(\overrightarrow{\psi_{(0)}} e^{-i\omega t} \right)^*, \tag{5}$$

where the constants a, b satisfy

$$|a|^2 - |b|^2 = 1, \tag{6}$$

and the asterisk stands for complex conjugation. Solution (5) was called in [10] a bi-frequency soliton since its components oscillate with frequencies $\pm\omega$. The aforementioned complex-valued parameter, which the previous studies of a perturbed GN soliton did not consider, is the above constant b. (The modulus of a is determined by $|b|$, and its phase can be absorbed in the overall phase of the single-frequency soliton (2), which will be explicitly written out later.)

In fact, the original motivation for this work was to explain features of the evolution of a perturbed GN soliton (2), observed in numerical experiments of [4, 7, 8], which could not be explained by the collective coordinates method. We will show below that the inclusion of b (or an equivalent parameter) in the perturbation theory explains the numerical results which the collective coordinates method of the previous studies, that was based only on the parameters of soliton (2), did not capture.

The linearized GN operator $\mathbf{L}$, mentioned before Eq. (5), was first studied in [12]. There, it was shown that it has only real eigenvalues (thus indicating that the soliton is linearly stable). Furthermore, it was shown that for sufficiently small ω, it has internal modes, i.e., localized modes that do not correspond to shifts of the parameters of soliton (2). Moreover, for *all* $\omega \in (0, 1)$, there are also internal modes with the eigenvalues $\pm 2\omega$; in [10], it was recognized that this mode is a consequence of the existence of the bi-frequency soliton (5).

In this work, we consider only the modes of the operator $\mathbf{L}$ corresponding to infinitesimal shifts of the soliton parameters (which, in the limit $b \to 0$, includes the 2ω-mode, as we will demonstrate). The main aspect in which our $\mathbf{L}$ differs from that considered in [12] is that it is written in the Galilean coordinate frame, i.e. for the variables $(x - vt, t)$. This was needed for the practical purpose of considering a soliton whose velocity is altered by the perturbation. Note that changing from the original (x, t) to the Galilean frame leads to a non-trivial (albeit, as we will show, simple) modification of $\mathbf{L}$, because the GN equation is Lorentzian-, but not Galilean-, invariant. Two other, minor, aspects of our $\mathbf{L}$ compared to that used in [12] will be pointed out as we present its form.

The main part of this work is organized as follows. In Sec. 2 we rewrite the GN equation for a four-component field made of $\vec{\psi}$ and $\vec{\psi}^*$; then we write the linearized GN equation, involving $\mathbf{L}$, in an analogous form. In Sec. 3 we derive modes of $\mathbf{L}$ associated with infinitesimal changes of the soliton parameters, and in Sec. 4 obtain inner products between such modes. In Sec. 5 we use the results of the previous sections to derive the equations for the

evolution of the soliton parameters under the action of a perturbation. Since in most cases those equations can be solved only numerically, we also explain how that can be done. In Secs. 6 and 7, we apply our perturbation theory to, respectively, the soliton perturbed by a linear potential, which was previously studied in [3, 4, 8], and that driven by parametric perturbations, one of which was considered in [7], by the collective coordinates method and direct numerical simulations. The focus of these two sections is on demonstrating that the perturbation theory can explain numerical results that could not be explained by the collective coordinates method. In Sec. 6, the results of the perturbation theory are in reasonable quantitative agreement with the numerics. In one of the examples presented in Sec. 7, the agreement between the two groups of results is excellent, whereas in the other example, it is only qualitative. In Sec. 8 we summarize our results and present two problems for future research that have been revealed but not answered by this work. The Appendices contain technical details of several derivations. A Matlab code implementing the evolution equations for soliton parameters for one of the perturbations considered in Sec. 7, can be found in Ref. [13].

2 Preliminaries

We begin by introducing a notation which will be extensively used below: $\boldsymbol{\Sigma}_i^j$ will denote a 4×4 matrix whose 2×2 *block* structure has the form of σ_i, while the nonzero blocks contain σ_j. For example:

$$\boldsymbol{\Sigma}_3^0 = \begin{pmatrix} \sigma_0 & \mathscr{o} \\ \mathscr{o} & -\sigma_0 \end{pmatrix}, \qquad \boldsymbol{\Sigma}_1^3 = \begin{pmatrix} \mathscr{o} & \sigma_3 \\ \sigma_3 & \mathscr{o} \end{pmatrix}, \tag{7a}$$

where $\mathscr{o}$ denotes the 2×2 zero matrix. Four other 4×4 matrices extensively used below are:

$$\boldsymbol{T} = \begin{pmatrix} T & \mathscr{o} \\ \mathscr{o} & T \end{pmatrix}, \qquad \boldsymbol{C} = \begin{pmatrix} a\,\sigma_0 & b\,\sigma_1 \\ b^*\sigma_1 & a^*\sigma_0 \end{pmatrix}, \qquad \boldsymbol{\Gamma}^{\pm} = \gamma \begin{pmatrix} \sigma_0 \pm v\sigma_1 & \mathscr{o} \\ \mathscr{o} & \mp(\sigma_0 \pm v\sigma_1) \end{pmatrix}, \tag{7b}$$

where T is defined in (4b) and a, b, in (5).

Next, in what follows it will be convenient to write (1) together with its complex conjugate as

$$i\left(\partial_t \overrightarrow{\boldsymbol{\psi}}_{\rm ext} + \boldsymbol{\Sigma}_0^1\,\partial_x \overrightarrow{\boldsymbol{\psi}}_{\rm ext}\right) + \boldsymbol{\Sigma}_3^3 \overrightarrow{\boldsymbol{\psi}}_{\rm ext}\left(\frac{1}{2}\overrightarrow{\boldsymbol{\psi}}_{\rm ext}^T \boldsymbol{\Sigma}_1^3 \overrightarrow{\boldsymbol{\psi}}_{\rm ext} - 1\right) = \epsilon\,\overrightarrow{\boldsymbol{r}}, \tag{8}$$

where

$$\overrightarrow{\boldsymbol{\psi}}_{\rm ext} = \begin{pmatrix} \vec{\psi} \\ \vec{\psi}^* \end{pmatrix}, \qquad \overrightarrow{\boldsymbol{r}} = \begin{pmatrix} \vec{r} \\ -\vec{r}^* \end{pmatrix}, \tag{9}$$

and subscript 'ext' stands for 'extended'. By a straightforward calculation (see the first relations in Eqs. (74) and (75) in Appendix A), it can be shown that the l.h.s. of (8) is invariant under the transformation

$$\overrightarrow{\boldsymbol{\psi}}_{\rm ext} \to \boldsymbol{C}\,\overrightarrow{\boldsymbol{\psi}}_{\rm ext}. \tag{10}$$

The corresponding derivation relies on (6) and the multiplication properties of the Pauli matrices:

$$\sigma_j^2 = \sigma_0, \qquad \sigma_k\sigma_j = -\sigma_j\sigma_k, \qquad j,k = 1,2,3, \quad i \neq j; \tag{11a}$$

$$\sigma_j\sigma_k = i\,\sigma_l, \qquad j,k,l \text{ are cyclic permutations of 1,2,3.} \tag{11b}$$

These properties are used extensively in the derivations below. Relation (10) generalizes (5) for an arbitrary solution (i.e., not just the soliton).

We will now write down the expression for $\overrightarrow{\boldsymbol{\psi}}_{\rm ext}$ focusing on its having a nonzero velocity v. Equation (8), of course, is also invariant under the four-dimensional extension of (4): $\overrightarrow{\boldsymbol{\psi}}'_{\rm ext} = \boldsymbol{T}\overrightarrow{\boldsymbol{\psi}}_{\rm ext}$. However, considering the soliton with $v = 0$ and then transforming it to the moving frame by the above — Lorentz — transformation will *not* fit our purpose of allowing a perturbed soliton to have a varying velocity. This is because the natural frame in which one observes a moving soliton (e.g., in numerical experiments) is related to the original (x,t) by the Galilean boost

$$(x,t) \to (x - vt, t) \tag{12}$$

rather than by the Lorentzian one, (4). Therefore, we will rewrite the expression for the moving soliton (2), (4) in the Galilean frame (12):

$$\overrightarrow{\boldsymbol{\psi}}_{\rm ext}(z,t) = \boldsymbol{C}\,\boldsymbol{T}\,\boldsymbol{\Phi}\,\overrightarrow{\boldsymbol{\psi}}_{(0)}, \qquad \boldsymbol{\Phi} = \begin{pmatrix} e^{i\phi(z,t)}\sigma_0 & \mathscr{O} \\ \mathscr{O} & e^{-i\phi(z,t)}\sigma_0 \end{pmatrix}, \qquad \overrightarrow{\boldsymbol{\psi}}_{(0)} = \begin{pmatrix} \overrightarrow{\psi_{(0)}}(z) \\ \overrightarrow{\psi_{(0)}}^{\,*}(z) \end{pmatrix}, \tag{13a}$$

where

$$z = \gamma\left(x - \int_0^t v(s)\,ds - x_0\right), \qquad \phi = -\int_0^t \frac{\omega(s)}{\gamma(s)}\,ds + v\omega\check{z} + \phi_0, \quad \check{z} = z + \gamma x_0. \tag{13b}$$

Here, x_0 and ϕ_0 are arbitrary (and, at this stage, constant) shifts of the soliton's center and phase, and we have written the terms vt and $(\omega/\gamma)t$ as integrals for future use when these parameters will be slow functions of time due to a perturbation. Note that matrix $\boldsymbol{T}$ commutes with both $\boldsymbol{\Phi}$ and $\boldsymbol{C}$, but the latter two matrices do not commute. Also note, from (2) and (3), that

$$\overrightarrow{\psi_{(0)}}^{\,*} = \sigma_3\,\overrightarrow{\psi_{(0)}}. \tag{14}$$

Since we have changed coordinates from (x,t) to (z,t), we rewrite (8) in these coordinates:

$$i\left(\partial_t\overrightarrow{\boldsymbol{\psi}}_{\rm ext} + \gamma\big(\boldsymbol{\Sigma}_0^1 - v\,\boldsymbol{\Sigma}_0^0\big)\,\partial_z\overrightarrow{\boldsymbol{\psi}}_{\rm ext}\right) + \boldsymbol{\Sigma}_3^3\overrightarrow{\boldsymbol{\psi}}_{\rm ext}\left(\frac{1}{2}\overrightarrow{\boldsymbol{\psi}}_{\rm ext}^{T}\boldsymbol{\Sigma}_1^3\overrightarrow{\boldsymbol{\psi}}_{\rm ext} - 1\right) = \epsilon\,\overrightarrow{\boldsymbol{r}}. \tag{15}$$

To complete this section, we present the linearization of the above equation. Specifically, we substitute

$$\overrightarrow{\boldsymbol{\psi}}_{\rm ext} = \boldsymbol{C}\,\boldsymbol{T}\,\boldsymbol{\Phi}\,\left(\overrightarrow{\boldsymbol{\psi}}_{(0)}(z) + \epsilon\,\boldsymbol{\delta}\overrightarrow{\boldsymbol{\psi}}(z,t)\right), \qquad \boldsymbol{\delta}\overrightarrow{\boldsymbol{\psi}} = \begin{pmatrix} \overrightarrow{\delta\psi}(z,t) \\ \overrightarrow{\delta\psi}^*(z,t) \end{pmatrix}, \tag{16}$$

into (15), multiply the result by $(\boldsymbol{C}\boldsymbol{T}\boldsymbol{\Phi})^{-1}$ and obtain:

$$\left(i\partial_t + \boldsymbol{\Gamma}^-\,\mathbf{L}_0\right)\boldsymbol{\delta}\overrightarrow{\boldsymbol{\psi}} = (\boldsymbol{C}\boldsymbol{T}\boldsymbol{\Phi})^{-1}\,\overrightarrow{\boldsymbol{r}}, \tag{17a}$$

where $\boldsymbol{\Gamma}^-$ is defined in (7b), and

$$\mathbf{L}_0 = \omega\,\boldsymbol{\Sigma}_3^0 + i\boldsymbol{\Sigma}_0^1\,\partial_z + n_{(0)}\,\boldsymbol{\Sigma}_3^3 + \mathbf{P}, \tag{17b}$$

$$n_{(0)} \equiv \left(\frac{1}{2}\overrightarrow{\boldsymbol{\psi}}_{(0)}^T\boldsymbol{\Sigma}_1^3\overrightarrow{\boldsymbol{\psi}}_{(0)} - 1\right), \tag{17c}$$

$$\mathbf{P} = \boldsymbol{\Sigma}_3^3\overrightarrow{\boldsymbol{\psi}}_{(0)}\overrightarrow{\boldsymbol{\psi}}_{(0)}^T\boldsymbol{\Sigma}_1^3 = \boldsymbol{\Sigma}_3^0\begin{pmatrix} P & P\,\sigma_3 \\ P^*\sigma_3 & P^* \end{pmatrix}, \qquad P = \sigma_3\,\overrightarrow{\psi_{(0)}}(z)\,\overrightarrow{\psi_{(0)}}^\dagger(z)\,\sigma_3, \tag{17d}$$

where † denotes Hermitian conjugation and $\overrightarrow{\psi_{(0)}}$ and $\overrightarrow{\boldsymbol{\psi}}_{(0)}$ are defined in (2) and (13), respectively. The derivation of this result is outlined in Appendix A.

Operator $\mathbf{L}_0$ is equivalent to the linearized operator presented in [12] for $v = 0$. The former differs from the latter cosmetically in two aspects. First, following the approach of [9], it is written for the vector $\boldsymbol{\delta}\overrightarrow{\boldsymbol{\psi}}$ instead of the real and imaginary parts of the components of $\overrightarrow{\delta\psi}$. This appears to be suitable for the GN equation given its invariance under transformation (10), which involves vector $\overrightarrow{\boldsymbol{\psi}}_{\rm ext}$. Second, the form in which we wrote the last term in (17b) and the 2×2 matrix P in (17d) reveals a structure that was not exhibited in [12]. Using this structure, one obtains four relations, which will be used in what follows:

$$\mathbf{P}\left(\begin{pmatrix} \sigma_2 & \mathscr{o} \\ \mathscr{o} & \mathscr{o} \end{pmatrix}\overrightarrow{\boldsymbol{\psi}}_{(0)}\right) = \mathbf{P}\left(\begin{pmatrix} \mathscr{o} & \mathscr{o} \\ \mathscr{o} & \sigma_2 \end{pmatrix}\overrightarrow{\boldsymbol{\psi}}_{(0)}\right) = \vec{\mathbf{0}}, \tag{18a}$$

$$\mathbf{P}(\boldsymbol{\Sigma}_0^1\overrightarrow{\boldsymbol{\psi}}_{(0)}) = \vec{\mathbf{0}}, \tag{18b}$$

$$\mathbf{P}(\boldsymbol{\Sigma}_3^0\overrightarrow{\boldsymbol{\psi}}_{(0)}) = \vec{\mathbf{0}}. \tag{18c}$$

Their derivation is also outlined in Appendix A.

The linearized operator of the GN equation written in a Galilean frame, i.e.,

$$\mathbf{L} \equiv \boldsymbol{\Gamma}^-\mathbf{L}_0 \tag{19}$$

(see (17a)) is related to $\mathbf{L}_0$ in a simple way, which, however, does not seem to be possible to predict at the outset without detailed calculations.

3 Neutral modes of L

3.1 Bi-frequency soliton (5) with arbitrary b

We now allow parameters of the GN soliton to be slow functions of time due to the action of perturbation $\vec{r}$. Neutral modes of $\mathbf{L}$ will be those modes corresponding to the shifts of the soliton parameters. Derivations in this section are quite technical, and the reader who is only interested in its main results may focus just on Eqs. (20), (21a), (22), (25), and (35).

Before continuing, we rewrite parameters a, b defined in (5), (6), in an equivalent form:

$$a = \cosh c, \qquad b = e^{i\Delta\phi} \sinh c, \qquad c \geq 0. \tag{20}$$

Note that the phase of a has been "absorbed" into ϕ_0 in (13b). Thus, $\phi_0, x_0, \omega, v, c, \Delta\phi$ are assumed to be functions of a 'slow' time $\tau \equiv \epsilon t$.

The operator ∂_t in (15) can then be written as

$$\partial_t \to \partial_t + \epsilon\partial_\tau \;=\; \partial_t + \epsilon \sum_{\alpha=\phi_0,x_0,\omega,v,c,\Delta\phi} \dot{\alpha}\, \widetilde{\vec{\boldsymbol{m}}_\alpha}, \tag{21a}$$

where the overdot stands for ∂_τ and, given the form of (16),

$$\widetilde{\vec{\boldsymbol{m}}_\alpha} = (\boldsymbol{C}\boldsymbol{T}\boldsymbol{\Phi})^{-1}\partial_\alpha \left(\boldsymbol{C}\boldsymbol{T}\boldsymbol{\Phi}\, \vec{\boldsymbol{\psi}}_{(0)}\right). \tag{21b}$$

By a straightforward calculation, one obtains from the last equation:

$$\widetilde{\vec{\boldsymbol{m}}_{\phi_0}} = \vec{\boldsymbol{m}}_{\phi_0} = \begin{pmatrix} i\,\overrightarrow{\psi_{(0)}}(z) \\ \overrightarrow{\mathrm{c.c.}} \end{pmatrix} \equiv i\,\boldsymbol{\Sigma}_3^0\, \vec{\boldsymbol{\psi}}_{(0)}, \tag{22a}$$

$$\widetilde{\vec{\boldsymbol{m}}_{x_0}} = -\gamma\vec{\boldsymbol{m}}_{x_0}, \qquad \vec{\boldsymbol{m}}_{x_0} = \begin{pmatrix} \partial_z\overrightarrow{\psi_{(0)}}(z) \\ \overrightarrow{\mathrm{c.c.}} \end{pmatrix}, \tag{22b}$$

$$\widetilde{\vec{\boldsymbol{m}}_{\omega}} = \vec{\boldsymbol{m}}_{\omega} + \gamma v x_0 \vec{\boldsymbol{m}}_{\phi_0}, \qquad \vec{\boldsymbol{m}}_{\omega} = \begin{pmatrix} \partial_\omega\overrightarrow{\psi_{(0)}}(z) \\ \overrightarrow{\mathrm{c.c.}} \end{pmatrix} + \begin{pmatrix} ivz\,\overrightarrow{\psi_{(0)}}(z) \\ \overrightarrow{\mathrm{c.c.}} \end{pmatrix}, \tag{22c}$$

$$\widetilde{\vec{\boldsymbol{m}}_{v}} = \gamma^2\left(\vec{\boldsymbol{m}}_{v} + \gamma\omega x_0 \vec{\boldsymbol{m}}_{\phi_0}\right), \qquad \vec{\boldsymbol{m}}_{v} = \begin{pmatrix} i\omega z\,\overrightarrow{\psi_{(0)}}(z) \\ \overrightarrow{\mathrm{c.c.}} \end{pmatrix} + \begin{pmatrix} vz\,\partial_z\overrightarrow{\psi_{(0)}}(z) \\ \overrightarrow{\mathrm{c.c.}} \end{pmatrix} + \frac{1}{2}\begin{pmatrix} \sigma_1\,\overrightarrow{\psi_{(0)}}(z) \\ \overrightarrow{\mathrm{c.c.}} \end{pmatrix}, \tag{22d}$$

$$\begin{gathered} \widetilde{\vec{\boldsymbol{m}}_{c}} = -e^{i\widetilde{\varphi}}\,\vec{\boldsymbol{m}}_{c+} - e^{-i\widetilde{\varphi}}\,\vec{\boldsymbol{m}}_{c-}, \qquad \widetilde{\varphi} \equiv \Delta\phi + 2\int_0^t \frac{\omega(s)}{\gamma(s)}ds - 2(\phi_0 + \gamma\omega v x_0), \\ \vec{\boldsymbol{m}}_{c+} = \begin{pmatrix} i\,e^{-2i\omega v z}\,\sigma_2\,\overrightarrow{\psi_{(0)}}(z) \\ \vec{0} \end{pmatrix}, \quad \vec{\boldsymbol{m}}_{c-} = \begin{pmatrix} \vec{0} \\ i\,e^{2i\omega v z}\,\sigma_2\,\overrightarrow{\psi_{(0)}}^{*}(z) \end{pmatrix}, \end{gathered} \tag{22e}$$

$$\widetilde{\overrightarrow{\boldsymbol{m}}_{\Delta\phi}} = \sinh c \, \left(\cosh c \, i \boldsymbol{\Sigma}_3^0 \widetilde{\overrightarrow{\boldsymbol{m}}_c} + \sinh c \, \overrightarrow{\boldsymbol{m}}_{\phi_0} \right), \tag{22f}$$

where '$\overrightarrow{\text{cc}}$' stands for the complex conjugate of the vector written immediately above it. The vectors $\overrightarrow{\boldsymbol{m}}_\alpha$ will be shown below to be eigenvectors or associate eigenvectors of $\mathbf{L}$. In writing (22b), we have used the fact that in (13b), $\check{z}$ and hence $\phi(z,t)$ do not depend on x_0. In computing (22c) and (22d), we used the fact that for any function $f(\omega)$,

$$\partial_{\omega(t)} \int^t f(\omega(s))\, ds = 0, \tag{23}$$

and similarly for the integral of a function of v. A similar fact was used in the derivation of modes in [9] but not explained there; therefore, we explain it in Appendix B. The terms proportional to x_0 in (22c) and (22d) are due to the relation between $\check{z}$ and z: see (13). In (22d), the second term in $\overrightarrow{\boldsymbol{m}}_v$ was obtained using

$$\partial_v z = (\partial_v \gamma/\gamma)\, z, \tag{24a}$$

and the third one, using

$$T^{-1}\partial_v T = (\gamma^2/2)\sigma_1. \tag{24b}$$

Finally, in obtaining (22e), we have used (14).

We will now demonstrate that:

$$\mathbf{L}\, \overrightarrow{\boldsymbol{m}}_{\{\phi_0,\, x_0\}} = \vec{0}, \tag{25a}$$

$$\mathbf{L}\, \overrightarrow{\boldsymbol{m}}_{\{\omega,\, v\}} = \frac{i}{\gamma}\, \overrightarrow{\boldsymbol{m}}_{\{\phi_0,\, x_0\}}, \tag{25b}$$

$$\mathbf{L}\, \overrightarrow{\boldsymbol{m}}_{c\,\pm} = \pm \frac{2\omega}{\gamma} \overrightarrow{\boldsymbol{m}}_{c\,\pm}. \tag{25c}$$

This demonstration relies on the fact that when $\overrightarrow{\boldsymbol{\psi}}_{\text{ext}} = \boldsymbol{C}\boldsymbol{T}\boldsymbol{\Phi}\, \overrightarrow{\boldsymbol{\psi}}_{(0)}$ is substituted into (15) with $\overrightarrow{\boldsymbol{r}} = \vec{\mathbf{0}}$, it results in

$$\boldsymbol{C}\boldsymbol{T}\boldsymbol{\Phi}\, \boldsymbol{\Gamma}^- \, \left(\omega \boldsymbol{\Sigma}_3^0 \overrightarrow{\boldsymbol{\psi}}_{(0)} + i\, \boldsymbol{\Sigma}_0^1\, \partial_z \overrightarrow{\boldsymbol{\psi}}_{(0)} + \boldsymbol{\Sigma}_3^3 \overrightarrow{\boldsymbol{\psi}}_{(0)}\, n_{(0)} \right) = \vec{\mathbf{0}}, \tag{26}$$

where $n_{(0)}$ is defined in (17c). The above equation is obtained using the relations

$$\boldsymbol{\Phi}^{-1}\partial_t \boldsymbol{\Phi} = -i\frac{\omega}{\gamma}\, \boldsymbol{\Sigma}_3^0, \qquad \boldsymbol{\Phi}^{-1}\partial_z \boldsymbol{\Phi} = i\omega v\, \boldsymbol{\Sigma}_3^0, \tag{27}$$

which follow from (13), and the commutation relations stated in Appendix A.

Differentiating (26) with respect to (w.r.t.) ϕ_0 and multiplying the result by $(\boldsymbol{C}\boldsymbol{T}\boldsymbol{\Phi})^{-1}$ on the left yields

$$i\boldsymbol{\Sigma}_3^0\, \boldsymbol{\Gamma}^- \, \left(\omega \boldsymbol{\Sigma}_3^0 \overrightarrow{\boldsymbol{\psi}}_{(0)} + i\, \boldsymbol{\Sigma}_0^1\, \partial_z \overrightarrow{\boldsymbol{\psi}}_{(0)} + \boldsymbol{\Sigma}_3^3 \overrightarrow{\boldsymbol{\psi}}_{(0)}\, n_{(0)} \right). \tag{28}$$

The first equation in (25a) then follows by noticing that $\boldsymbol{\Sigma}_3^0$ commutes with all the other matrices in (28) and using (18c). Combining the above facts, one can rewrite (28) as the first equation in (25a).

The second equation in (25a) is obtained straightforwardly since none of $\boldsymbol{C}$, $\boldsymbol{T}$, $\boldsymbol{\Phi}$, and $\boldsymbol{\Gamma}^-$ depend on x_0.

To obtain the first equation in (25b), one differentiates (26) w.r.t. ω, multiplies by $(\boldsymbol{CT\Phi})^{-1}$ and, using a relation $\boldsymbol{\Phi}^{-1}\partial_\omega \boldsymbol{\Phi} = ivz\,\boldsymbol{\Sigma}_3^0$, obtains:

$$ivz\boldsymbol{\Sigma}_3^0\boldsymbol{\Gamma}^-(\mathbf{L}_0-\mathbf{P})\overrightarrow{\boldsymbol{\psi}}_{(0)} + \mathbf{L}\partial_\omega\overrightarrow{\boldsymbol{\psi}}_{(0)} + \boldsymbol{\Gamma}^-\boldsymbol{\Sigma}_3^0\overrightarrow{\boldsymbol{\psi}}_{(0)} = \vec{0}. \tag{29}$$

The first term above is equivalently rewritten as:

$$ivz\boldsymbol{\Sigma}_3^0\boldsymbol{\Gamma}^-(\mathbf{L}_0-\mathbf{P})\overrightarrow{\boldsymbol{\psi}}_{(0)} = \boldsymbol{\Gamma}^-(\mathbf{L}_0-\mathbf{P})\left(ivz\boldsymbol{\Sigma}_3^0\overrightarrow{\boldsymbol{\psi}}_{(0)}\right) + \boldsymbol{\Gamma}^-\boldsymbol{\Sigma}_3^1 v\overrightarrow{\boldsymbol{\psi}}_{(0)}. \tag{30}$$

The last term in (30) was obtained using the identity

$$z\,\partial_z\overrightarrow{\boldsymbol{\psi}}_{(0)} = \partial_z\left(z\overrightarrow{\boldsymbol{\psi}}_{(0)}\right) - \overrightarrow{\boldsymbol{\psi}}_{(0)}. \tag{31}$$

This term combined with the last term on the l.h.s. of (29) yields the negative of the r.h.s. of the sought equation in (25b). Then, using (18c) in the first term on the l.h.s. (29) and combining the result with the second term on the l.h.s. of (29), one obtains the l.h.s. of the first equation in (25b).

The derivation of the second equation in (25b) is even more technical and therefore is presented in Appendix C.

Equations (25c), unlike (25a) and (25b), have no counterpart in other nonlinear wave equations (e.g., in the Nonlinear Schrödinger equation). To begin their derivation, one differentiates (26) w.r.t. c and computes

$$(\boldsymbol{CT\Phi})^{-1}\,\partial_c\,\boldsymbol{CT\Phi} = \begin{pmatrix} \mathscr{O} & e^{\,i(\widetilde{\varphi}-2wvz)}\sigma_1 \\ e^{-i(\widetilde{\varphi}-2wvz)}\sigma_1 & \mathscr{O} \end{pmatrix}, \tag{32a}$$

where $\widetilde{\varphi}$ is defined in (22e). Also, using the definition of $\overrightarrow{\boldsymbol{\psi}}_{(0)}$ and relation (14) (see also the derivation of (80)), one notices that

$$\begin{pmatrix} \mathscr{O} & e^{\,i(\widetilde{\varphi}-2wvz)}\sigma_1 \\ e^{-i(\widetilde{\varphi}-2wvz)}\sigma_1 & \mathscr{O} \end{pmatrix}\overrightarrow{\boldsymbol{\psi}}_{(0)} = \widetilde{\overrightarrow{\boldsymbol{m}}}_c. \tag{32b}$$

Since no term other than $\boldsymbol{C}$ in (26) depends on c, our goal now is to "move" the matrix on the r.h.s. of (32a) so as to place it in front of $\overrightarrow{\boldsymbol{\psi}}_{(0)}$ in (26). This matrix commutes with $\boldsymbol{\Gamma}^-$, $\boldsymbol{\Sigma}_0^1$, and $\boldsymbol{\Sigma}_3^3$ and anti-commutes with $\boldsymbol{\Sigma}_3^0$. Therefore, using the form of $\phi(z,t)$ from (13b) and the identity (31), one has:

$$\boldsymbol{\Gamma}^-\left(-2\omega\boldsymbol{\Sigma}_3^0 + \omega\boldsymbol{\Sigma}_3^0 + i\boldsymbol{\Sigma}_0^1\partial_z + \boldsymbol{\Sigma}_3^3 n_{(0)}\right)\widetilde{\overrightarrow{\boldsymbol{m}}}_c + 2\omega v\boldsymbol{\Sigma}_3^1\widetilde{\overrightarrow{\boldsymbol{m}}}_c = \vec{0}. \tag{33}$$

Combining the first and last terms on the l.h.s. of (33), one finds:

$$\boldsymbol{\Gamma}^{-}(\mathbf{L}_0 - \mathbf{P})\widetilde{\overrightarrow{\boldsymbol{m}}_c} = \mathbf{L}\widetilde{\overrightarrow{\boldsymbol{m}}_c} = \frac{2\omega}{\gamma}\boldsymbol{\Sigma}_3^0\widetilde{\overrightarrow{\boldsymbol{m}}_c}. \tag{34}$$

The first equation above was obtained using the definition (22e) of $\overrightarrow{\boldsymbol{m}}_c$ and Eqs. (18a). The last equation in (34) is not quite the regular eigenvalue problem because of the presence of $\boldsymbol{\Sigma}_3^0$. However, note that $\boldsymbol{\Gamma}^{-}(\mathbf{L}_0 - \mathbf{P})$ is a block-diagonal operator, and therefore one can rewrite (34) as two independent equations for its 2×2 diagonal blocks, which yields both equations in (25c).

3.2 Single-frequency soliton (2)

This case is obtained from the case of the bi-frequency soliton (5) in the limit $c \to 0$. The only change that occurs in the analysis of the previous subsection is that in this limit, $\widetilde{\overrightarrow{\boldsymbol{m}}_{\Delta\phi}} = \vec{0}$.

Also, in this case it can be more intuitive to exclude $\alpha = \{c, \Delta\phi\}$ from the sum in (21a) and instead consider $\boldsymbol{\delta}\overrightarrow{\boldsymbol{\psi}}$ to be a linear combination of modes $\overrightarrow{\boldsymbol{m}}_{c\pm}$:

$$\boldsymbol{\delta}\overrightarrow{\boldsymbol{\psi}} = k_+(t)\,\overrightarrow{\boldsymbol{m}}_{c+}\,\exp\left[2i\int_0^t \frac{\omega(s)}{\gamma(s)}\,ds\right] \,+\, k_-(t)\,\overrightarrow{\boldsymbol{m}}_{c-}\,\exp\left[-2i\int_0^t \frac{\omega(s)}{\gamma(s)}\,ds\right]. \tag{35}$$

One can view the exponentials above as being inherited from $\widetilde{\varphi}$ in (22e). Alternatively, one can say that they are due to the eigenvalues of the modes $\overrightarrow{\boldsymbol{m}}_{c\pm}$ in (25c). These modes are the '2ω-modes', first found (in a different form and for $v = 0$ only) in [12]. Finally, note that the dependence of the coefficients $k_\pm$ on time can occur due to the perturbation $\vec{r}$.

4 Inner products between the neutral modes

The main results of this section are contained in Eqs. (41), (44), and (45). The reader not interested in their derivation can skip right to these equations.

Equations of the adiabatic perturbation theory, which we will present in the next section, require the values of inner products between the neutral modes. Those, in turn, require the weight with which the inner product is taken. By analogy with [9], we expect this weight to be related to $\boldsymbol{\Sigma}_3^0$. We clarify this expectation below.

The linearized GN operator $\mathbf{L} = \boldsymbol{\Gamma}^{-}\mathbf{L}_0$ can be equivalently written as

$$\mathbf{L} = \left(\boldsymbol{\Gamma}^{-}\boldsymbol{\Sigma}_3^0\right)\mathbf{L}_H, \qquad \mathbf{L}_H = \omega\boldsymbol{\Sigma}_0^0 + i\boldsymbol{\Sigma}_3^1\partial_z + n_{(0)}\boldsymbol{\Sigma}_0^3 + \begin{pmatrix} P & P\sigma_3 \\ P^*\sigma_3 & P^* \end{pmatrix}. \tag{36}$$

Since P is Hermitian (see (17d)), then so is $\mathbf{L}_H$. Next, consider the eigenvalue problem for $\mathbf{L}$,

$$\left(\boldsymbol{\Gamma}^{-}\boldsymbol{\Sigma}_3^0\right)\mathbf{L}_H\overrightarrow{\boldsymbol{v}} = \lambda\overrightarrow{\boldsymbol{v}}, \tag{37a}$$

and its adjoint,

$$\overrightarrow{\boldsymbol{v}}^{\,A}\left(\boldsymbol{\Gamma}^{-}\boldsymbol{\Sigma}_3^0\right)\mathbf{L}_H=\lambda\,\overrightarrow{\boldsymbol{v}}^{\,A}, \tag{37b}$$

where $\overrightarrow{\boldsymbol{v}}$ and $\overrightarrow{\boldsymbol{v}}^{\,A}$ are the corresponding eigenvectors. In the last equation, the derivative in $\mathbf{L}_H$ is assumed to be applied to $\overrightarrow{\boldsymbol{v}}^{\,A}$. Taking the Hermitian conjugate of (37b) and multiplying both sides by $\left(\boldsymbol{\Gamma}^{-}\boldsymbol{\Sigma}_3^0\right)$ yields:

$$\left(\boldsymbol{\Gamma}^{-}\boldsymbol{\Sigma}_3^0\right)\mathbf{L}_H\left(\boldsymbol{\Gamma}^{-}\boldsymbol{\Sigma}_3^0\right)\left(\overrightarrow{\boldsymbol{v}}^{\,A}\right)^{\dagger}=\lambda\left(\boldsymbol{\Gamma}^{-}\boldsymbol{\Sigma}_3^0\right)\left(\overrightarrow{\boldsymbol{v}}^{\,A}\right)^{\dagger}, \tag{38}$$

whence, by comparison with (37a):

$$\left(\boldsymbol{\Gamma}^{-}\boldsymbol{\Sigma}_3^0\right)\left(\overrightarrow{\boldsymbol{v}}^{\,A}\right)^{\dagger}=\overrightarrow{\boldsymbol{v}}\qquad\Rightarrow\qquad\overrightarrow{\boldsymbol{v}}^{\,A}=\overrightarrow{\boldsymbol{v}}^{\,\dagger}\,\boldsymbol{\Gamma}^{+}, \tag{39}$$

where $\boldsymbol{\Gamma}^{+}$ is defined in (7b). The structure of the inner product is defined by the relation that follows from, say, (37a) with $\lambda=\lambda_1$:

$$\overrightarrow{\boldsymbol{v}}^{\,A}(z,\lambda_2)\mathbf{L}\,\overrightarrow{\boldsymbol{v}}(z,\lambda_1)=\lambda_1\,\overrightarrow{\boldsymbol{v}}^{\,A}(z,\lambda_2)\,\overrightarrow{\boldsymbol{v}}(z,\lambda_1); \tag{40a}$$

similarly, from (37b) with $\lambda=\lambda_2$ one has:

$$\overrightarrow{\boldsymbol{v}}^{\,A}(z,\lambda_2)\mathbf{L}\,\overrightarrow{\boldsymbol{v}}(z,\lambda_1)=\lambda_2\,\overrightarrow{\boldsymbol{v}}^{\,A}(z,\lambda_2)\,\overrightarrow{\boldsymbol{v}}(z,\lambda_1). \tag{40b}$$

Together, Eqs. (39) and (40) imply that the eigenvectors of $\mathbf{L}$ corresponding to different eigenvalues are orthogonal with respect to the inner product defined as:

$$\langle\overrightarrow{\boldsymbol{v}}_1|\overrightarrow{\boldsymbol{v}}_2\rangle\equiv\int_{-\infty}^{\infty}dz\;\overrightarrow{\boldsymbol{v}}_1^{\,\dagger}\boldsymbol{\Gamma}^{+}\overrightarrow{\boldsymbol{v}}_2. \tag{41}$$

Thus, the weight of the inner product is $\boldsymbol{\Gamma}^{+}$, and the orthogonality relation between eigenvectors of $\mathbf{L}$ with different eigenvalues is:

$$\langle\overrightarrow{\boldsymbol{v}}(z,\lambda_1)|\overrightarrow{\boldsymbol{v}}(z,\lambda_2)\rangle=0\qquad\text{for }\lambda_1\neq\lambda_2. \tag{42}$$

Also, for *any* two vectors, one has

$$\langle\overrightarrow{\boldsymbol{v}}_2|\overrightarrow{\boldsymbol{v}}_1\rangle\;=\;\langle\overrightarrow{\boldsymbol{v}}_1|\overrightarrow{\boldsymbol{v}}_2\rangle^{\,*}. \tag{43}$$

As a side note, we point out a more intuitive, albeit less rigorous, way to derive the weight for the inner product. First, one observes that for $v=0$ and $c=0$, that weight is $\boldsymbol{\Sigma}_3^0$. This can be established either along the above lines, or from the structure of the linearized Lagrangian density [14]. Then, for $v\neq0$ and $c\neq0$, one uses the form (16) of the perturbation to the soliton and thereby finds that the matrix "inserted" between $\overrightarrow{\boldsymbol{v}}^{\,\dagger}$ and $\overrightarrow{\boldsymbol{v}}$ must be

$$(\boldsymbol{C}\boldsymbol{T}\boldsymbol{\Phi})^{\dagger}\,\boldsymbol{\Sigma}_3^0\,(\boldsymbol{C}\boldsymbol{T}\boldsymbol{\Phi})=\boldsymbol{\Gamma}^{+}.$$

The inner products between the zero eigenvectors $\overrightarrow{\boldsymbol{m}}_{\phi_0,x_0}$ of $\mathbf{L}$ and their associate eigenvectors $\overrightarrow{\boldsymbol{m}}_{\omega,v}$ can be found by direct calculation with Wolfram Mathematica:

$$\langle \overrightarrow{\boldsymbol{m}}_{\phi_0} | \overrightarrow{\boldsymbol{m}}_{\omega} \rangle = \frac{2i\gamma}{\omega^2\sqrt{1-\omega^2}}, \tag{44a}$$

$$\langle \overrightarrow{\boldsymbol{m}}_{x_0} | \overrightarrow{\boldsymbol{m}}_{v} \rangle = 2i\gamma \, \ln \frac{1-\sqrt{1-\omega^2}}{\omega}. \tag{44b}$$

Inner products involving any other pairs of these modes are zero. This is similar to the inner product relations presented in [9]. Additionally,

$$\langle \overrightarrow{\boldsymbol{m}}_{c\pm} | \overrightarrow{\boldsymbol{m}}_{c\pm} \rangle = \pm \frac{2\gamma\sqrt{1-\omega^2}}{\omega}. \tag{45}$$

The inner products of $\overrightarrow{\boldsymbol{m}}_{c\pm}$ with any of $\overrightarrow{\boldsymbol{m}}_{\phi_0,x_0,\omega,v}$ are zero, in agreement with (42).

5 Equations of the adiabatic perturbation theory and their solution

We will first derive the aforementioned equations for the (general) case of the bi-frequency soliton and then point out how they need to be modified for the single-frequency soliton. The main results for the bi-frequency soliton are Eqs. (50a)–(50d) and (51). For the single-frequency one, Eqs. (51) are to be replaced with (53).

Then we will explain how these equations can be solved numerically, because, with the exception of perhaps a few simplest forms of the perturbation $\vec{r}$, the inner products in those equations cannot be evaluated analytically. For two specific forms of the perturbation (54), given by Eqs. (55) and (62), the main results are given by Eqs. (56) and (63), respectively.

5.1 Bi-frequency soliton

Substituting (21a) into (17a), one obtains:

$$\left(i\partial_t + \boldsymbol{\Gamma}^- \mathbf{L}_0\right) \boldsymbol{\delta}\overrightarrow{\boldsymbol{\psi}} = (\boldsymbol{C}\boldsymbol{T}\boldsymbol{\Phi})^{-1} \overrightarrow{\boldsymbol{r}} \; - \; i \sum_{\alpha=\phi_0,x_0,\omega,v,c,\Delta\phi} \dot{\alpha} \, \widetilde{\overrightarrow{\boldsymbol{m}}_{\alpha}} \, . \tag{46}$$

By the standard argument of avoiding secular growth of any part of $\boldsymbol{\delta}\overrightarrow{\boldsymbol{\psi}}$ (see Appendix D), we require that the inner products of certain frequency components of the r.h.s. of (46) with any of the modes listed on the l.h.s. of (25) vanish. Specifically, let us introduce a notation for any function $f(t)$ of the 'fast' time t:

$$[\![f]\!]_{\eta} = \epsilon \int_0^{1/\epsilon} ds \; f(s) \, e^{-i\eta s}, \tag{47}$$

where the separation of the 'fast' and 'slow' times was introduced in (21a). That is, $[\![f]\!]_\eta$ is the Fourier amplitude of the frequency η in $f(t)$. Then, on the account of Eqs. (22), the aforementioned secularity conditions yield the following equations for the slow evolution of the soliton parameters:

$$\dot{\omega}\ (i\langle\overrightarrow{\boldsymbol{m}}_{\phi_0}|\overrightarrow{\boldsymbol{m}}_{\omega}\rangle) = \langle\overrightarrow{\boldsymbol{m}}_{\phi_0}|\ [\![(\boldsymbol{C}\boldsymbol{T}\boldsymbol{\Phi})^{-1}\ \overrightarrow{\boldsymbol{r}}\ ]\!]_0\ \rangle\,, \tag{48a}$$

$$\dot{v}\ (i\gamma^2\langle\overrightarrow{\boldsymbol{m}}_{x_0}|\overrightarrow{\boldsymbol{m}}_{v}\rangle) = \langle\overrightarrow{\boldsymbol{m}}_{x_0}|\ [\![(\boldsymbol{C}\boldsymbol{T}\boldsymbol{\Phi})^{-1}\ \overrightarrow{\boldsymbol{r}}\ ]\!]_0\ \rangle\,, \tag{48b}$$

$$\Big(\dot{\phi}_0 + \gamma x_0(\dot{\omega}v + \gamma^2\omega\dot{v}) + \dot{\Delta}\phi\ \sinh^2 c\Big)\ (-i\langle\overrightarrow{\boldsymbol{m}}_{\phi_0}|\overrightarrow{\boldsymbol{m}}_{\omega}\rangle) = \langle\overrightarrow{\boldsymbol{m}}_{\omega}|\ [\![(\boldsymbol{C}\boldsymbol{T}\boldsymbol{\Phi})^{-1}\ \overrightarrow{\boldsymbol{r}}\ ]\!]_0\ \rangle\,, \tag{48c}$$

$$\dot{x_0}\ (i\gamma\langle\overrightarrow{\boldsymbol{m}}_{x_0}|\overrightarrow{\boldsymbol{m}}_{v}\rangle) = \langle\overrightarrow{\boldsymbol{m}}_{v}|\ [\![(\boldsymbol{C}\boldsymbol{T}\boldsymbol{\Phi})^{-1}\ \overrightarrow{\boldsymbol{r}}\ ]\!]_0\ \rangle\,, \tag{48d}$$

$$\Big(\dot{c} + i\,\dot{\Delta}\phi\ \cosh c\,\sinh c\Big)\ \big(-i\,\langle\overrightarrow{\boldsymbol{m}}_{c+}|\overrightarrow{\boldsymbol{m}}_{c+}\rangle\big) = \langle\overrightarrow{\boldsymbol{m}}_{c+}|\ [\![(\boldsymbol{C}\boldsymbol{T}\boldsymbol{\Phi})^{-1}\ \overrightarrow{\boldsymbol{r}}\ e^{-i\widetilde{\varphi}}\ ]\!]_0\ \rangle\,, \tag{48e}$$

$$\Big(\dot{c} - i\,\dot{\Delta}\phi\ \cosh c\,\sinh c\Big)\ \big(-i\,\langle\overrightarrow{\boldsymbol{m}}_{c-}|\overrightarrow{\boldsymbol{m}}_{c-}\rangle\big) = \langle\overrightarrow{\boldsymbol{m}}_{c-}|\ [\![(\boldsymbol{C}\boldsymbol{T}\boldsymbol{\Phi})^{-1}\ \overrightarrow{\boldsymbol{r}}\,e^{i\widetilde{\varphi}}\ ]\!]_0\ \rangle\,. \tag{48f}$$

In writing (48c) and (48d), we have used (43). Recall that while x_0 affects the location of the soliton's center, it *is not* the soliton's center. The equation for the latter is given below in (57).

Equations (48) can be simplified slightly. To that end, we notice: (i) the $\overrightarrow{\text{cc}}$ in $\overrightarrow{\boldsymbol{m}}_{\phi_0,x_0,\omega,v}$; (ii) that the second 2×2 diagonal block of $\boldsymbol{\Gamma}^+$, which is the weight in (41), is opposite of the first such a block; and (iii) that $(\boldsymbol{C}\boldsymbol{T}\boldsymbol{\Phi})^{-1}$ does not change the 2×1 block *structure* of $\overrightarrow{\boldsymbol{r}}$, shown in (9). Then, denoting for any four-component vectors $\overrightarrow{\boldsymbol{u}}^{\,T} = [\overrightarrow{u}_1^{\,T},\ \overrightarrow{u}_2^{\,T}]$ and $\overrightarrow{\boldsymbol{v}}^{\,T} = [\overrightarrow{v}_1^{\,T},\ \overrightarrow{v}_2^{\,T}]$ a "reduced" inner product

$$\langle\overrightarrow{\boldsymbol{u}}\,|\,\overrightarrow{\boldsymbol{v}}\rangle_{\text{red}} \equiv \int_{-\infty}^{\infty} dz\ \overrightarrow{u}_1^{\,\dagger}\,\gamma(\sigma_0 + v\sigma_1)\ \overrightarrow{v}_1, \tag{49}$$

one rewrites (48a)–(48d) as:

$$\dot{\omega}\ (i\langle\overrightarrow{\boldsymbol{m}}_{\phi_0}|\overrightarrow{\boldsymbol{m}}_{\omega}\rangle) = 2\text{Re}\,\langle\overrightarrow{\boldsymbol{m}}_{\phi_0}|\ [\![(\boldsymbol{C}\boldsymbol{T}\boldsymbol{\Phi})^{-1}\ \overrightarrow{\boldsymbol{r}}]\!]_0\ \rangle_{\text{red}}, \tag{50a}$$

$$\dot{v}\ (i\gamma^2\langle\overrightarrow{\boldsymbol{m}}_{x_0}|\overrightarrow{\boldsymbol{m}}_{v}\rangle) = 2\text{Re}\,\langle\overrightarrow{\boldsymbol{m}}_{x_0}|\ [\![(\boldsymbol{C}\boldsymbol{T}\boldsymbol{\Phi})^{-1}\ \overrightarrow{\boldsymbol{r}}]\!]_0\ \rangle_{\text{red}}, \tag{50b}$$

$$\Big(\dot{\phi}_0 + \gamma x_0(\dot{\omega}v + \gamma^2\omega\dot{v}) + \dot{\Delta}\phi\ \sinh^2 c\Big)\ (-i\langle\overrightarrow{\boldsymbol{m}}_{\phi_0}|\overrightarrow{\boldsymbol{m}}_{\omega}\rangle) = 2\text{Re}\,\langle\overrightarrow{\boldsymbol{m}}_{\omega}|\ [\![(\boldsymbol{C}\boldsymbol{T}\boldsymbol{\Phi})^{-1}\ \overrightarrow{\boldsymbol{r}}]\!]_0\ \rangle_{\text{red}}, \tag{50c}$$

$$\dot{x_0}\ (i\gamma\langle\overrightarrow{\boldsymbol{m}}_{x_0}|\overrightarrow{\boldsymbol{m}}_{v}\rangle) = 2\text{Re}\,\langle\overrightarrow{\boldsymbol{m}}_{v}|\ [\![(\boldsymbol{C}\boldsymbol{T}\boldsymbol{\Phi})^{-1}\ \overrightarrow{\boldsymbol{r}}]\!]_0\ \rangle_{\text{red}}. \tag{50d}$$

Furthermore, using the relation (22e) between $\overrightarrow{\boldsymbol{m}}_{c-}$ and $\overrightarrow{\boldsymbol{m}}_{c+}$ and also the relation (45) between their inner products, one rewrites Eqs. (48e), (48f) as

$$\dot{c} = \text{Re}\,\left(\frac{i}{\langle\overrightarrow{\boldsymbol{m}}_{c+}|\overrightarrow{\boldsymbol{m}}_{c+}\rangle}\,\langle\overrightarrow{\boldsymbol{m}}_{c+}|\,[\![\,(\boldsymbol{C}\boldsymbol{T}\boldsymbol{\Phi})^{-1}\ \overrightarrow{\boldsymbol{r}}\ e^{-i\widetilde{\varphi}}\ ]\!]_0\,\rangle_{\text{red}}\right), \tag{50e}$$

$$\dot{\Delta}\phi = \frac{1}{\cosh c\,\sinh c}\,\text{Im}\,\left(\frac{i}{\langle\overrightarrow{\boldsymbol{m}}_{c+}|\overrightarrow{\boldsymbol{m}}_{c+}\rangle}\,\langle\overrightarrow{\boldsymbol{m}}_{c+}|\,[\![\,(\boldsymbol{C}\boldsymbol{T}\boldsymbol{\Phi})^{-1}\ \overrightarrow{\boldsymbol{r}}\ e^{-i\widetilde{\varphi}}\ ]\!]_0\,\rangle_{\text{red}}\right). \tag{50f}$$

Let us remark that in Eqs. (48) or (50), the presence of the $[\![\ldots]\!]_0$ symbol is essential. Indeed, these equations were derived by the method of multiple scales, whereby their r.h.s.'s are allowed to depend only on the 'slow' time, but not on the 'fast' one. We will discuss this further in Sec. 7 with illustrating examples.

5.1.1 On the solution of Eqs. (50e) and (50f)

While the numerical solution of all of the Eqs. (50) is addressed in Sec. 5.3, here we address an issue specific to Eqs. (50e) and (50f). Namely, if at any point during the evolution c vanishes, that will present a problem for the evaluation of the r.h.s. of (50f). Therefore, below we show how Eqs. (50e) and (50f) can be recast in a form that is free of that problem.

Using the definition of $\widetilde{\varphi}$ from (22e), multiplying Eq. (50e) by $\mathrm{sech}^2 c\, e^{i\Delta\phi}$ and Eq. (50f) by $i \tanh c\, e^{i\Delta\phi}$ and adding the results, one obtains:

$$\dot{p} \;=\; \frac{i\,(1-|p|^2)}{\langle \overrightarrow{\boldsymbol{m}}_{c+} | \overrightarrow{\boldsymbol{m}}_{c+} \rangle} \, \langle \overrightarrow{\boldsymbol{m}}_{c+} | \, [\![\, (\boldsymbol{C}\boldsymbol{T}\boldsymbol{\Phi})^{-1} \, \overrightarrow{\boldsymbol{r}} \, e^{2i\widetilde{\phi}} \,]\!]_0 \, \rangle_{\mathrm{red}} \, . \tag{51a}$$

where

$$p \equiv \tanh c \; e^{i\Delta\phi}, \qquad \widetilde{\phi} = -\int_0^t \frac{\omega(s)}{\gamma(s)} ds + (\phi_0 + \gamma\omega v x_0). \tag{51b}$$

The real part of the numerator in the fraction on the r.h.s. of (51a) is just $\mathrm{sech}^2 c$. The quantities a and b, which enter the expression for matrix $\boldsymbol{C}$ in (7b), are expressed through p as:

$$a = 1/\sqrt{1-|p|^2}, \qquad b = p/\sqrt{1-|p|^2}. \tag{52}$$

It is clear that the numerical integration of (51a) will not present a problem at or near $p = 0$, which corresponds to $c = 0$.

5.2 Single-frequency soliton

The case of a single-frequency background soliton corresponds to $c \to 0$. Then equations (48a)–(48d) or, equivalently, (50a)–(50d), remain the same except that in the third equation, one needs to set $\dot{\Delta\phi} \sinh^2 c \equiv 0$. Furthermore, instead of considering $\dot{c}$ and $\dot{\Delta\phi}$, it is more intuitive to consider the decomposition (35), since $\overrightarrow{\boldsymbol{m}}_{c\pm}$ are eigenvectors of the linearized GN operator $\mathbf{L}$. Substituting (35) into (46), where now the sum does *not* include $\alpha = c, \Delta\phi$, and taking inner products with $\overrightarrow{\boldsymbol{m}}_{c+}$ and $\overrightarrow{\boldsymbol{m}}_{c-}$, one finds, upon using (25c):

$$\frac{dk_+}{dt} = \frac{-i}{\langle \overrightarrow{\boldsymbol{m}}_{c+} | \overrightarrow{\boldsymbol{m}}_{c+} \rangle} \left\langle \overrightarrow{\boldsymbol{m}}_{c+} \right| (\boldsymbol{T}\boldsymbol{\Phi})^{-1} \, \overrightarrow{\boldsymbol{r}} \, \exp\left[-2i \int_0^t ds\, \omega(s)/\gamma(s) \right] \Big\rangle_{\mathrm{red}}, \tag{53a}$$

where we have also used that $\lim_{c\to 0} \boldsymbol{C} = \boldsymbol{\Sigma}_0^0$; similarly,

$$k_- = k_+^*. \tag{53b}$$

Let us note that these equations (with the exponent replaced by $-i\widetilde{\varphi}$) can also be obtained from (50e) and (50f) (or from (48e)). Indeed, comparing (5) where $|b| \ll 1$ with (35) and using the second line of (22e) in the former and (14) in the latter, one concludes that $b = -\epsilon k_+$. Finally, the desired statement follows from $b \approx c\,\exp[i\Delta\phi]$ for $c \ll 1$ (see (20)).

An important difference between the r.h.s. of (53a) and those of (50) (and (51a)) is that in the former, one does *not* need to perform the averaging $[\![\ldots]\!]_0$. This is because k_+ pertains to the perturbation of the soliton's shape $\boldsymbol{\delta}\overrightarrow{\boldsymbol{\psi}}(x,t)$, which is allowed to depend on the 'fast' time t (see (16)).

5.3 Solving Eqs. (50) numerically

For brevity, we will refer to solving Eqs. (50), even though instead of Eqs. (50e) and (50f) one will need to solve a complex-valued Eq. (51a).

As we announced earlier, the inner products on the r.h.s.'s of Eqs. (50) cannot, in most cases, be evaluated analytically. Therefore, here we describe how they can be solved numerically for a particular but wide class of perturbations in Eq. (1):

$$\vec{r} = \Big[f_1\big(\vec{\psi},\, \vec{\psi}^{\,*}\big)\, g_1(x)\ ,\, f_2\big(\vec{\psi},\, \vec{\psi}^{\,*}\big)\, g_2(x) \Big]^T , \tag{54}$$

where $f_{1,2}$ are arbitrary functions of their arguments and $g_{1,2}$ are either linear functions or sine or cosine. The perturbations considered in the next two sections have the above form. Other forms of perturbations, e.g., polynomial in x and/or involving time, can also be handled by the procedure described below.

Consider the case where

$$g_1(x) = g_2(x) = x. \tag{55}$$

(The case where g_1 and g_2 are different linear functions and/or contain a constant term, is conceptually the same.) Then, using the first relation in (13b), any of the equations in (50) (or Eq. (53)) can be written as

$$\dot{\alpha} = h_{\alpha,\,1}(\omega, v, \{c, \Delta\phi\}_e) + h_{\alpha,\,0}(\omega, v, \{c, \Delta\phi\}_e)\, x_{\rm cent}, \tag{56}$$

where: α denotes the corresponding soliton parameter; the notation $\{\cdots\}_e$ will be explained shortly;

$$x_{\rm cent} = \int_0^t v(s)\, ds + x_0 \qquad \Leftrightarrow \qquad d\, x_{\rm cent}/dt = v(\epsilon t) + \epsilon \dot{x}_0; \tag{57}$$

and $h_{\alpha,\,j}$, $j = 0, 1$ are obtained from the inner products in (50) by substituting $[f_1, f_2]^T (z/\gamma)^j$ for $\vec{r}$. An example of one of the $h_{\alpha,\,j}$, on which the other such terms can be mimicked, is:

$$h_{\omega,\,j} = \frac{2{\rm Re}\,\langle \overrightarrow{\boldsymbol{m}}_{\phi_0}|\ [\![(\boldsymbol{C}\boldsymbol{T}\boldsymbol{\Phi})^{-1}\, [f_1, f_2, \overrightarrow{\rm cc}]^T\ ]\!]_0\ (z/\gamma)^j\ \rangle_{\rm red}}{i \langle \overrightarrow{\boldsymbol{m}}_{\phi_0} | \overrightarrow{\boldsymbol{m}}_{\omega} \rangle}, \tag{58}$$

where $\overrightarrow{\mathrm{cc}}$ was defined after (22). (The sign in front of this $\overrightarrow{\mathrm{cc}}$ is unimportant as per (49).)

To solve a system of equations (56) and (57), which involves both 'fast' and 'slow' times, one relates the derivative w.r.t. the 'slow' time τ, denoted by the overdot, to that w.r.t. the 'fast' time: see (21a). Then, for the purpose of numerical integration over time, the two aforementioned equations are equivalently rewritten as

$$d\alpha/dt = \epsilon\ (h_{\alpha,\,1}(\omega, v, \{c, \Delta\phi\}_e) + h_{\alpha,\,0}(\omega, v, \{c, \Delta\phi\}_e)\, x_{\mathrm{cent}}\,)\,, \tag{59}$$

$$d\, x_{\mathrm{cent}}/dt = v(t) + \epsilon\, dx_0/dt. \tag{60}$$

Now, note that when computing $h_{\alpha,\,j}$, their dependence on ω will enter *implicitly* through the inner products, which, in most cases, can be found only via numerical integration over z. This is due to a complicated dependence of $\overrightarrow{\psi_{(0)}}$ and, consequently, of $\overrightarrow{\boldsymbol{m}}_\alpha$ on ω. On the other hand, the only place where these inner products may depend on $\{c, \Delta\phi\}$ can come from the $\boldsymbol{C}$ matrix, and hence is much simpler. The ensuing dependence of $h_{\alpha,\,j}$ on these parameters can therefore be treated explicitly; thus, the notation $\{\cdots\}_e$ in (56) stands for 'explicit dependence'. Furthermore, $h_{\alpha,\,j}$ can also be expressed as explicit functions of v. However, this dependence is more complicated than that on $\{c, \Delta\phi\}$ and, therefore, we have chosen to compute it numerically, as for ω.

The way that $h_{\alpha,\,j}$ are computed as functions of $\{\omega,\ v\}$ and then used in the integration of Eqs. (56) is as follows. One turns the range $[0 < \omega < 1] \times [-1 < v < 1]$ into a numerical grid $\{(\omega_{\ell_\omega},\, v_{\ell_v})\}$, where ℓ_ω, ℓ_v are the respective grid indices. Correspondingly, one finds values $h_{\alpha,\,j}[\ell_\omega, \ell_v](\{c, \Delta\phi\}_e) \equiv h_{\alpha,\,j}(\omega_{\ell_\omega}, v_{\ell_v},\, \{c, \Delta\phi\}_e)$ on this grid. Next, let the grid for numerical integration of Eqs. (56) *over time* consist of points t_n for a range of n values. Suppose that at some $t = t_n$ one finds $\omega(t_n) = \omega_n,\ v(t_n) = v_n$. Then one finds the corresponding values of $h_{\alpha,\,j}$ by interpolation. Below we give an example of using linear interpolation while suppressing the implied dependence on $c, \Delta\phi$, which does not require interpolation:

$$h_{\alpha,\,j}(\omega_n, v_n) = (1 - u_\omega - u_v) h_{\alpha,\,j}\,[\,\lfloor \ell_\omega \rfloor, \lfloor \ell_v \rfloor\,] + u_\omega\, h_j\,[\,\lceil \ell_\omega \rceil, \lfloor \ell_v \rfloor\,] + u_v\, h_{\alpha,\,j}\,[\,\lfloor \ell_\omega \rfloor, \lceil \ell_v \rceil\,]\,, \tag{61a}$$

where $\lfloor \ell_{\omega,v} \rfloor$, $\lceil \ell_{\omega,v} \rceil$ are the pairs of consecutive indices on the (ω, v)-grid such that $\omega_{\lfloor \ell_\omega \rfloor} \leq \omega_n \leq \omega_{\lceil \ell_\omega \rceil}$, and similarly for v; and

$$u_\omega = \frac{\omega_n - \omega_{\lfloor \ell_\omega \rfloor}}{\omega_{\lceil \ell_\omega \rceil} - \omega_{\lfloor \ell_\omega \rfloor}}, \qquad u_v = \frac{v_n - v_{\lfloor \ell_v \rfloor}}{v_{\lceil \ell_v \rceil} - v_{\lfloor \ell_v \rfloor}}. \tag{61b}$$

Recall that the values of x_{cent}, c, and $\Delta\phi$ do not need to be interpolated.

Finally, let us comment on how the case where g_j in (54) are sines or cosines differs from that considered above. As an example, consider

$$g_1(x) = g_2(x) = \sin(q_1 x + q_0) \tag{62}$$

for constant q_1, q_0. Again, using the first relation in (13b) and the trigonometric identity for the sine of two arguments, any of the Eqs. (50) can be put in the form similar to that of (59):

$$d\alpha/dt = \epsilon \left(h_{\alpha,\,\sin}(\omega, v, \{c, \Delta\phi\}_e)\, \cos(q_1 x_{\rm cent}) + h_{\alpha,\,\cos}(\omega, v, \{c, \Delta\phi\}_e)\, \sin(q_1 x_{\rm cent}) \right), \qquad (63)$$

where, e.g.,

$$h_{\omega,\,\sin} = \frac{2{\rm Re}\, \langle \overrightarrow{\boldsymbol{m}}_{\phi_0}|\, [\![\, (\boldsymbol{C}\boldsymbol{T}\boldsymbol{\Phi})^{-1}\, [f_1, f_2, \overrightarrow{\rm cc}]^T\,]\!]_0\, \sin(q_1 z/\gamma + q_0)\, \rangle_{\rm red}}{i \langle \overrightarrow{\boldsymbol{m}}_{\phi_0} | \overrightarrow{\boldsymbol{m}}_{\omega} \rangle}, \qquad (64)$$

and similarly for $h_{\omega,\,\cos}$. The rest of the numerical integration of Eqs. (63) is analogous to that of (56).

In the next two sections, we apply this algorithm to two illustrative examples.

As far as the numerical integration of Eqs. (59), (60) or (63), (60) over time, one can choose any finite-difference method. In the results reported below, we used the second-order accurate explicit modified Euler method. As both the step of time integration and the mesh sizes on the (ω, v)-grid, we used the value 0.01.

6 Soliton perturbed by a linear potential

Here we consider Eq. (1) with

$$\vec{r} = x\, \vec{\psi}. \qquad (65)$$

This is the case (54), (55) with $[f_1, f_2]^T = \vec{\psi}$. It was considered in [3, 4] and recently revisited in [8], with the initial condition being the single-frequency soliton (2). In all of these earlier studies, it was shown that the collective coordinates method very accurately predicted the location of the soliton's center, i.e., $x_{\rm cent}$ in (57), when compared with the results of numerically simulating Eq. (1). However, in Fig. 9 of Ref. [4] and Fig. 3 of Ref. [8], the profile of the soliton with $\omega = 0.3$ is clearly seen to be asymmetric, which has not been explained. Below we demonstrate that this asymmetry is due to the '2ω-modes' $\overrightarrow{\boldsymbol{m}}_{c\,\pm}$.

One can handle the effect of any perturbation by the bi-frequency soliton perturbation theory, i.e., by all of Eqs. (50), even when the initial condition is a single-frequency soliton. However, if the perturbation does not lead to considerable growth of the '2ω-mode' (i.e., the latter remains small compared to the background soliton), then one can use the technically simpler single-frequency soliton perturbation theory, i.e., Eqs. (50a)–(50d) with $c = 0$ and Eqs. (35) and (54). The extent to which the '2ω-mode' would grow can be estimated from either (50e) or (53a).

For example, substituting $\overrightarrow{\boldsymbol{r}} = x\, \boldsymbol{T}\boldsymbol{\Phi}\, \overrightarrow{\psi_{(0)}}$ into the latter equation, one sees that its r.h.s. oscillates (with an instantaneous frequency $2\omega/\gamma$) and hence its average, $[\![\ldots]\!]_0$, vanishes. Therefore, the initially small amplitude k_+ of the '2ω-mode' oscillates on the 'fast'

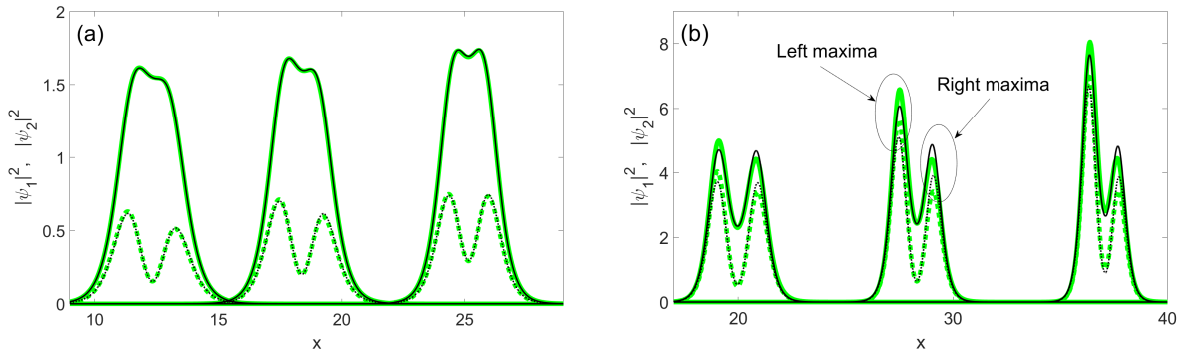

Figure 1: Intensities of the soliton's components for the perturbation (65) with $\epsilon = -0.01$ at $t = 40$, 50, and 60 (from left to right). Thick green lines: from numerical simulations of (1); thinner black lines: from the perturbation theory. Solid lines: $|\psi_1|^2$; dotted lines: $|\psi_2|^2$. Panels (a) and (b) are for $\omega = 0.3$ and 0.1, respectively. In (a), the numerical and theoretical results appear indistinguishable.

time scale but cannot grow considerably over the 'slow' scale, and hence we have used the single-frequency soliton perturbation equations in this case.

In Fig. 1 we display a sequence of snapshots of the soliton profile for $\omega = 0.3$ and 0.1. One set of data is obtained from the above perturbation theory with $\epsilon = -0.01$ (as in [8]), while the other is found by the simulation of Eq. (1) by the Method-of-Characteristics-based Split-Step method described in Sec. 2.4 of [15] with the parameters listed there. The agreement is seen to be good. In particular, as announced earlier, the perturbation theory explains the asymmetry of the soliton profile.

In regards to the perturbation theory, one can easily show from (58) and (50a) that $d\omega/dt = 0$. Therefore, the only evolution equations one needs are for dv/dt, dx_0/dt, and dk_+/dt. Solving them as described in Sec. 5.3 above, one then computes

$$\vec{\psi} = \boldsymbol{T}\boldsymbol{\Phi}\left(\overrightarrow{\psi_{(0)}} + \epsilon k_+ \, i\sigma_2 \overrightarrow{\psi_{(0)}} \, \exp\left[2i\int_0^t \omega(s)/\gamma(s)\, ds - 2i\omega v z\right]\right), \tag{66}$$

where we have used Eqs. (16), (22e), and (35), and $\overrightarrow{\psi_{(0)}}(z)$ is used with the instantaneous values of ω, v. From (66), one obtains the quantities displayed in Fig. 1.

In Fig. 2(a,b) we demonstrate that the perturbation theory also predicts the location of the soliton center, $x_{\rm cent}$, well, as did the collective coordinates method in the earlier studies [3, 4, 8]. Following them, in the simulations of Eq. (1), we compute this quantity as the center of mass:

$$x_{\rm cm} = \int x \left\|\vec{\psi}\right\|^2 dx \,/ \int \left\|\vec{\psi}\right\|^2 dx, \qquad \left\|\vec{\psi}\right\|^2 \equiv |\psi_1|^2 + |\psi_2|^2, \tag{67}$$

where the integration is performed over the computational window. In Fig. 2(c,d) we show the evolutions, computed by the perturbation theory, of velocity v and the amplitude of the

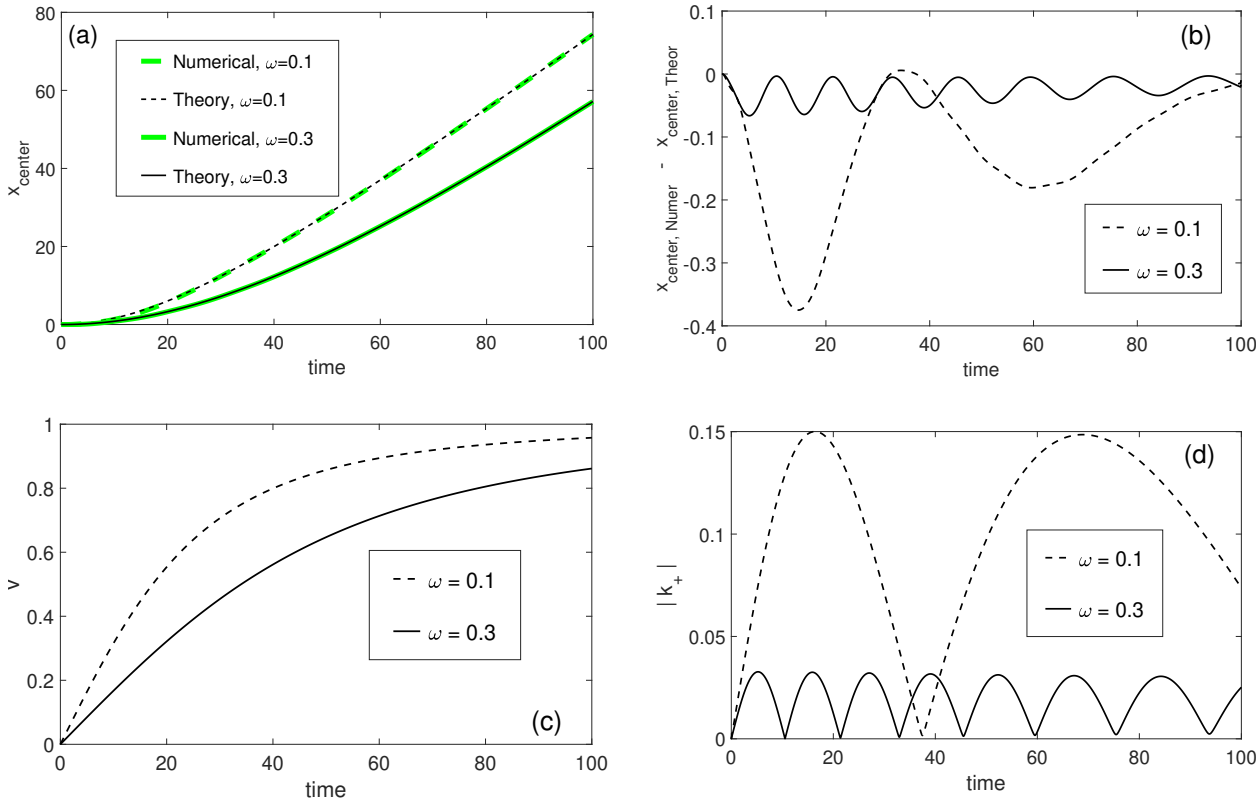

Figure 2: (a) Evolution of the soliton center, computed by numerically solving Eq. (1) (thicker green lines) and by the perturbation theory (thinner black lines) for the perturbation (65) with $\epsilon = -0.01$. The numerical and theoretical results appear indistinguishable. (b) The difference between these two sets of results. (c) Evolution of the soliton velocity. (d) Evolution of the amplitude of the '2ω-mode'.

'2ω-mode'. In particular, the latter figure corroborates both the earlier statement that the amplitude of this mode does not grow considerably and the observation from Fig. 1 that this amplitude — and hence the asymmetry of the soliton — is, overall, much greater for the soliton with a smaller ω.

Finally, in Fig. 3 we display the evolutions of the values of the left and right maxima of $\left\|\vec{\psi}\right\|^2$ and $|\psi_2|^2$, as labeled in Fig. 1. One set of data is computed by the perturbation theory and the other, by direct numerical simulations of Eq. (1). Unlike in the plots shown in Fig. 1, here we transform the soliton into the frame where it is stationary. Technically, this is achieved by *not* multiplying the r.h.s. of (66) by $\boldsymbol{T}$ and by multiplying the direct solution of (1) by $\boldsymbol{T}^{-1}$. This transformation is done so as not to obscure the oscillations of the soliton's left and right maxima by their growth, which occurs due to the increasing v. Again, the results of the perturbation theory and direct simulations agree with each other well. The reason that the agreement is better for $\omega = 0.3$ than for $\omega = 0.1$ is that in the latter case, the linearized GN operator $\mathbf{L}$ has an additional internal mode of the same parity (i.e., evennes/oddness of the two components) as the '2ω-mode'; see Figs. 6 and 7 in [12].

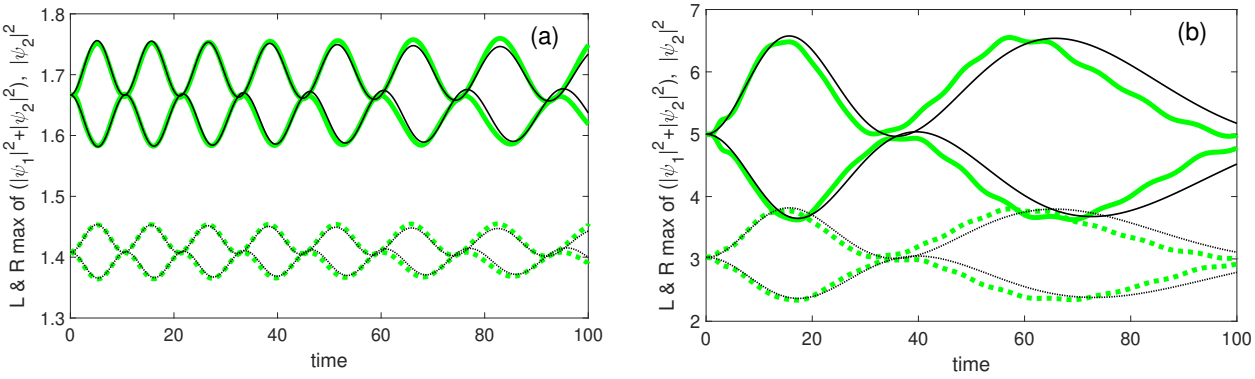

Figure 3: Evolutions of the left and right maxima, as marked in Fig. 1(b), for the perturbation (65) with $\epsilon = -0.01$. The field is first transformed into the non-moving frame by the transformation $\boldsymbol{T}^{-1}$, as explained in the text. Instead of plotting the maxima of $|\psi_1|^2$, we have found it to be more convenient to plot those of $|\psi_1|^2 + |\psi_2|^2$ because the latter maxima are more discernible than the former ones for $\omega = 0.3$ (see Fig. 1(a)). Also, we shifted the curves for the maxima of $|\psi_2|^2$ by 1 unit upwards in both panels in order to reduce the vertical range. Thick green lines: from numerical simulations of (1); thinner black lines: from the perturbation theory. Solid lines: $|\psi_1|^2 + |\psi_2|^2$; dotted lines: $|\psi_2|^2$. Panels (a) and (b) are for $\omega = 0.3$ and 0.1, respectively.

This internal mode, therefore, is also excited by the perturbation and contributes to the evolution of the soliton.

For larger values of ω, e.g., $\omega > 0.5$, one has $|\psi_{2(0)}| \ll |\psi_{1(0)}|$, and therefore the oscillations of the soliton profile become less conspicuous than for the smaller values of ω.

7 Soliton under the action of parametric perturbations

Here we will consider two forms of the perturbation in Eq. (1):

$$\vec{r} = \sigma_1 \vec{\psi}^*, \tag{68a}$$

and

$$\vec{r} = \sin(q_1 x + q_0)\, \sigma_3 \vec{\psi}^*. \tag{68b}$$

Perturbation (68b) was considered in [7], where for some values of parameters the numerical results were reported to significantly deviate from the prediction of the collective coordinates method, which did not take into account the possibility that $b \neq 0$ in (5). We consider perturbation (68a), which, to our knowledge, has not been considered before, as it presents a simpler case where a strong growth of parameter b is observed. Everywhere in this section we use the value $\epsilon = 0.02$, which was used in [7].

Unlike the case of the linear potential perturbation, where the factor on the right of the vertical bar on the r.h.s.'s of (48) (or (50)) simplifies to $x\,\vec{\psi}$, in the case of the parametric

perturbations (68), we first need to present an expression of that factor. This expression is, for the perturbation (68a):

$$(\boldsymbol{CT\Phi})^{-1}\,\overrightarrow{\boldsymbol{r}} = \boldsymbol{\Sigma}_3^0 \begin{pmatrix} (a^2+b^2)\sigma_1 e^{-2i\phi} & a(b+b^*)\sigma_0 \\ a(b+b^*)\sigma_0 & (a^2+(b^*)^2)\sigma_1 e^{2i\phi} \end{pmatrix} \overrightarrow{\boldsymbol{\psi}}^*_{(0)}\,; \tag{69a}$$

and for the perturbation (68b) it is:

$$(\boldsymbol{CT\Phi})^{-1}\,\overrightarrow{\boldsymbol{r}} = \left(\boldsymbol{\Gamma}^+\right)^{-1} \boldsymbol{\Sigma}_0^3 \begin{pmatrix} (a^2-b^2)\sigma_0 e^{-2i\phi} & a(b^*-b)\sigma_1 \\ a(b-b^*)\sigma_1 & (a^2-(b^*)^2)\sigma_0 e^{2i\phi} \end{pmatrix} \overrightarrow{\boldsymbol{\psi}}^*_{(0)}\,. \tag{69b}$$

Substitution of either of these equations into (51a) shows that the integrand on its r.h.s. will have a time-independent term since the time-dependent part of $(\widetilde{\phi}-\phi)$ vanishes:

$$\widetilde{\phi}-\phi = -\,v\omega z\,; \tag{70}$$

see Eqs. (13b) and (51b). This means that a significant change in parameter p (and hence in a and b) can be caused by those perturbations, and therefore one needs to use the general form of the perturbation theory developed above rather than its single-frequency soliton form. On the other hand, the off-diagonal blocks in both Eqs. (69) do not contribute to the evolution of p since they are zeroed out by the averaging $[\![\ldots]\!]_0$. Conversely, the diagonal (off-diagonal) blocks in (69) do not (do) contribute to the slow evolutions of ϕ_0, x_0, ω, v. The corresponding perturbation theory-based solution $\vec{\psi}$ is then found from the first relation in (13a), where the parameters of the three matrices and of the vector $\overrightarrow{\boldsymbol{\psi}}_{(0)}$ are found by solving Eqs. (50a)–(50d) and (51), (52) as explained in Sec. 5.3.

Solving the latter set of equations for the case (69a) with the initial conditions $v = \phi_0 = 0$ (and any ω and x_0) leads to p being purely imaginary and also being the only soliton parameter affected by this perturbation, (68a). The intensities of the soliton components found by the above perturbation theory-based approach are compared in Fig. 4 with those found by solving the GN equation (1) with perturbation (68a) by a 3rd-order accurate Method of Characteristics described in [16]. (A minor technical note: We used nonreflecting boundary conditions and, in addition, applied the same absorbing boundary function as in [15]. We did not use the 4th-order accurate method from [16], also based on a pseudo-Runge–Kutta scheme, because it requires a more complicated treatment of the boundary conditions, whereas we found the 3rd-order accurate method adequate for our purpose.) In Fig. 5(a) we plot the evolution of the soliton's charge, defined by Eq. (72) below, for the same perturbation and the same initial soliton as in Fig. 4, also computed by the above two methods. The agreement between the two sets of results is excellent. We verified that it remains such regardless of the soliton's ω. (Incidentally, the growth rate of the soliton charge — dictated by that of the parameter p — is also independent of ω, which can be shown using

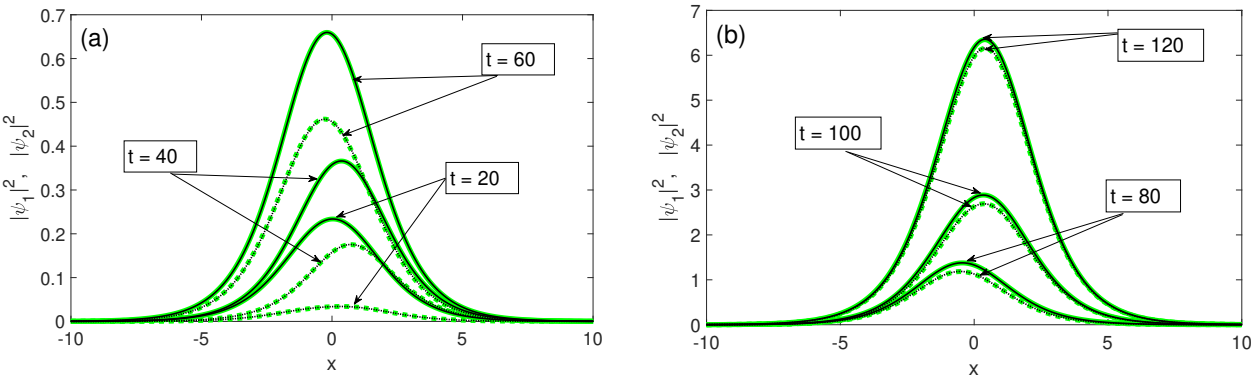

Figure 4: Intensities of the soliton's components for the perturbation (68a) with $\epsilon = 0.02$ and the initial soliton with $\omega = 0.9$, $v = \phi_0 = x_0 = b = 0$ at $t = 20, 40, 60$ (panel (a)) and $t = 80, 100, 120$ (panel (b)). Two panels instead of one are used to improve visibility, given their disparate vertical scales. Thick green lines: from numerical simulations of (1); thinner black lines: from the perturbation theory. Solid lines: $|\psi_1|^2$; dotted lines: $|\psi_2|^2$. All numerical results appear indistinguishable from the corresponding theoretical ones.

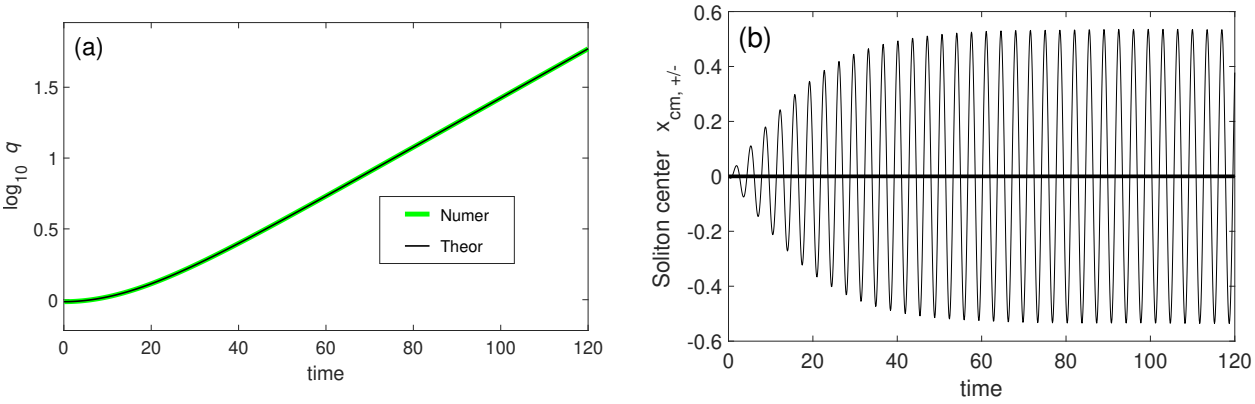

Figure 5: (a) Evolution of the soliton's charge, defined by (72), on the logarithmic scale. The numerical and perturbation-theoretical results appear indistinguishable: they differ by less than half percent. The growth of the charge, as well as that of parameters a and b, is seen to be asymptotically exponential, which can be (with some effort) deduced from (50e). (b) Evolutions of the soliton's centers of masses $x_{\mathrm{cm},+}$ (thin line) and $x_{\mathrm{cm},-}$ (thick line) computed from the numerical solution with Eq. (73). The evolution of the soliton's center computed with the perturbation theory (see (57)) coincides with that of $x_{\mathrm{cm},-}$.

(51a) and the aforementioned fact that the 2×2 diagonal blocks of (69a) are proportional to $\overrightarrow{\boldsymbol{m}}_{c+}$.)

For the foregoing discussion, we define two "densities":

$$\rho_{\pm} = |\psi_1|^2 \pm |\psi_2|^2, \tag{71}$$

where the former one is conventionally referred to as the density of the soliton charge $\mathscr{q}$:

$$\mathscr{q} = \int \rho_+ \, dx \tag{72}$$

(recall that the integration is implied over the computational window). We now use these densities to compute two versions of the soliton's center of mass, defined similarly to (67):

$$x_{\mathrm{cm},\pm} = \int x \, \rho_{\pm} \, dx \, / \int \rho_{\pm} \, dx. \tag{73}$$

Note that $x_{\mathrm{cm},+}$ coincides with x_{cm} defined in (67).

In the last sentence to the caption to Fig. 5(b) we point out that the location of the soliton's center found by the perturbation theory coincides with the numerically computed quantity $x_{\mathrm{cm},-}$ rather than with the more conventionally used $x_{\mathrm{cm},+}$ (see, e.g., [7]). Note that the oscillations of $x_{\mathrm{cm},+}$ observed in Fig. 5(b) are consistent with the soliton's center in Fig. 4 being visibly shifted away from 0. Yet, the perturbation theory-based graphs in Fig. 4 are plotted using the theoretical value $x_{\mathrm{cent}} = 0$. Thus, those shifts in Fig. 4 must have a different explanation, and, most likely, that explanation is that the $\psi_{2(0)}$-component of the soliton has an anti-symmetric shape. The oscillations of the location of the soliton's visible center then occur due to a combined effect of matrices $\boldsymbol{C}$ and $\boldsymbol{\Phi}$ ($\boldsymbol{T}$ for the case of perturbation (68a) is the identity matrix) on the components of $\overrightarrow{\boldsymbol{\psi}}_{(0)}$. On the other hand, the density ρ_- remains unchanged by the action of $\boldsymbol{C}$, $\boldsymbol{T}$, and $\boldsymbol{\Phi}$: see the sentence after Eqs. (75) in Appendix A.

Following the above discussion, we will use the quantity $x_{\mathrm{cm},-}$ when comparing the numerical results for the other perturbation, (69b), with those predicted by the perturbation theory. We now turn to considering that perturbation. For it, we found that p, and hence b, remain purely real during the evolution; therefore, the off-diagonal blocks in the matrix in (69b) vanish.

In Fig. 6 we compare intensities of the soliton components obtained by direct numerical simulations of (1) with term (68b) and by using the perturbation theory. We use the same parameters for which in [7], the collective coordinates method yielded results substantially different from those obtained by direct numerical simulations. In particular, the latter revealed formation of a large-amplitude soliton: see Fig. 7 there. Our Fig. 6 shows reasonable agreement between the numerical and perturbation theory-based results until the amplitude of the soliton increases significantly.

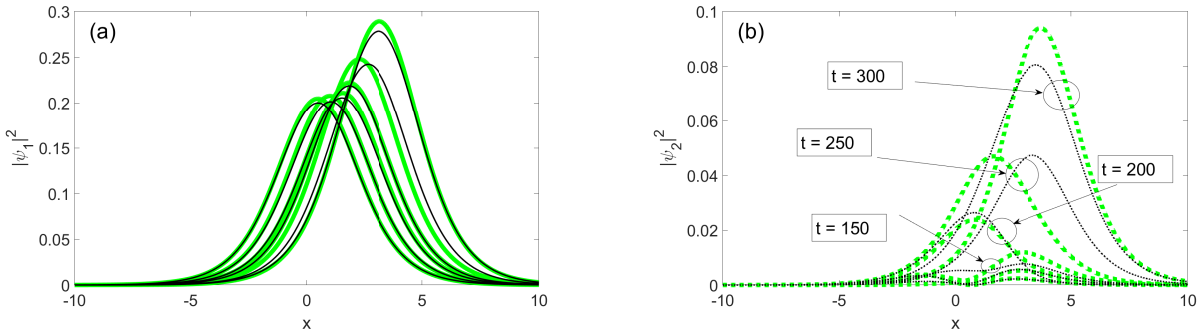

Figure 6: Intensities of the soliton's components for the perturbation (68b) with $\epsilon = 0.02$, $q_1 = 5\pi/32$, $q_0 = \pi/2$ and the initial soliton with $\omega = 0.9$, $v = 0.01$, and $\phi_0 = x_0 = b = 0$ at $t = 50 : 50 : 300$. Two panels instead of one are used to improve visibility, given their disparate vertical scales, but organized differently than in Figs. 1 or 4. Thick green lines: from numerical simulations of (1); thinner black lines: from the perturbation theory. In panel (a), where the agreement between the numerical and theoretical results is good, the respective pairs of graphs correspond to the time increasing from left to right. In panel (b), where the agreement is worse, we used grouping ellipses for the respective pairs except the first two, where the intensities are very small.

In Fig. 7 we show the evolution of the soliton charge, (72), and the soliton center $x_{\mathrm{cm},-}$, (73), for the same parameters as used for Fig. 6 but up to $t = 1500$, for which the simulations in [7] were run. One point that this figure conveys is that the perturbation theory predicts, albeit not very accurately, the formation of a large-amplitude soliton, which occurs due to the growth of parameters a and b. The other point is that the use of the averaging $[\![\ldots]\!]_0$ in the perturbation equations can be essential, as argued in Sec. 5.3. Indeed, the evolution predicted by the perturbation theory with this averaging (labeled as 'correct') agrees with the numerics considerably better than the evolution predicted by the perturbation equations without averaging (labeled as 'wrong') for this case of a substantial growth of the parameter b.

Given the last observation, it is the more surprising to report a case where the situation is reversed: the evolution predicted by the 'correct' perturbation theory agrees with the numerics much worse than the evolution predicted by the 'wrong' theory. We found this to occur for a *small* growth of the parameter b, as shown in Fig. 8. Moreover, we have also found that the performance of both versions of the perturbation theory, the 'correct' and 'wrong' ones, relative to the numerics depends on parameter ω of the soliton. Namely, we repeated the simulations with parameters as in Fig. 8 but for $\omega = 0.9$ and 0.95. Their results are summarized in Tables 1 and 2; they show that the evolution of the soliton's charge is predicted, at best, qualitatively, whereas that of the soliton's center is not reliably predicted by either version of the perturbation theory. We do not have an explanation for this situation

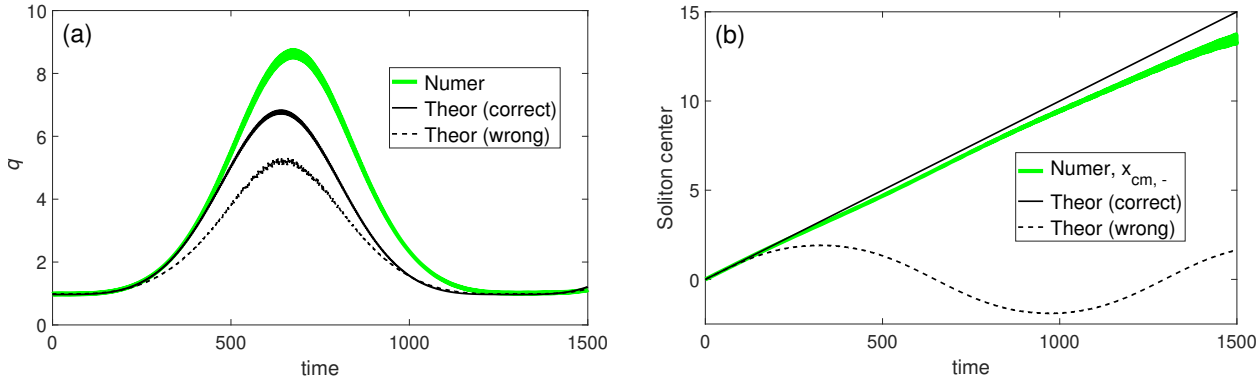

Figure 7: (a) Evolution of the soliton's charge, defined by (72), for the same parameters as in Fig. 6. Solid black line is for the perturbation theory which uses the averaging $[\![\ldots]\!]_0$, as explained in Sec. 5.3; dashed black line is for the perturbation theory which uses the expressions on the r.h.s.'s of (50) without the averaging. The "thickening" of all curves near the maxima is due to the small oscillations with frequency 2ω, which was also noted in [7]. (b) Evolutions of the soliton's center computed from numerical simulations and by the two versions of the perturbation theory, as explained in the caption to panel (a).

and leave it as a problem for future research.

	numerics	'correct' pert. theory	'wrong' pert. theory
$\max \mathcal{q}$	~ 3.5	~ 17	~ 1.8
time of $\max \mathcal{q}$	~ 800	~ 1100	~ 650
$x_{\rm cent}$ behavior	oscillates, period ~ 1600	increases linearly at rate v_0	oscillates, period ~ 1300

Table 1: Summary of the simulations with the same parameters as reported in Fig. 8 but for $\omega = 0.9$. 'Time of $\max \mathcal{q}$' is the time where the soliton charge first reaches it peak value.

8 Summary and Discussion

The main result of this work are the equations — Eqs. (50) (and also (51)), derived in Sec. 5 — for the evolution of the GN soliton parameters under the action of a small perturbation. We present both the general form of those equations, valid for the bi-frequency soliton (5) (or, equivalently, (13)), and their form specialized to the single-frequency soliton (2). The latter form can be obtained from the former in the limit of parameter b in (5) being very small. Accounting for this parameter is the key difference between our perturbation equations and those of the collective coordinates method, used in earlier studies. Another substantial

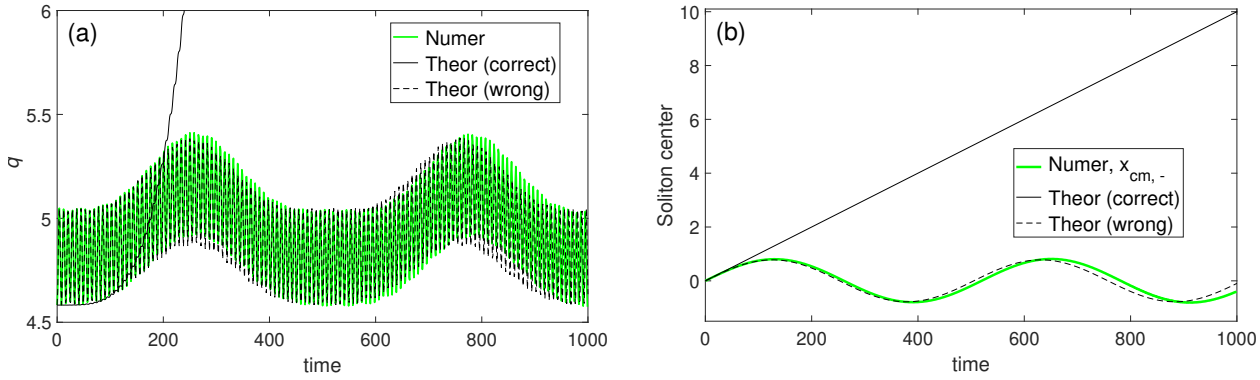

Figure 8: Same as in Fig. 7, but for $q_1 = 3\pi/32$ and the initial soliton having $\omega = 0.4$. To keep the meanings of the line styles/colors the same as in Fig. 7, we denoted those obtained with the perturbation theory that does not use the averaging as 'wrong', even though it has a much better agreement with direct numerics than the 'correct' theory. The charge predicted by the 'correct' theory peaks at a value around 1500 slightly past $t = 1000$, so that only a small part of that curve is shown in panel (a). The fast oscillations (with frequency 2ω) of the other two curves occur due to such oscillations in ω, whereas the envelope of the evolution occurs due to the (slow) changes of b.

	numerics	'correct' pert. theory	'wrong' pert. theory
$\max \mathcal{Q}$	~ 40	~ 6.5	~ 2.4
time of $\max \mathcal{Q}$	~ 1500	~ 1100	~ 900
$x_{\rm cent}$ behavior	increases linearly at rate $\sim 0.7 v_0$	increases linearly at rate v_0	oscillates, period ~ 1800

Table 2: Same as in Table 1, but for $\omega = 0.95$.

difference between the two approaches is that ours is based on the linearized operator of the GN equation.

The inclusion of parameter b allows one to explain both asymmetry of the soliton's profile (Sec. 6) and large growth (for certain types of the perturbation) of the soliton's components (Sec. 7). The latter can occur because both a and b in Eq. (5) can grow, as long as constraint (6) holds.

Since the expressions arising as driving terms in the equations of the perturbation theory can be, in most cases, evaluated only numerically (for a given value of the soliton frequency ω and velocity v), we also described — in Sec. 5.3 — how those equations can be solved by a finite-difference scheme in time. In Ref. [13] the reader can find an example of a code that we used to solve those equations for one of the perturbations considered in Sec. 7.

Let us now comment on two technical aspects of our calculations. First, in Sec. 2 we

explained (see the text near Eq. (12)) that the perturbation theory for a moving GN soliton needs to be developed in the coordinate frame related to the original (x,t) coordinates by the Galilean rather than Lorentzian boost, even though the GN equation is invariant under the latter. This led to the linearized GN operator in the moving frame, $\mathbf{L}$, being different from that in the original frame, $\mathbf{L}_0$. Their relation, Eq. (19), although simple, appears not to be obvious from the outset and needs to be derived by somewhat involved calculations.

Second, we wrote the $\overrightarrow{\psi_{(0)}}$-dependent part of $\mathbf{L}_0$ in a form, (17), that allowed us to establish relations (18), which were not mentioned in the earlier studies of the linearized GN operator [12]. Each of these relations was then used in Sec. 3 to derive relations satisfied by the neutral modes. In particular, relations (18a) were used to derive the 2ω-modes $\overrightarrow{\boldsymbol{m}}_{c\pm}$, which are "responsible" for infinitesimal changes of parameter b.

Finally, we point out two open problems. First, the fact that the derivation of relations (25c) satisfied by the neutral modes $\overrightarrow{\boldsymbol{m}}_{c\pm}$ is technically involved, can be attributed to the transformation (10) being unique to the GN equation (and some of its generalizations); indeed, its counterpart does not exist for the Nonlinear Schrödinger and (most, if not all) other nonlinear wave equations. However, relations (25b), satisfied by the neutral modes $\overrightarrow{\boldsymbol{m}}_{\omega,v}$, appear to be generic. For example, they hold for the Nonlinear Schrödinger equation; see Eqs. (29) in [9]. Then, is there a simpler (and possibly generic) way to derive them, thereby bypassing the very technical derivations found in Sec. 3 and Appendix C?

Second, at the end of Sec. 7 we gave two examples which presented conflicting views as to whether the averaging $[\![\ldots]\!]_0$ needs to be used in the perturbation equations (50) and (51a). On one hand, the averaging must be used by the very essence of the method of multiple scales, by which those equations were derived; see the end of Sec. 5.1. We also gave an example — see Fig. 7 — where the form of or equation with the averaging agreed considerably better with direct numerics than the form without the averaging. However, we also presented examples of the same perturbation with different parameters where either the perturbation theory *without* the averaging agrees with the direct numerics much better than that with the averaging (Fig. 8), or both versions of the perturbation theory do not give reliable predictions about the soliton evolution (Tables 1 and 2). The question of whether proceeding to the second order of the perturbation theory, where one considers the effect of the fast, non-adiabatic field $\boldsymbol{\delta}\overrightarrow{\boldsymbol{\psi}}$ (see (16)) on the soliton parameters, could improve the results of the 'correct' form of the perturbation theory (i.e., that with the averaging), is left for future research.

Conflict of interest

The author declares that he has no conflict of interest.

Data and Code availability

There are no data in this work other than those reported in the Figures. Those data are available from the author upon reasonable request. The code used to obtain some of the results reported in Sec. 7 is available in [13].

Acknowledgments

The author thanks A. Comech for pointing out the bi-frequency soliton equation (5) and N.R. Quintero for useful discussions about various types of perturbations of the GN equation.

References

[1] D.J. Gross and A. Neveu. Dynamical symmetry breaking in asymptotically free field theories. *Phys. Rev. D*, 10:3235–3253, 1974.

[2] M. Soler. Classical, stable, nonlinear spinor field with positive rest energy. *Phys. Rev. D*, 1:2766–2769, 1970.

[3] Y. Nogami, F.M. Toyama, and Z. Zhao. Nonlinear Dirac soliton in an external field. *J. Phys. A: Math. Gen.*, 28:1413–1424, 1995.

[4] F.G. Mertens, N.R. Quintero, F. Cooper, A. Khare, and A. Saxena. Nonlinear Dirac equation solitary waves in external fields. *Phys. Rev. E*, 86:046602, 2012.

[5] N.V. Alexeeva, I.V. Barashenkov, and A. Saxena. Spinor solitons and their PT-symmetric offspring. *Ann. Phys.*, 403:198–223, 2019.

[6] N.R. Quintero and B. Sánchez-Rey. Exact stationary solutions of the parametrically driven and damped nonlinear Dirac equation. *Chaos*, 29:093129, 2019.

[7] N.R. Quintero, B. Sánchez-Rey, F. Cooper, and F.G. Mertens. Length-scale competition in the parametrically driven nonlinear Dirac equation with a spatially periodic force. *J. Phys. A: Math. Theor.*, 52:285201, 2019.

[8] D. Mellado-Alcedo and N.R. Quintero. Stability of nonlinear Dirac solitons under the action of external potential. *Chaos*, 34:013140, 2024.

[9] D.J. Kaup. Perturbation theory for solitons in optical fibers. *Phys. Rev. A*, 42:5689–5694, 1990.

[10] N. Boussaïd and A. Comech. Spectral stability of bi-frequency solitary waves in Soler and Dirac-Klein-Gordon models. *Commun. Pure Appl. Anal.*, 17:1331–1347, 2018.

[11] A. Galindo. Remarkable invariance of classical Dirac Lagrangians. *Lett. Nuovo Cimento*, 20:210–212, 1977.

[12] G. Berkolaiko and A. Comech. On spectral stability of solitary waves of nonlinear Dirac equation in 1D. *Math. Model. Nat. Phenom.*, 7:13–31, 2012.

[13] T.I. Lakoba. https://github.com/tlakoba/Perturbation-theory-for-nonlinear-Dirac. *Github Repository*, 2026.

[14] D.J. Kaup and T.I. Lakoba. Variational method: How it can generate false instabilities. *J. Math. Phys.*, 37:3442–3462, 1996.

[15] T.I. Lakoba. Numerical study of solitary wave stability in cubic nonlinear Dirac equations in 1D. *Phys. Lett. A*, 382:300–308, 2018.

[16] T.I. Lakoba and J.S. Jewell. Higher-order explicit schemes based on the method of characteristics for hyperbolic equations with crossing straight-line characteristics. *Numer Meth. Partial Differ. Equ.*, 37:2742–2780, 2021.

A Derivation of Eqs. (17) and (18)

The first term on the r.h.s. of (17b) is obtained straightforwardly, given that $\gamma(\boldsymbol{\Sigma}_0^1 - v\boldsymbol{\Sigma}_0^0) \equiv \boldsymbol{\Gamma}^-\,\boldsymbol{\Sigma}_0^1$ commutes with all of $\boldsymbol{C}$, $\boldsymbol{T}$, and $\boldsymbol{\Phi}$. To obtain the second term, one takes into account the t- and z-derivatives of $\boldsymbol{\Phi}$. The third term is obtained by noticing that

$$\boldsymbol{C}^{-1}\boldsymbol{\Sigma}_3^3\boldsymbol{C} = \boldsymbol{\Sigma}_3^3, \qquad \boldsymbol{T}^{-1}\boldsymbol{\Sigma}_3^3\boldsymbol{T} = \boldsymbol{\Gamma}^-\boldsymbol{\Sigma}_3^3, \qquad \boldsymbol{\Phi}^{-1}\boldsymbol{\Gamma}^-\boldsymbol{\Sigma}_3^3\boldsymbol{\Phi} = \boldsymbol{\Gamma}^-\boldsymbol{\Sigma}_3^3, \tag{74}$$

and that

$$\boldsymbol{C}^T\boldsymbol{\Sigma}_1^3\boldsymbol{C} = \boldsymbol{\Sigma}_1^3, \qquad \boldsymbol{T}^T\boldsymbol{\Sigma}_1^3\boldsymbol{T} = \boldsymbol{\Sigma}_1^3, \qquad \boldsymbol{\Phi}^T\boldsymbol{\Sigma}_1^3\,\boldsymbol{\Phi} = \boldsymbol{\Sigma}_1^3. \tag{75}$$

Let us note that relations (75) follow from the invariance of the nonlinearity, i.e., the $\overrightarrow{\boldsymbol{\psi}}_{\rm ext}^T\boldsymbol{\Sigma}_1^3\overrightarrow{\boldsymbol{\psi}}_{\rm ext}$, in (13) with respect to the transformations of $\overrightarrow{\boldsymbol{\psi}}_{\rm ext}$ with matrices $\boldsymbol{C}$, $\boldsymbol{T}$, and $\boldsymbol{\Phi}$, respectively.

To derive the last term in (17b), we first note that for any vector $\vec{u}$ and a constant and symmetric matrix S, one has the identity

$$\delta\left(\vec{u}^T S\,\vec{u}\right) = 2\big(\vec{u}^T S\,\delta\vec{u}\big). \tag{76}$$

We now apply this equation to taking the variation of $\overrightarrow{\boldsymbol{\psi}}^T_{\rm ext}\boldsymbol{\Sigma}^3_1\overrightarrow{\boldsymbol{\psi}}_{\rm ext}$. Noting from (15) that this variation is to be multiplied by $\boldsymbol{\Sigma}^3_3\overrightarrow{\boldsymbol{\psi}}_{\rm ext}$ on the left, one rewrites

$$\boldsymbol{\Sigma}^3_3\overrightarrow{\boldsymbol{\psi}}_{(0)}\overrightarrow{\boldsymbol{\psi}}^T_{(0)}\boldsymbol{\Sigma}^3_1=\boldsymbol{\Sigma}^0_3\,\boldsymbol{\Sigma}^3_0\overrightarrow{\boldsymbol{\psi}}_{(0)}\overrightarrow{\boldsymbol{\psi}}^\dagger_{(0)}\boldsymbol{\Sigma}^3_0, \tag{77}$$

from which the last term in (17b) with P defined in (17d) follows.

Next, we outline the derivation of (18a) and then, of (18b) and (18c). Consider

$$\partial_c\Big((\boldsymbol{C}\overrightarrow{\boldsymbol{\psi}}_{(0)})^T\boldsymbol{\Sigma}^3_1(\boldsymbol{C}\overrightarrow{\boldsymbol{\psi}}_{(0)})\Big)=0, \tag{78}$$

where c is defined in (20). The equality above follows from the first relation in (75). On the other hand, using (76) and the relation

$$\boldsymbol{C}^T\boldsymbol{\Sigma}^3_1\,\partial_c\boldsymbol{C}=\boldsymbol{\Sigma}^3_1\boldsymbol{C}^{-1}\partial_c\boldsymbol{C}=\boldsymbol{\Sigma}^3_1\begin{pmatrix}\mathscr{o} & e^{i\Delta\phi}\sigma_1\\ e^{-i\Delta\phi}\sigma_1 & \mathscr{o}\end{pmatrix}, \tag{79}$$

one has:

$$-\frac{i}{2}\,\partial_c\Big((\boldsymbol{C}\overrightarrow{\boldsymbol{\psi}}_{(0)})^T\boldsymbol{\Sigma}^3_1\,(\boldsymbol{C}\overrightarrow{\boldsymbol{\psi}}_{(0)})\Big)=\overrightarrow{\boldsymbol{\psi}}^T_{(0)}\,\boldsymbol{\Sigma}^3_1\begin{pmatrix}e^{-i\Delta\phi}\sigma_2 & \mathscr{o}\\ \mathscr{o} & e^{i\Delta\phi}\sigma_2\end{pmatrix}\overrightarrow{\boldsymbol{\psi}}_{(0)}. \tag{80}$$

The last equation is obtained similarly to the identity $\boldsymbol{\Sigma}^1_1\overrightarrow{\boldsymbol{\psi}}_{(0)}=-i\,\boldsymbol{\Sigma}^2_0\overrightarrow{\boldsymbol{\psi}}_{(0)}$, which follows from (14). Since (80) holds for any $\Delta\phi$, then one must have:

$$\overrightarrow{\boldsymbol{\psi}}^T_{(0)}\boldsymbol{\Sigma}^3_1\begin{pmatrix}\sigma_2 & \mathscr{o}\\ \mathscr{o} & \mathscr{o}\end{pmatrix}\overrightarrow{\boldsymbol{\psi}}_{(0)}=\overrightarrow{\boldsymbol{\psi}}^T_{(0)}\boldsymbol{\Sigma}^3_1\begin{pmatrix}\mathscr{o} & \mathscr{o}\\ \mathscr{o} & \sigma_2\end{pmatrix}\overrightarrow{\boldsymbol{\psi}}_{(0)}=\vec{\boldsymbol{0}}, \tag{81}$$

which can also be confirmed by the identity

$$\overrightarrow{\psi_{(0)}}^T\sigma_2\overrightarrow{\psi_{(0)}}=0. \tag{82}$$

Relations (81) yield (18a).

By differentiating, respectively, $(\boldsymbol{T}\overrightarrow{\boldsymbol{\psi}}_{(0)})^T\boldsymbol{\Sigma}^3_1(\boldsymbol{T}\overrightarrow{\boldsymbol{\psi}}_{(0)})$ and $(\boldsymbol{\Phi}\overrightarrow{\boldsymbol{\psi}}_{(0)})^T\boldsymbol{\Sigma}^3_1(\boldsymbol{\Phi}\overrightarrow{\boldsymbol{\psi}}_{(0)})$ with respect to v and ϕ_0, one can see that the corresponding expressions yield relations

$$\overrightarrow{\boldsymbol{\psi}}^T_{(0)}\boldsymbol{\Sigma}^1_0\overrightarrow{\boldsymbol{\psi}}_{(0)}=0,\qquad\overrightarrow{\boldsymbol{\psi}}^T_{(0)}\boldsymbol{\Sigma}^0_3\overrightarrow{\boldsymbol{\psi}}_{(0)}=0, \tag{83}$$

which imply (18b) and (18c).

Finally, we note that using $\partial_{\Delta\phi}$ instead of ∂_c in (78) yields the already known relations (82) and (18c).

B Derivation of Eq. (23)

The integral on the r.h.s. of that equation can be thought of as the Riemann sum:

$$\int_0^t f(\omega(s))\,ds = \lim_{ds\to 0}\sum_{i=0}^{t/ds} f(\omega(t-i\,ds))\;ds. \tag{84}$$

Then

$$\partial_{\omega(t)}\int_0^t f(\omega(s))\,ds = \lim_{ds\to 0} df(\omega(t))/d\omega(t)\;ds = 0. \tag{85}$$

C Derivation of the second equation in (25b)

Differentiating (26) with respect to v, multiplying the result by $(\boldsymbol{CT\Phi})^{-1}$, and using (24) and relation $\boldsymbol{\Phi}^{-1}\partial_v\boldsymbol{\Phi} = i\omega\check{z}(1+v^2\gamma^2)\boldsymbol{\Sigma}_3^0$, one obtains a sum of five terms $\mathcal{T}_j$, $j=1,\ldots,5$, with:

$$\mathcal{T}_1 = \frac{\gamma^2}{2}\,\boldsymbol{\Sigma}_0^1\boldsymbol{\Gamma}^-\,(\mathbf{L}_0-\mathbf{P})\,\overrightarrow{\boldsymbol{\psi}}_{(0)}, \tag{86a}$$

$$\mathcal{T}_2 = \gamma^2\,i\omega\check{z}\,\boldsymbol{\Sigma}_3^0\boldsymbol{\Gamma}^-\,(\mathbf{L}_0-\mathbf{P})\,\overrightarrow{\boldsymbol{\psi}}_{(0)}, \tag{86b}$$

$$\mathcal{T}_3 = \partial_v\boldsymbol{\Gamma}^-\,(\mathbf{L}_0-\mathbf{P})\,\overrightarrow{\boldsymbol{\psi}}_{(0)}, \tag{86c}$$

$$\mathcal{T}_4 = \mathbf{L}\left(vz\,\partial_z\overrightarrow{\boldsymbol{\psi}}_{(0)}\right), \tag{86d}$$

$$\mathcal{T}_5 = \boldsymbol{\Gamma}^-\,i\boldsymbol{\Sigma}_0^1\partial_v(1/\gamma)\partial_x\overrightarrow{\boldsymbol{\psi}}_{(0)}. \tag{86e}$$

Terms $\mathcal{T}_{1,2}$ originated from, respectively, $\partial_v\boldsymbol{T}$ and $\partial_v\boldsymbol{\Phi}$. Term $\mathcal{T}_4$ originated from differentiating the expression in the main parentheses in (26) and using (24a). The origin of $\mathcal{T}_5$ is as follows. Note that $\partial_z = (1/\gamma)\partial_x$. Therefore,

$$\partial_v\big(\partial_z(\ldots)\big) = \partial_v(\gamma^{-1})\,\partial_x(\ldots) + \partial_z\big(\partial_v(\ldots)\big). \tag{87}$$

The last term above is already included in $\mathcal{T}_4$, and hence $\mathcal{T}_5$ accounts for the first term on the r.h.s. of (87).

To begin simplifying $\sum_{j=1}^5\mathcal{T}_j$, we note that $\mathcal{T}_3 = -2\mathcal{T}_1$ and write

$$\begin{aligned}\mathcal{T}_1+\mathcal{T}_3 &= -\frac{\gamma^2}{2}\boldsymbol{\Sigma}_0^1\boldsymbol{\Gamma}^-\left(\omega\boldsymbol{\Sigma}_3^0 + i\boldsymbol{\Sigma}_0^1\partial_z + \boldsymbol{\Sigma}_3^3 n_{(0)}\right)\overrightarrow{\boldsymbol{\psi}}_{(0)}\\ &= \gamma^2\boldsymbol{\Gamma}^-\left(-\omega\boldsymbol{\Sigma}_3^1 - i\boldsymbol{\Sigma}_0^0\partial_z\right) + \boldsymbol{\Gamma}^-\,(\mathbf{L}_0-\mathbf{P})\left(\frac{\gamma^2}{2}\boldsymbol{\Sigma}_0^1\overrightarrow{\boldsymbol{\psi}}_{(0)}\right),\end{aligned} \tag{88}$$

where $n_{(0)}$ was defined in (17c). The second equality above follows because $\boldsymbol{\Sigma}_0^1$ commutes with the first three matrices but anti-commutes with $\boldsymbol{\Sigma}_3^3$. Using the relation (18b), one has:

$$\text{the last term in (88)} = \mathbf{L}\left(\frac{\gamma^2}{2}\boldsymbol{\Sigma}_0^1\overrightarrow{\boldsymbol{\psi}}_{(0)}\right). \tag{89}$$

Next,

$$\text{the } \boldsymbol{\Sigma}^0_0\text{-term in (88)} + \mathcal{T}_5 = -i\gamma\partial_z \overrightarrow{\boldsymbol{\psi}}_{(0)}. \tag{90}$$

Now, to relate $\mathcal{T}_2$ to the first term in the expression of $\overrightarrow{\boldsymbol{m}}_v$ in (22d), one needs to "move" $i\omega\check{z}\boldsymbol{\Sigma}^0_3$ from the left to the right of the parentheses in (86b). Matrix $\boldsymbol{\Sigma}^0_3$ commutes with $\boldsymbol{\Gamma}^-$ and all three matrices in $(\mathbf{L}_0 - \mathbf{P})$, and therefore the only "extra" term that will occur in the above "moving" step would appear in light of the identity (31). Hence,

$$\mathcal{T}_2 = \mathbf{L}\left(\gamma^2\, i\omega\check{z}\boldsymbol{\Sigma}^0_3\overrightarrow{\boldsymbol{\psi}}_{(0)}\right) + \gamma^2\boldsymbol{\Sigma}^1_3\omega\overrightarrow{\boldsymbol{\psi}}_{(0)}. \tag{91}$$

The first term above is obtained using the relation (18c). In it, $\check{z}$ can be replaced with z (see (13b)) in view of the first equation in (25a). The last term in (91) cancels with the $\boldsymbol{\Sigma}^1_3$-term in (88). Combining this with all the previous relations in this Appendix, one obtains the desired result.

D Avoiding secular growth of the solution of (46)

Taking the inner product of (46) with $\overrightarrow{\boldsymbol{m}}_{\phi_0}$, using the fact that operator $\boldsymbol{\Gamma}^+\mathbf{L}$ is Hermitian (see (36) and (39)) and then using the first equation in (25a), one finds:

$$i\,\partial_t\,\langle\overrightarrow{\boldsymbol{m}}_{\phi_0}|\boldsymbol{\delta}\overrightarrow{\boldsymbol{\psi}}\rangle = \langle\overrightarrow{\boldsymbol{m}}_{\phi_0}|\,\text{r.h.s. of (46)}\rangle. \tag{92}$$

Using the notation (47), the condition to avoid the linear growth of $\langle\overrightarrow{\boldsymbol{m}}_{\phi_0}|\boldsymbol{\delta}\overrightarrow{\boldsymbol{\psi}}\rangle$ with the 'fast' time t is:

$$\langle\overrightarrow{\boldsymbol{m}}_{\phi_0}|\,[\![\,\text{r.h.s. of (46)}\,]\!]_0\,\rangle = 0, \tag{93}$$

from which (48a) follows. One can then set

$$[\![\,\langle\overrightarrow{\boldsymbol{m}}_{\phi_0}|\boldsymbol{\delta}\overrightarrow{\boldsymbol{\psi}}\rangle\,]\!]_0 = 0, \tag{94}$$

since if the l.h.s. of (94) is a nonzero constant instead, it can be absorbed into a shift of ω.

Taking the inner product of (46) with $\overrightarrow{\boldsymbol{m}}_\omega$, following the above lines and also using (94), one obtains the condition

$$\langle\overrightarrow{\boldsymbol{m}}_{\omega}|\,[\![\,\text{r.h.s. of (46)}\,]\!]_0\,\rangle = 0, \tag{95}$$

which entails (48c). The analogous conditions involving $\overrightarrow{\boldsymbol{m}}_{x_0,v}$ are obtained similarly.

The derivation of the conditions involving $\overrightarrow{\boldsymbol{m}}_{c\pm}$ requires a minor change to the above. Taking the inner product of (46) with, say, $\overrightarrow{\boldsymbol{m}}_{c+}$, following the above lines and also using (25c), one obtains an equation

$$i\,\partial_t\,\langle\overrightarrow{\boldsymbol{m}}_{c+}|\boldsymbol{\delta}\overrightarrow{\boldsymbol{\psi}}\rangle + \frac{2\omega}{\gamma}\,\langle\overrightarrow{\boldsymbol{m}}_{c+}|\boldsymbol{\delta}\overrightarrow{\boldsymbol{\psi}}\rangle = \langle\overrightarrow{\boldsymbol{m}}_{c+}|\,\text{r.h.s. of (46)}\rangle. \tag{96}$$

The condition to avoid a secular growth of $\langle \overrightarrow{\boldsymbol{m}}_{c+} | \overrightarrow{\boldsymbol{\delta\psi}} \rangle$ is

$$\langle \overrightarrow{\boldsymbol{m}}_{c+} | \, [\![\text{ r.h.s. of (46) }]\!]_{2\omega/\gamma} \, \rangle = 0, \tag{97}$$

and similarly for $\overrightarrow{\boldsymbol{m}}_{c-}$, with the frequency being now $-2\omega/\gamma$. Then equations (48e), (48f) follow upon noticing that the 'fast'-time dependence of $\exp[2i(\omega/\gamma)\, t]$ and $\exp[i\widetilde{\varphi}]$ is the same.