# Averaged dynamics of optical pulses described by a nonlinear Schrödinger equation with periodic coefficients

Andreas Wingen,[1] Karl H. Spatschek,[1] and Serguei B. Medvedev[1,2]
[1]*Institut für Theoretische Physik, Heinrich-Heine-Universität Düsseldorf, D-40225 Düsseldorf, Germany*
[2]*Institute of Computational Technologies, Siberian Branch, Russian Academy of Science, Novosibirsk 630090, Russia*
(Received 7 February 2003; published 23 October 2003)

A nonlinear Schrödinger equation with periodic coefficients, as it appears, e.g., in nonlinear optics, is considered. The high-frequency, variable part of the dispersion may be even much larger than the mean value. The ratio of the length of the dispersion map to the period of a solution is assumed as one small parameter. The second one corresponds to the integral over the variable part of the dispersion. For the averaged dynamics, we propose a procedure based on the Bogolyubov method. As a result, we obtain the asymptotic equation in the dominating order, as well as with the next corrections. The equation is valid for all combinations of the small parameters. The explicit forms of the coefficients are presented for a two-step dispersion map with an exponential loss function. The forms of the bright and black soliton solutions are discussed. The results are compared to those from other averaging methods, namely, the multiple-scale method and the method based on Lie transformations.

DOI: 10.1103/PhysRevE.68.046610 PACS number(s): 42.81.Dp, 42.65.Tg, 42.81.Uv, 42.79.Sz

## I. INTRODUCTION

Dispersion management is a novel technique in high-bit-rate optical data transmission. The dispersion-managed (DM) transmission systems use periodic alternations of fiber pieces with positive and negative dispersion coefficients, respectively. The idea of dispersion management is to minimize the path-averaged dispersion of a line, keeping high enough local dispersion. Pulse propagation in optical fibers is generically modeled by a nonlinear Schrödinger equation (NLSE) [1,2]. Dispersion management results in a NLSE model with periodic coefficients. Usually, one is interested in the long-distance dynamics of pulses when the averaged variations are slow compared to the variations of the dispersion. Then averaging methods should be applied.

The question, to which extent the *averaged* pulse propagation is robust, is of high practical relevance in dispersion-managed fiber optics communication systems. Although a pulse will locally change its form when propagating through one fiber piece (dispersion map), on an average (when propagating through many fiber pieces) it can be quite stable. One would like to know the (order of magnitude of the) distance up to which the averaged pulse shape is practically unchanged. That distance will depend on the parameters of the dispersion compensation. The idea to answer the question of robust *averaged* pulse propagation is based on transformations to equations for the averaged pulse (soliton). If the latter have a dominating integrable part, the estimate of the relevant distances can be explicitly performed and compared with experiments.

One averaging technique being applicable to this problem is the guiding-center theory based on the Lie transform [2–4]. At the beginning, only small variations of the dispersion were considered. Recently, the revised guiding-center theory was developed [5], but without a detailed analysis of different limits for the coefficients. Another approach was used in Refs. [6–8]. There, the multiple-scale method was applied to study the DM soliton power enhancement. Corrections to the bright soliton were obtained analytically. In those investigations the dispersion variation can be of the order of the mean value. Also, a normal form theory was applied to eliminate some nonimportant terms for the long-distance behavior. The elimination reflects the fact that the averaged equation is not of Hamiltonian form, and the nonlinear terms dependent on a phase. The final equations for the averaged dynamics contain fifth power nonlinearities. The assumption was made that the averaged dispersion is of zeroth order. The multiple-scale technique was applied in Ref. [9] to the Fourier transformed problem. The Fourier transformation helps via the Floquet-Lyapunov transformation to eliminate the large, variable part of the dispersion. Then, the so-called Gabitov-Turitsyn equation in spectral form [10] appears. Important in the present context is also Ref. [11]. There, the averaged equation was obtained using the Fourier expansion of the periodic coefficients. However, it was assumed that the variable part of the dispersion is small. Another approach to study existence and stability of DM solitons was made in Ref. [12]. For an averaged integral NLS equation, using averaging methods, the first correction to the Hamiltonian was derived to describe the long-range behavior via canonical equations.

Here, we will present another averaging method. We shall compare its result with those of the other two known ones, i.e., the direct or multiple-scale method and the Lie-transform technique, respectively. The present method is much more systematic than the direct method. The method proposed here is based on a Bogolyubov transformation. The latter is very elegant in practice, but requires the standard form for a correct application. The results of all the known averaging methods will be compared for the NLSE with fast varying coefficients.

The main aims of this paper are twofold. First, we want to find an averaged equation whose validity allows for a maximum range of parameter variation. Second, we intend to evaluate the corrections to the fundamental soliton solutions

 ©2003 The American Physical Society

in detail. The paper is organized as follows. In Sec. II, the model is presented. Section III defines the Bogolyubov transform. Explicit results for the averaged equation are obtained in Sec. IV. The corresponding soliton solutions are discussed in Sec. V. The paper is concluded by a short summary and outlook. In the Appendix we compare the present findings with those by other methods.

## II. MODEL

We consider a NLSE in dimensionless form

$$iu_z + d(z)u_{tt} + \varepsilon c(z)|u|^2 u = 0, \tag{1}$$

where the functions $d(z)$ and $c(z)$ are periodic, $d(z)=d(z+1)$ and $c(z)=c(z+1)$. The $z$ scale is made dimensionless by using a common period $z_0$ of the dimensional functions $d(z)$ and $c(z)$. The parameter $d(z)$ describes a fast variation of the local dispersion. The variation of the parameter $c(z)$ is, e.g., motivated by driving and damping in the original system. The parameter $\epsilon$ takes care of the small ratio of the periodicity length $z_0$ to the so-called nonlinear length [1]. Let us assume that in the original fiber line damping is compensated by lumped amplification in the form

$$i\frac{\partial w}{\partial z} + d(z)\frac{\partial^2 w}{\partial t^2} + \varepsilon|w|^2 w = -i\gamma w + i(\sqrt{G}-1)\sum_{n=0}^{N-1}\delta(z-nz_a)w, \tag{2}$$

where $z_a$ is the amplifier spacing, $\gamma$ the damping, and $G$ the gain coefficient; $z_a$ is assumed to be small compared to the dispersion length. By the transformation $w(z,t)=a(z)u(z,t)$ with

$$\frac{\partial a}{\partial z} + \left[\gamma - (\sqrt{G}-1)\sum_{n=0}^{N-1}\delta(z-nz_a)\right]a = 0, \tag{3}$$

the driven damped NLSE (2) is transformed into Eq. (1) with $c(z)=a(z)^2$.

For applications in nonlinear optics, the main aim is to find a localized, periodic solution of Eq. (1). We shall solve this problem using averaging procedures and the method of normal forms.

## III. BOGOLYUBOV TRANSFORMATION

### A. Standard form

A naive averaging of Eq. (1) fails since, in general, the function $d(z)$ is not small. To overcome this obstacle we apply the Floquet-Lyapunov transformation to eliminate the large variation $\tilde{d}$ of $d(z)$, i.e.,

$$u(t,z) = e^{iR(z)\Delta}A(t,z), \tag{4}$$

where

$$\Delta := \frac{\partial^2}{\partial t^2}, \quad R(z) = [d](z) + R_0. \tag{5}$$

We have introduced the notations

$$[d](z) := \int_0^z \tilde{d}(s)ds, \quad \tilde{d}(z) := d(z) - \langle d\rangle,$$

$$\langle d\rangle := \int_0^1 d(z)dz. \tag{6}$$

The angular brackets denote an averaging over one period. The tilde denotes the mean-free variable part of the corresponding variable. Finally, the square brackets denote the integral operator important in the theory of averaging.

The equation for $A$ takes the standard Bogolyubov form for a one-frequency system,

$$A_z = \varepsilon F(A,z) := \varepsilon(LA + P[|QA|^2 QA]), \tag{7}$$

with the compact notations for operators

$$L := i\langle d\rangle\Delta, \quad P := ic(z)\exp\{-iR(z)\Delta\},$$

$$Q := \exp\{iR(z)\Delta\}. \tag{8}$$

We assume that the mean of $d(z)$ is small (of order $\varepsilon$). Then, Eq. (7) is a one-frequency system, and it does not have any resonances. Averaging methods can be applied with arbitrary accuracy. If the mean of $d(z)$ is not small, resonances are possible, and one should apply the average theory for resonant systems. The latter theory is much more complicated (at least for partial differential equations).

We underline the importance to put the system into the standard form [13] before averaging. Otherwise, averaging may lead to wrong results. See, e.g., the finite-dimensional examples in Ref. [14] which illustrate this statement.

The strength of the dispersion management is characterized by $R(z)$,

$$R(z) = [d](z) + R_0, \tag{9}$$

where $R_0$ is, at this point, an arbitrary constant. The operators $P$ and $Q$ are bounded for any real function $R(z)$. Therefore, $F(A,z)$ is bounded too, and the form of the function $R(z)$ is not essential for performing the averaging. But the smallness of $R(z)$ will be helpful for a systematic simplification of the averaged equations.

### B. Averaging

Next, we apply the Bogolyubov transformation

$$A(z,t) = B(z,t) + \varepsilon V(B,z) + \varepsilon^2 W(B,z) + \varepsilon^3 Y(B,z) + \cdots. \tag{10}$$

The ansatz is made to eliminate the variable parts in successive orders of $\varepsilon$, and to transform Eq. (7) into an equation of the form

$$B_z = \varepsilon G(B), \tag{11}$$

with a $z$-independent functional $G(B)$.

Substituting Eq. (10) into Eq. (7) leads to

$$\vec{A}_z=[E+\varepsilon\hat{V}+\varepsilon^2\hat{W}+\varepsilon^3\hat{Y}]\vec{B}_z+\varepsilon\vec{V}_z+\varepsilon^2\vec{W}_z+\varepsilon^3\vec{Y}_z, \tag{12}$$

where the vector $\vec{A}$, the unit operator $E$, and the operators $\hat{V}$ are

$$\vec{A}=\begin{pmatrix} A_z \\ A_z^* \end{pmatrix}, \quad E=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{and} \quad \hat{V}=\begin{pmatrix} \frac{\partial V}{\partial B} & \frac{\partial V}{\partial B^*} \\ \frac{\partial V^*}{\partial B} & \frac{\partial V^*}{\partial B^*} \end{pmatrix}, \tag{13}$$

respectively. The notations for the other variables are similar.

In the following, we keep the terms up to third order in $\varepsilon$ in the equation for $B$,

$$\vec{B}_z=\varepsilon(\vec{F}_0-\vec{V}_z)+\varepsilon^2[\vec{F}_1-\vec{W}_z-\hat{V}(\vec{F}_0-\vec{V}_z)]+\varepsilon^3[\vec{F}_2-\vec{Y}_z-\hat{V}(\vec{F}_1-\vec{W}_z)-\hat{W}(\vec{F}_0-\vec{V}_z)+\hat{V}^2(\vec{F}_0-\vec{V}_z)], \tag{14}$$

where $F_n$ is the $n$th term of the expansion of $F$ with respect to the small parameter $\varepsilon$. Straightforward calculations lead to

$$F_0=LB+P(|QB|^2QB), \tag{15}$$

$$F_1=LV+P[2|QB|^2QV+(QB)^2(QV)^*], \tag{16}$$

$$F_2=LW+P[2|QV|^2QB+(QV)^2(QB)^*+2|QB|^2QW+(QB)^2(QW)^*]. \tag{17}$$

Using this, one finds the bounded mean-free coefficient $V(B,z)$ from

$$V_z=\tilde{F}_0=F_0-\langle F_0\rangle, \quad \langle V\rangle=0. \tag{18}$$

Then Eq. (11) takes the form

$$\vec{B}_z=\varepsilon\langle\vec{F}_0\rangle+\varepsilon^2(\vec{F}_1-\vec{W}_z-\hat{V}\langle\vec{F}_0\rangle)+\varepsilon^3[\vec{F}_2-\vec{Y}_z-\hat{V}(\vec{F}_1-\vec{W}_z)-\hat{W}\langle\vec{F}_0\rangle+\hat{V}^2\langle\vec{F}_0\rangle]. \tag{19}$$

The coefficient $W(B,z)$ can be determined from

$$\vec{W}_z=\tilde{\vec{F}}_1-\hat{V}\langle\vec{F}_0\rangle, \quad \langle W\rangle=0, \tag{20}$$

and the $\vec{B}$ equation is simplified to

$$\vec{B}_z=\varepsilon\langle\vec{F}_0\rangle+\varepsilon^2\langle\vec{F}_1\rangle+\varepsilon^3(\vec{F}_2-\vec{Y}_z-\hat{V}\langle\vec{F}_1\rangle-\hat{W}\langle\vec{F}_0\rangle). \tag{21}$$

To reduce the variable part in third order, we demand

$$\vec{Y}_z=\tilde{\vec{F}}_2-\hat{V}\langle\vec{F}_1\rangle-\hat{W}\langle\vec{F}_0\rangle, \quad \langle Y\rangle=0. \tag{22}$$

As a result, one obtains the averaged equation up to the third order

$$\vec{B}_z=\varepsilon G(B)=\varepsilon\langle\vec{F}_0\rangle+\varepsilon^2\langle\vec{F}_1\rangle+\varepsilon^3\langle\vec{F}_2\rangle. \tag{23}$$

The details for $\langle F_0\rangle$, $\langle F_1\rangle$, and $\langle F_2\rangle$ are

$$\langle F_0\rangle=LB+\langle P(|QB|^2QB)\rangle, \tag{24}$$

$$\langle F_1\rangle=\langle P[2|QB|^2QV+(QB)^2(QV)^*]\rangle, \tag{25}$$

$$\langle F_2\rangle=\langle P[2|QV|^2QB+(QV)^2(QB)^*+2|QB|^2QW+(QB)^2(QW)^*]\rangle. \tag{26}$$

We emphasize that this general equation is obtained for arbitrary $R(z)$. However, it is impossible to obtain tractable forms without additional assumptions; Eq. (23) cannot be solved analytically. Thus it is important to reduce Eq. (23) to a simpler form. When only the lowest order term is taken into account, Eq. (23) was solved numerically in Refs. [9,15].

## IV. EXPANSION FOR SMALL *R*

Since one cannot evaluate the operator $e^{iR(z)\Delta}$, in general, we consider approximations for this operator. Formally, we can introduce the parameter

$$\rho=\max|R(z)| \tag{27}$$

and renormalize $R(z)=\rho r(z)$. If the parameter $\rho$ is small, then we can expand the operators $P$ and $Q$ with respect to $\rho$, i.e.,

$$P:=ic(z)\exp\{-i\rho r\Delta\}=ic(z)(1-i\rho r\Delta-\rho^2r^2\Delta^2/2+\cdots), \tag{28}$$

$$Q:=\exp\{i\rho r\Delta\}=1+i\rho r\Delta-\rho^2r^2\Delta^2/2+\cdots. \tag{29}$$

### A. Case $R\equiv 0$

In the limit $R=0$, we obtain $P=ic(z)$ and $Q=E$. Substituting these expressions into $V$, $W$, $\langle F_0\rangle$, $\langle F_1\rangle$, and $\langle F_2\rangle$, we get in the first order

$$F_{00}=i\langle d\rangle\Delta B+ic(z)|B|^2B, \tag{30}$$

$$V=i\tilde{v}_0(z)|B|^2B, \quad v_0(z)=[c](z), \tag{31}$$

$$F_{10}=-\langle d\rangle\tilde{v}_0\Delta(|B|^2B)-c\tilde{v}_0|B|^4B. \tag{32}$$

Hereinafter, the second index of $F$ denotes the corresponding order of an expansion with respect to $R$ (or more formally with respect to $\rho$).

It is easy to show that

$$\langle c\tilde{v}_0\rangle=\langle\tilde{c}\tilde{v}_0\rangle=0 \quad \text{and} \quad \langle c[c]\rangle=\langle c\rangle\langle[c]\rangle \tag{33}$$

for arbitrary $c(z)$. Therefore $\langle F_1\rangle=0$ in this limit.

For the next coefficient we have the equation

$$W_z = -\langle d\rangle \tilde{v}_0 N_1(B) - \widetilde{c\tilde{v}_0}|B|^4 B, \tag{34}$$

with the solution

$$W = \tilde{w}_{03}(z) N_1(B) + \tilde{w}_{05}(z)|B|^4 B, \tag{35}$$

where

$$w_{03} = -\langle d\rangle [v_0], \quad w_{05} = -[\widetilde{c\tilde{v}_0}]. \tag{36}$$

The third term is

$$F_{20} = i\langle d\rangle \tilde{w}_{03}\Delta(N_1(B)) + i\langle d\rangle \tilde{w}_{05}\Delta(|B|^4 B) + ic\tilde{w}_{03}[2|B|^2\Delta(|B|^2 B) + B^2\Delta(|B|^2 B^*) - 3|B|^4\Delta B] + ic[\tilde{v}_0^2(z) + 3\tilde{w}_{05}(z)]|B|^6 B, \tag{37}$$

leading to

$$\langle F_{20}\rangle = i(\langle [c]^2\rangle - \langle [c]\rangle^2)\{\langle d\rangle [2|B|^2\Delta(|B|^2 B) + B^2\Delta(|B|^2 B^*) - 3|B|^4\Delta B] + \langle c\rangle |B|^6 B\}. \tag{38}$$

### B. Case $c(z) \equiv c_0$

Next, assuming $c(z) = c_0$ as constant, we present the expansions for small $\rho$. The first order gives

$$F_0 = i\langle d\rangle \Delta B + ic_0|B|^2 B + \rho c_0 r N_1(B) - \frac{i}{2}\rho^2 c_0 r^2 N_2(B), \tag{39}$$

where

$$N_1(B) := \Delta(|B|^2 B) + B^2\Delta B^* - 2|B|^2\Delta B, \tag{40}$$

$$N_2(B) := \Delta^2(|B|^2 B) + B^2\Delta^2 B^* + 2|B|^2\Delta^2 B - 4B|\Delta B|^2 + 2\Delta(B^2\Delta B^* - 2|B|^2\Delta B) + 2B^*(\Delta B)^2. \tag{41}$$

The averaged nonlinear term being proportional to $r$ can be eliminated when choosing the free constant $r_0$ such that $\langle r\rangle = 0$. The expansion of the first term is

$$\langle F_0\rangle = i\langle d\rangle \Delta B + ic_0|B|^2 B - \frac{i}{2}c_0\rho^2\langle r^2\rangle N_2(B). \tag{42}$$

Obviously, $\langle r^2\rangle = 0$ only if $r(z) \equiv 0$ Therefore this parameter is essential.

The first coefficient of the transformation is

$$V = \rho\tilde{v}_1 N_1(B) + i\rho^2\tilde{v}_2 N_2(B) + O(\rho^3), \tag{43}$$

where

$$v_1(z) = c_0[r], \tag{44}$$

$$v_2(z) = -\tfrac{1}{2}c_0[r^2]. \tag{45}$$

Therefore, we conclude that $\langle F_1\rangle \sim \rho^2$. However, this contribution can be neglected when compared with the last term on the right-hand side of Eq. (42). The expansion of $\langle F_2\rangle$ gives, in main order, $\rho^2$ contributions. These corrections can be neglected here. As a result, the main terms in Eq. (42) are valid up to the accuracy $\varepsilon\rho^2$.

### C. Case $c \not\equiv c_0$ and $R(z) \not\equiv 0$

Now, we present expansions for small $R(z)$ and arbitrary $c(z)$. The dominating terms are

$$F_0 = i\langle d\rangle \Delta B + ic(z)|B|^2 B + \rho c(z) r(z) N_1(B) - \frac{i}{2}\rho^2 c(z) r^2(z) N_2(B). \tag{46}$$

The averaged nonlinear term being proportional to $cr$ can be eliminated by choosing $r(z)$ with $\langle c(z) r(z)\rangle = 0$. The expansion of the first term leads to

$$\langle F_0\rangle = i\langle d\rangle \Delta B + i\langle c(z)\rangle |B|^2 B - \frac{i}{2}\rho^2\langle cr^2\rangle N_2(B). \tag{47}$$

It is obvious that $\langle c(z) r^2(z)\rangle = 0$ only if $c(z) = 0$ or $r(z) = 0$. Therefore this correction is essential.

The first coefficient of the transformation is now

$$V = i\tilde{v}_0|B|^2 B + \rho^2\tilde{v}_1 N_1(B) + i\rho^2\tilde{v}_2 N_2(B) + O(\rho^3), \tag{48}$$

where

$$v_1(z) = [cr], \quad v_2(z) = -\tfrac{1}{2}[cr^2]. \tag{49}$$

The expansion of $F_{11}$ gives

$$F_{11} = -\langle d\rangle \tilde{v}_0\Delta(|B|^2 B) + i\rho\langle d\rangle \tilde{v}_1\Delta(N_1(B)) + i\rho cr\tilde{v}_0[\Delta(|B|^4 B) - B^2\Delta(|B|^2 B^*) - 2|B|^2\Delta(|B|^2 B) + 2B^2|B|^2\Delta B^*] + i\rho c\tilde{v}_1[B^2\Delta(|B|^2 B^*) + 2|B|^2\Delta(|B|^2 B) - 3|B|^4\Delta B]. \tag{50}$$

For its averaged part we obtain

$$\langle F_{11}\rangle = i\rho\langle [c]cr\rangle [\Delta(|B|^4 B) - 2B^2\Delta(|B|^2 B^*) - 4|B|^2\Delta(|B|^2 B) + 2B^2|B|^2\Delta B^* + 3|B|^4\Delta B]. \tag{51}$$

As a result, we arrive at an equation that contains all the additional terms necessary for the relevant orderings in the parameters $\varepsilon$ and $\rho$, namely,

$$B_z = \varepsilon(i\langle d\rangle \Delta B + i\langle c\rangle |B|^2 B + \rho^2\langle F_{02}\rangle + \varepsilon\rho\langle F_{11}\rangle + \varepsilon^2\langle F_{20}\rangle), \tag{52}$$

where

$$\langle F_{02}\rangle = -\frac{i}{2}\langle c(z) r^2(z)\rangle N_2(B). \tag{53}$$

The term $\langle F_{02}\rangle$ is an essential term since it gives the first correction to the NLSE if $1 \gg \rho \gg \varepsilon$. The term $\langle F_{11}\rangle$ appears together with the other two terms on the right-hand side of Eq. (52), if $\rho\sim\varepsilon$. Finally, the term $\langle F_{20}\rangle$ has to be taken into account for $\rho\ll\varepsilon$ and when the coefficient $\langle[c]^2\rangle-\langle[c]\rangle^2$ is not small.

## V. SOLITON SOLUTIONS

Equation (52) has the so-called fundamental soliton solutions in the main order. The small perturbations induce corrections to the fundamental soliton solution. First, we evaluate the form of the coefficients of Eq. (52) for a given dispersion map. Then we will consider the perturbed solutions in the case of bright solitons ($\langle d\rangle>0$) and black solitons ($\langle d\rangle<0$), respectively.

### A. Two-step dispersion map

To be specific, we consider a symmetric two-step dispersion map $d(z)$

$$d(z)=\begin{cases}\varepsilon\langle d\rangle+D, & 0\leqslant z<1/4\\ \varepsilon\langle d\rangle-D, & 1/4\leqslant z<3/4\\ \varepsilon\langle d\rangle+D, & 3/4\leqslant z<1\end{cases} \tag{54}$$

and the exponential loss function $c(z)$ on the unit length $c(z)=c_0\exp(-2\gamma z)$. Here $\varepsilon\langle d\rangle$ is a normalized mean of $d(z)$, $D\geqslant 0$ is the amplitude of the variable part of $d(z)$, and $\gamma$ is the loss coefficient. The parameter $\rho$ is equal to $\rho=D/4$.

We denote the explicitly $z$-dependent factors appearing in $\langle F_{20}\rangle$, $\langle F_{11}\rangle$, and $\langle F_{02}\rangle$ (as functions of $\gamma$) by

$$f_{20}(\gamma)=\langle c\rangle(\langle[c]^2\rangle-\langle[c]\rangle^2),\quad f_{11}(\gamma)=\langle[c]cr\rangle,$$

$$\text{and}\quad f_{02}(\gamma)=\langle cr^2\rangle/2, \tag{55}$$

respectively. Then, we compare them with the coefficient $\langle c\rangle$ of the main order nonlinear term. These ratios have the following asymptotics:

$$\frac{f_{20}}{\langle c\rangle}=c_0^2\frac{\gamma^2}{180},\quad \frac{f_{11}}{\langle c\rangle}=c_0\frac{\gamma^2}{384},$$

$$\frac{f_{02}}{\langle c\rangle}=\left(\frac{1}{96}-\frac{13\gamma^2}{5760}\right)\quad\text{as}\quad\gamma\to 0, \tag{56}$$

$$\frac{f_{20}}{\langle c\rangle}=c_0^2\frac{1}{48\gamma^2},\quad \frac{f_{11}}{\langle c\rangle}=c_0\frac{1}{16\gamma^2},\quad \frac{f_{02}}{\langle c\rangle}=\frac{1}{8\gamma^2}\quad\text{as}\quad\gamma\to\infty. \tag{57}$$

We recognize that for small losses $\gamma$ the coefficient $f_{02}$ is relevant. The other coefficients are close to zero. Beyond some critical value $\gamma^*$ all the coefficients have the same order, and all have to be taken into account. For very large $\gamma$, all coefficients decrease as $\gamma^{-2}$, and can be neglected.

In the case of a lossless system ($\gamma=0$), the coefficients $\langle R^n\rangle$ take the forms

$$\langle R^{2k}\rangle=\frac{D^{2k}}{(2k+1)4^{2k}},\quad \langle R^{2k-1}\rangle=0\quad\text{for}\quad k=1,2,3,\ldots. \tag{58}$$

If the first term $\langle R^2\rangle$ has the order of $\varepsilon$, then we get for the amplitude $D=\sqrt{48\varepsilon}$. The next term is $\langle R^4\rangle=\frac{9}{5}\varepsilon^2$.

The two-step dispersion map used here is simple but important, since most limits of short-scale or long-scale dispersion maps have the present asymptotic form of the two-step dispersion map [16].

### B. Evaluation of soliton solutions

Now we are ready to solve the averaged equation (52). Because $\rho=\sqrt{\varepsilon}\gg\varepsilon$, we include only the dominating term, identify $v\equiv B$, and write the simplified version of Eq. (52) in the form

$$iv_z+\langle d\rangle v_{tt}+\langle c\rangle|v|^2v=\tfrac{1}{2}\langle cR^2\rangle N_2(v), \tag{59}$$

with $\langle d\rangle\sim O(\varepsilon)$, $\langle c\rangle\sim O(\varepsilon)$, and $R\sim O(\rho)$. For bright solitons we assume $\langle d\rangle>0$. To find a suitable ansatz for the perturbed soliton solution, we have to know the solution of the unperturbed equation. Without the perturbation on the right-hand side, Eq. (59) is the integrable NLSE with constant coefficients. Such an equation has, among other solutions, the bright one-soliton solution

$$v(z,t)=\eta\sqrt{\frac{2\langle d\rangle}{\langle c\rangle}}\operatorname{sech}(\eta t)e^{i\eta^2\langle d\rangle z}=G(t)e^{i\eta^2\langle d\rangle z}. \tag{60}$$

A proper ansatz for the solution of Eq. (59) is by perturbation theory,

$$v(z,t)=[G(t)+F(t)+H(t)]e^{i\eta^2\langle d\rangle z}, \tag{61}$$

with $F(t)\sim O(\varepsilon)$, $H(t)\sim O(\varepsilon^2)$, and $F$ as well as $H$ being real. Inserting this ansatz into Eq. (59), the zeroth order is trivially satisfied. In the first order of $\varepsilon$ we find

$$F_{tt}+\eta^2[6\operatorname{sech}^2(\eta t)-1]F=\frac{1}{2}\frac{\langle cR^2\rangle}{\langle d\rangle}N_2(G). \tag{62}$$

Next, we evaluate $N_2(G)$ using $G=k\operatorname{sech}(\eta t)$. A short calculation leads to the inhomogeneous ordinary differential equation (ODE)

$$F_{tt}+\eta^2[6\operatorname{sech}^2(\eta t)-1]F=\lambda\eta^2[76\operatorname{sech}^7(\eta t)-84\operatorname{sech}^5(\eta t)+16\operatorname{sech}^3(\eta t)] \tag{63}$$

for $F(t)$, with

$$\lambda=2\frac{\langle cR^2\rangle}{\langle d\rangle}\eta^2k^3=4\frac{\langle cR^2\rangle}{\langle c\rangle}\eta^5\sqrt{\frac{2\langle d\rangle}{\langle c\rangle}}. \tag{64}$$

To solve this ODE we introduce

$$y=\operatorname{sech}(\eta t) \tag{65}$$

to eliminate the trigonometric functions. For the time derivative we obtain the operator

$$\frac{\partial^2}{\partial t^2}=\eta^2\left[(y-2y^3)\frac{\partial}{\partial y}+y^2(1-y^2)\frac{\partial^2}{\partial y^2}\right], \tag{66}$$

and the ODE reads

$$(y-2y^3)F_y+y^2(1-y^2)F_{yy}+(6y^2-1)F$$
$$=\lambda(76y^7-84y^5+16y^3). \tag{67}$$

Using a power series expansion $F(y)=\lambda\Sigma_{n=0}^{\infty}c_n y^n$ we obtain the solution of Eq. (67) in the form $F(y)=\lambda(\frac{4}{3}y+\frac{4}{3}y^3-\frac{19}{6}y^5)$. After reinserting transformation (65) into $F(y)$ we find the solution of Eq. (63) for $F(t)$,

$$F(t)=\lambda\left(\frac{4}{3}\operatorname{sech}(\eta t)+\frac{4}{3}\operatorname{sech}^3(\eta t)-\frac{19}{6}\operatorname{sech}^5(\eta t)\right)$$
$$=\frac{16}{3}\frac{\langle cR^2\rangle}{\langle c\rangle}\eta^4 G(t)\left(1+\operatorname{sech}^2(\eta t)-\frac{19}{8}\operatorname{sech}^4(\eta t)\right). \tag{68}$$

The second-order (in $\varepsilon$) equation for $H$ can be solved in a similar manner. Starting with

$$H_{tt}+\eta^2[6\operatorname{sech}^2(\eta t)-1]H$$
$$=\frac{1}{2}\frac{\langle cR^2\rangle}{\langle d\rangle}N_2(G+F)-3\frac{\langle c\rangle}{\langle d\rangle}GF^2, \tag{69}$$

or after proper expansion of the right-hand side,

$$H_{tt}+\eta^2[6\operatorname{sech}^2(\eta t)-1]H$$
$$=\eta^2\chi\left[\frac{66\,557}{8}\operatorname{sech}^{11}(\eta t)+\frac{25\,725}{2}\operatorname{sech}^9(\eta t)\right.$$
$$\left.-4916\operatorname{sech}^7(\eta t)+116\operatorname{sech}^5(\eta t)+40\operatorname{sech}^3(\eta t)\right], \tag{70}$$

with

$$\chi=\frac{64}{3}\eta^9\frac{\langle cR^2\rangle^2}{\langle c\rangle^2}\sqrt{\frac{2\langle d\rangle}{\langle c\rangle}}, \tag{71}$$

we find

$$H(t)=\chi\left[\frac{66\,557}{672}\operatorname{sech}^9(\eta t)-\frac{10\,372}{105}\operatorname{sech}^7(\eta t)\right.$$
$$\left.+\frac{509}{70}\operatorname{sech}^5(\eta t)+\frac{1024}{105}\operatorname{sech}^3(\eta t)-\frac{998}{105}\operatorname{sech}(\eta t)\right]. \tag{72}$$

Inserting all these results into Eq. (61), we obtain the solution of the averaged equation (59) of the form

$$v(z,t)=\eta\sqrt{\frac{2\langle d\rangle}{\langle c\rangle}}\operatorname{sech}(\eta t)e^{i\eta^2\langle d\rangle z}\left\{1+\frac{16}{3}\frac{\langle cR^2\rangle}{\langle c\rangle}\right.$$
$$\times\eta^4\left[1+\operatorname{sech}^2(\eta t)-\frac{19}{8}\operatorname{sech}^4(\eta t)\right]$$
$$+\frac{64}{3}\frac{\langle cR^2\rangle^2}{\langle c\rangle^2}\eta^8\left[-\frac{998}{105}+\frac{1024}{105}\operatorname{sech}^2(\eta t)\right.$$
$$+\frac{509}{70}\operatorname{sech}^4(\eta t)-\frac{10\,372}{105}\operatorname{sech}^6(\eta t)$$
$$\left.\left.+\frac{66\,557}{672}\operatorname{sech}^8(\eta t)\right]\right\}. \tag{73}$$

Note that only the phase is $z$ dependent. Otherwise, we have $z$-independent factors which indicate that the (averaged) pulse amplitudes should be constant during propagation. Let us now compare this expression with the solution of the original equation (1). For that we take into account all corrections up to order $\rho^5$. Note that the next correction to $v$ is of order $\rho^6$. We choose the special case $c(z)\equiv\langle c\rangle\equiv\varepsilon$ for the reason of demonstration. We obtain for $u$,

$$u(z,t)=v(z,t)+iR\Delta v(z,t)-\tfrac{1}{2}R^2\Delta^2[v_0(z,t)+v_1(z,t)]$$
$$-\tfrac{1}{6}iR^3\Delta^3(v_0+v_1)+\tfrac{1}{24}R^4\Delta^4 v_0+\tfrac{1}{120}iR^5\Delta^5 v_0$$
$$+\varepsilon[R]N_1(v_0), \tag{74}$$

where $v_0(z,t)$ is the bright fundamental soliton and $v_1(z,t)$ is the first correction.

Let us now calculate the explicit forms of the averaged soliton $v$ and compare with the actual solution $u$ for different values of $\varepsilon$ and $\rho=\sqrt{\varepsilon}$. For simplicity, we assume $\eta=1$. Different values of $\varepsilon$ correspond to different strengths of the dispersion management. In the following, three cases will be presented for $\varepsilon=0.01$, $\varepsilon=0.04$, and $\varepsilon=0.1$. The averaged soliton $v$ is given by Eq. (73). The actual soliton $u$ can be calculated by inserting the averaged soliton $v$ into transformation (74) and using all terms up to the relevant order, as shown in Eq. (74).

Figure 1 shows the squares of the absolute amplitudes of $v$ (broken line) and $u$ (solid line) for $\varepsilon=0.01$. The amplitudes of $u$ and $v$ are nearly identical. Both solutions have the typical sech form. The expression $v$ is the so-called guiding-center soliton, while $u$ is called the dressed soliton. The dressed soliton shows small variations over the period length of the dispersion map during propagation. Breathing is only a small perturbation and the main form of the dressed soliton is dominated by the guiding-center soliton. This behavior is the typical behavior for weak dispersion management.

Figure 2 shows the squares of the absolute amplitudes of $v$ (broken line) and $u$ (solid line) for a stronger dispersion management ($\varepsilon=0.04$) compared to the previous case. Now the averaged soliton and the soliton solution $u$ are no longer very close to each other. While the averaged soliton $v$ still shows similarities to a sech pulse, the solution $u$ has a dip at

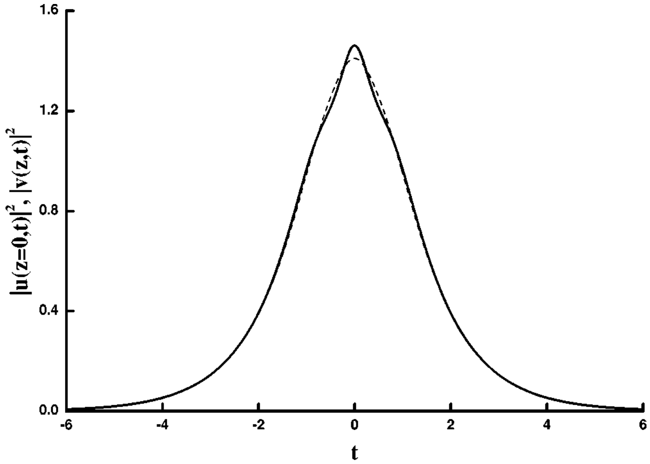


FIG. 1. Squared amplitudes $|u(z,t)|^2$ of the actual soliton at $z=0$ (solid line) and of the averaged soliton $|v(z,t)|^2$ (broken line) for $\varepsilon=0.01$ and $\eta=1$.

each side, and even indicates a second dip. This change in the behavior characterizes the beginning of strong dispersion management effects.

For even stronger dispersion management also the averaged soliton $v$ loses its sech form. Figure 3 shows the squares of the absolute amplitudes of $v$ (broken line) and $u$ (solid line) on a logarithmic scale for $\varepsilon=0.1$. The functions $u$ and $v$ show the characteristic behaviors of strong dispersion management. In the past, the so-called dispersion-managed solitons (DM solitons) were only investigated numerically. An analytic formula was unknown. The typical form of a DM soliton was presented in Ref. [17]. There we recognize regular dips at each side of the pulse. Here we achieved three dips at each side by using a perturbation theory.

We conclude this section by some results on black solitons. They can be treated in a similar fashion. If $\langle d\rangle<0$, then the integrable NLSE has dark soliton solutions. We consider the black soliton. Equation (52) describes the corrections to

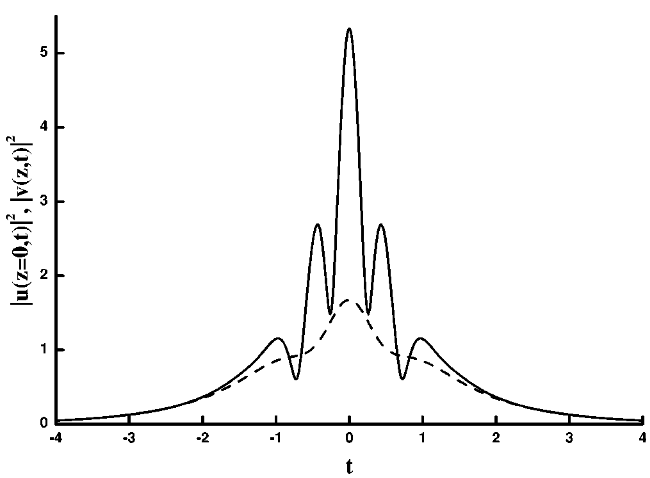


FIG. 2. Squared amplitudes $|u(z,t)|^2$ of the actual soliton at $z=0$ (solid line) and of the averaged soliton $|v(z,t)|^2$ (broken line) for $\varepsilon=0.04$ and $\eta=1$.

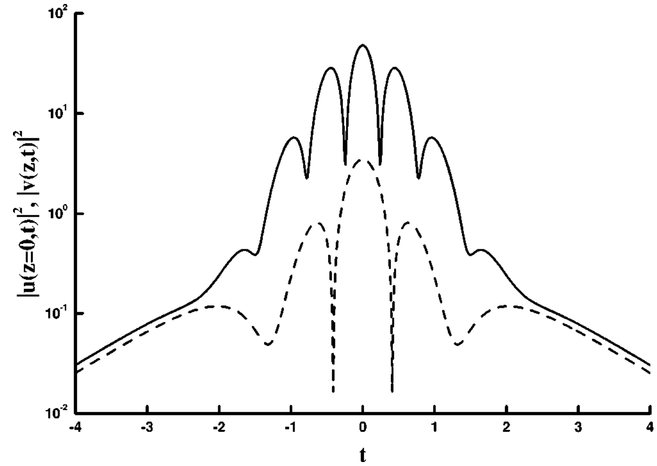


FIG. 3. Squared amplitudes $|u(z,t)|^2$ of the actual soliton at $z=0$ (solid line) and of the averaged soliton $|v(z,t)|^2$ (broken line) in a logarithmic scale for $\varepsilon=0.1$ and $\eta=1$.

the dark solitons, since we did not restrict the boundary conditions during any transformation. Using the ansatz $B(z)=e^{-2i\varepsilon\langle d\rangle\alpha^2 z}f(t)$, we find the fundamental soliton, in the main order, and its first correction,

$$B(z,t)=\exp\{-2i\varepsilon\langle d\rangle\alpha^2 z\}\alpha\left(-\frac{2\langle d\rangle}{\langle c\rangle}\right)^{1/2}(\tanh(\alpha t) \tag{75}$$

$$+\frac{2\rho^2\alpha^2\langle cr^2\rangle}{3\langle c\rangle}\operatorname{sech}^5(\alpha t)[3\sinh(\alpha t)$$
$$+22\sinh^3(\alpha t)-12(\alpha t)\cosh^3(\alpha t)]+\cdots). \tag{76}$$

To get the solution in the original variable $u$, we again apply the expansion of the exponential operator

$$u(z,t)=u_0(z,t)-i\rho r(z)\Delta u_0(z,t)+u_1(z,t)$$
$$-\tfrac{1}{2}\rho^2 r^2(z)\Delta^2 u_0(z,t), \tag{77}$$

where $u_0(z,t)$ is the black soliton,

$$u_0(z,t)=\exp\{-2i\varepsilon\langle d\rangle\alpha^2 z\}\alpha\left(-\frac{2\langle d\rangle}{\langle c\rangle}\right)^{1/2}\tanh(\alpha t), \tag{78}$$

and $u_1(z,t)$ is the correction,

$$u_1(z,t)=\rho^2\exp\{-2i\alpha^2\langle d\rangle z\}\alpha\left(-\frac{2\langle d\rangle}{\langle c\rangle}\right)^{1/2}\frac{2\alpha^2\langle cr^2\rangle}{3\langle c\rangle}$$
$$\times\operatorname{sech}^5(\alpha t)[3\sinh(\alpha t)+22\sinh^3(\alpha t)$$
$$-12(\alpha t)\cosh^3(\alpha t)]+i\varepsilon\tilde{v}_0(z)u_0^2(z,t)$$
$$+\varepsilon\rho\tilde{v}_1(z)N_1(u_0(z,t)). \tag{79}$$

For the black soliton, we have the opposite situation to the bright soliton. Now $r(z)=0$ corresponds to the maximum black soliton width, and the maximum of $r^2(z)$ corresponds to the minimal width.

## VI. DISCUSSION AND CONCLUSIONS

The main result of the present paper are the averaged equations (23) and (52) which generalize all previously known models. Now, we discuss more details of the averaged equations. A comparison with the previous results is made in the Appendix.

The first term on the right-hand side of Eq. (23) is known and was obtained in Ref. [10]. It leads to the Gabitov-Turitsyn equation which was also discussed in Ref. [9]. The next order terms of Eq. (23) are explicitly written out here. The domain of applicability of the full equation is limited by the required smallness of the ratio of the period of the coefficients to a characteristic length of a solution. The higher order approximations can be easly determined in arbitrary order since the initial system is a one-frequency system [13].

Equation (52) is valid for all combinations of the small parameters $\varepsilon$ and $\rho$, for arbitrary forms of the coefficients $\tilde{d}(z)$ and $c(z)$, and for an arbitrary value of the averaged dispersion $\langle d\rangle$. Equation (52), without the two last terms on the right-hand side, was obtained in Ref. [11] for a small variation of the dispersion; the full equation is new (to the best of our knowledge) and has not been presented in literature so far.

We remark that the equations obtained in Refs. [3,6] do not coincide with Eq. (52), even for $c=1$. Actually, in the latter case, the term $\langle F_{02}\rangle$ contains the third power nonlinearity and the fourth-order differentiation. The equation derived in Ref. [3] has the fifth power nonlinearity and the second-order differentiation. The Hasegawa-Kodama equation was obtained under the assumption that the averaged dispersion has the same order as the coefficient in front of the nonlinear term. Equation (9) from Ref. [6] has also the third power nonlinearity and the fourth-order differentiation. But not all terms are important for long-scale solutions; some terms can be eliminated by a transformation of the dependent variable. It is easy to check that Eq. (52), without the last two terms, is Hamiltonian. Equation (17) from Ref. [11] is equal to Eq. (9) from Ref. [6]. Equation (52), without the two last terms on the right-hand-side, can be obtained from the equations just mentioned by a quasi-identical transformation.

We remind the reader that for large parameters $\langle d\rangle$ (in comparison to $\langle c\rangle$), Eq. (A35) (see Appendix) can be reduced to the classical NLSE [18]. For $\langle d\rangle \sim \langle c\rangle$, Eq. (A35) is identical with the Hasegawa-Kodama equation. For a small parameter $\langle d\rangle$, Eq. (A35) is valid but the soliton solutions (73) and (75) loose their applicability.

In conclusion, we demonstrated the applicability of three methods for averaging a nonlinear Schrödinger equation with periodic coefficients. All methods arrive at the same results (see Appendix), although they are very different in practice. The multiple-scale method is difficult to handle and is very elaborate. It requires an *a priori* scaling of variables (for the two small-parameter situation). The Lie transformation is highly systematical. The ansatz is straightforward, and equation Eq. (A19) holds without any assumptions for the starting equation. The method works fast and efficiently, but the calculation of the Lie brackets is complex and elaborate. Also, the Lie method works with powers of small parameters. Sometimes it does not allow us to do a global transformation for additional small parameters. The Bogolyubov transformation is the most elegant method; the main result is reached very fast. Before applying the Bogolyubov transformation, one should formulate the basic equation in standard form. Otherwise the result may be wrong [14].

We considered the two-parameter situation, assuming that the high-frequency, variable part of the dispersion is larger than the mean value. The first small parameter is the ratio of the length of the dispersion map to the length of a solution. The second parameter is the value of the (mean-free) integral for the variable part of a dispersion. This is an important difference compared to previous assumptions on the amplitude of the variable part of dispersion. Note that, for short-scale management, the amplitude can be very large while the function $R(z)$ has a small amplitude [19]. We obtained the asymptotic equation in the main order, and with first corrections. The calculations are valid for all ranges of parameters as long as they are small. The coefficients of the obtained equation were evaluated for a two-step dispersion map with an exponential loss function.

The corrections to the fundamental bright and black soliton solutions were obtained. These corrections show different behaviors. The bright soliton has a minimal width at a point where $R^2(z)=0$ while the dark soliton has a maximum width at this point.

## ACKNOWLEDGMENTS

The work was performed under the auspices of the Sonderforschungsbereich No. 591. Discussions with Wolfgang Laedke, Fabio Mancin, Alexander Posth, and Marcus Neuer are gratefully acknowledged.

## APPENDIX: COMPARISON WITH OTHER METHODS

In this Appendix we compare the Bogolyubov method with other averaging methods, such as the multiple-scale averaging and a technique based on Lie transformation.

### 1. Multiple-scale averaging

We start with the NLSE in the form

$$iu_z+[\varepsilon\langle d\rangle+\delta\hat{d}(z)]u_{tt}+\varepsilon c(z)|u|^2u=0, \tag{A1}$$

where $\hat{d}\sim 1$ and $\hat{d}(z)=\hat{d}(z+1)$. We are interested in relatively large variations of the dispersion. Therefore we assume $\delta \gg \varepsilon$. Let us introduce the slow variable $x=\delta z$. Then we get

$$iu_x+\left[\frac{\varepsilon}{\delta}\langle d\rangle+\hat{d}\left(\frac{x}{\delta}\right)\right]u_{tt}+\frac{\varepsilon}{\delta}c(z)|u|^2u=0. \tag{A2}$$

The dependent variable $u$ should have a derivative $u_x$ of the order of 1. In addition, a fast oscillating coefficient $\hat{d}(x/\delta)$ appears for $\delta \ll 1$. In this case, the equation is ready for application of a multiple-scale averaging.

To be specific, we choose $\delta=\varepsilon^{1/2}$. The case with $\delta=\varepsilon$ was considered by Yang and Kath [6]. We also choose $c(z)\equiv 1$ for simplicity. Next, we expand $u$ in a power series of $\varepsilon^{1/2}$,

$$u=U+\varepsilon^{1/2}u_1+\varepsilon u_2+\cdots, \tag{A3}$$

with the abbreviations $U=U(z_0,z_1,z_2,\ldots)$ and $u_k=u_k(z,U)$ for $k=1,2,\ldots$. We use the scale $u_k\sim O(\varepsilon^{k/2})$, introduce a fast variable $\xi=\varepsilon^{-1/2}x=z$, and the slow variables $z_n=\varepsilon^{k/2}x$. We assume that $U$ is independent of the fast variable $\xi$, $\langle u\rangle=U$, and $\langle u_k\rangle=0$. Substituting Eq. (A3) into Eq. (A2) and using

$$\frac{\partial U}{\partial x}=\frac{\partial U}{\partial z_0}+\varepsilon^{1/2}\frac{\partial U}{\partial z_1}+\varepsilon\frac{\partial U}{\partial z_2}+\cdots \tag{A4}$$

we get

$$\begin{aligned}&i\left[\frac{\partial U}{\partial z_0}+\frac{\partial u_1}{\partial \xi}+\varepsilon^{1/2}\left(\frac{\partial U}{\partial z_1}+\frac{\partial u_1}{\partial z_0}+\frac{\partial u_2}{\partial z}\right)+\varepsilon\left(\frac{\partial U}{\partial z_2}+\frac{\partial u_1}{\partial z_1}+\frac{\partial u_2}{\partial z_0}\right.\right.\\&\quad\left.\left.+\frac{\partial u_3}{\partial \xi}\right)+\cdots\right]+\varepsilon^{1/2}\langle d\rangle\left(\frac{\partial^2 U}{\partial t^2}+\varepsilon^{1/2}\frac{\partial^2 u_1}{\partial t^2}+\cdots\right)\\&\quad+\tilde{d}\left(\frac{\partial^2 U}{\partial t^2}+\varepsilon^{1/2}\frac{\partial^2 u_1}{\partial t^2}+\cdots\right)+\varepsilon^{1/2}[|U|^2U\\&\quad+\varepsilon^{1/2}(2|U|^2u_1+U^2u_1^*)+\varepsilon(U^*u_1^2+2|U|^2u_2\\&\quad+2U|u_1|^2+U^2u_2^*)+\cdots]=0.\end{aligned} \tag{A5}$$

Next, we collect terms of the same orders and eliminate the secular terms. In the zeroth order one obtains

$$i\frac{\partial U}{\partial z_0}+i\frac{\partial u_1}{\partial \xi}+\tilde{d}(\xi)\frac{\partial^2 U}{\partial t^2}=0. \tag{A6}$$

By averaging this equation, where the fast varying terms cancel, we get

$$i\frac{\partial U}{\partial z_0}=0\quad\text{and}\quad\frac{\partial u_1}{\partial \xi}=i\tilde{d}\frac{\partial^2 U}{\partial t^2}. \tag{A7}$$

By integration over $\xi$, where the slowly varying terms are assumed to be constant during the integration, we achieve

$$u_1=i\tilde{r}U_{tt}\sim O(\varepsilon^{1/2}), \tag{A8}$$

with $\tilde{r}=r-\langle r\rangle$ and $r=\int_0^{\xi}(d-\langle d\rangle)\,d\xi'$. Using this result in the next order $\varepsilon^{1/2}$, we find

$$i\frac{\partial U}{\partial z_1}=-\langle d\rangle\frac{\partial^2 U}{\partial t^2}-|U|^2U\quad\text{and}\quad u_2=-\frac{1}{2}U_{tttt}(\tilde{r}^2-2M), \tag{A9}$$

with $M=\frac{1}{2}(\langle r^2\rangle-\langle r\rangle^2)$.

In the first order we obtain

$$\frac{\partial U}{\partial z_2}=0\quad\text{and}$$

$$u_3=2(s-\langle s\rangle)[U_t^2U^*+2U|U_t|^2+U^2U_{tt}^*]-\frac{i}{2}U_{6t}P_1(r), \tag{A10}$$

where $s=[r](\xi)$.

In the order 3/2 we finally arrive at

$$\begin{aligned}i\frac{\partial U}{\partial z_3}&=4M[U^2U_{tttt}^*+6UU_tU_{ttt}^*+7U_t^2U_{tt}^*+10|U_t|^2U_{tt}\\&\quad+2U_{ttt}(|U|^2)_t+5U|U_{tt}|^2+\tfrac{5}{2}U^*U_{tt}^2].\end{aligned} \tag{A11}$$

By inserting all results into Eq. (A4) we get

$$\begin{aligned}&iU_x+\varepsilon^{1/2}\langle d\rangle U_{tt}+\varepsilon^{1/2}|U|^2U\\&\quad=4\varepsilon^{3/2}M[U^2U_{tttt}^*+6UU_tU_{ttt}^*+7U_t^2U_{tt}^*+10|U_t|^2U_{tt}\\&\qquad+2U_{ttt}(|U|^2)_t+5U|U_{tt}|^2+\tfrac{5}{2}U^*U_{tt}^2]+O(\varepsilon^2).\end{aligned} \tag{A12}$$

Next, we multiply both sides of Eq. (A12) with $\delta=\varepsilon^{1/2}$ to rewrite

$$\begin{aligned}&iU_z+\varepsilon\langle d\rangle U_{tt}+\varepsilon|U|^2U\\&\quad=4\varepsilon^2M[U^2U_{tttt}^*+6UU_tU_{ttt}^*+7U_t^2U_{tt}^*+10|U_t|^2U_{tt}\\&\qquad+2U_{ttt}(|U|^2)_t+5U|U_{tt}|^2+\tfrac{5}{2}U^*U_{tt}^2]+O(\varepsilon^{5/2}).\end{aligned} \tag{A13}$$

The transformation between the averaged and actual variables is

$$\begin{aligned}u&=U+i(R-\langle R\rangle)U_{tt}-\tfrac{1}{2}U_{tttt}[(R-\langle R\rangle)^2-2M]\\&\quad-\tfrac{1}{2}iU_{6t}P_1(R)+\varepsilon^{1/2}\tilde{S}N_1+O(\varepsilon^2).\end{aligned} \tag{A14}$$

We have used the relations $R=\varepsilon^{1/2}r$ and $S=\varepsilon s$. The expression $P_1(R)\sim O(\varepsilon^{3/2})$ is a polynomial in $R$.

### 2. Lie transformation

The Lie-transformation technique was developed for ordinary differential equations, but we can also use it for a partial differential equation, by interpreting the latter as a infinite system of ordinary differential equations. To use the Lie transformation here, we rearrange Eq. (A2) in the form

$$\frac{\partial u}{\partial z}=i\varepsilon^{1/2}\langle d\rangle u_{tt}+i\varepsilon^{1/2}|u|^2u+i\tilde{d}u_{tt}=X_0+\tilde{d}X_{0D}=X[u,u^*]. \tag{A15}$$

Thereby, $X$ depends on the infinite set of variables $(u,u_t,u_{tt},\ldots,u^*,u_t^*,u_{tt}^*,\ldots)$, indicated by $X[u,u^*]$.

Obviously it is

$$X_0\sim O(\varepsilon^{1/2}),\quad X_{0D}\sim O(1). \tag{A16}$$

The Lie transformation is applied in the form

$$u=e^{\vec{\phi}\vec{\nabla}}v=v+\phi+\frac{1}{2!}(\vec{\phi}\cdot\vec{\nabla})\phi+\frac{1}{3!}[(\vec{\phi}\cdot\vec{\nabla})\phi\cdot\vec{\nabla}]\phi+\cdots \tag{A17}$$

to transform Eq. (A15) into

$$\frac{\partial v}{\partial z}=Y[v,v^*]. \tag{A18}$$

Thereby $\vec{\phi}=(\phi,\phi_t,\ldots)$ and $\vec{\nabla}=(\partial/\partial v,\partial/\partial v_t,\ldots)$ are infinite-dimensional vectors.

Inserting transformation (A17) into Eq. (A15), we can derive the following relation between $Y$ and $X$ (thereby $X=X[v,v^*]$ holds):

$$Y+\frac{\partial\phi}{\partial z}+\frac{1}{2!}\left[\phi,\frac{\partial\phi}{\partial z}\right]+\frac{1}{3!}\left[\phi,\left[\phi,\frac{\partial\phi}{\partial z}\right]\right]+\cdots$$
$$=X+[\phi,X]+\frac{1}{2!}[\phi,[\phi,X]]+\cdots. \tag{A19}$$

In this equation, the Lie bracket is introduced via

$$[A,B]=(\vec{A}\cdot\vec{\nabla})B-(\vec{B}\cdot\vec{\nabla})A, \tag{A20}$$

with

$$(\vec{A}\cdot\vec{\nabla})B=\sum_{n=0}^{\infty}\left(A_{(nt)}\frac{\partial}{\partial v_{(nt)}}+A^*_{(nt)}\frac{\partial}{\partial v^*_{(nt)}}\right)B. \tag{A21}$$

Next, we expand $Y$ and $\phi$,

$$Y=Y_0+Y_1+Y_2+Y_3+\cdots,\quad \phi=\phi_1+\phi_2+\phi_3+\cdots, \tag{A22}$$

with the scalings

$$Y_k\sim O(\varepsilon^{k/2})\quad\text{and}\quad \phi_k\sim O(\varepsilon^{k/2}), \tag{A23}$$

respectively. Since $\phi$ scales with the fast variation, we assume

$$\frac{\partial\phi_k}{\partial z}\sim O(\phi_{k-1})\sim O(\varepsilon^{(k-1)/2}). \tag{A24}$$

Using the expansion, we get in zeroth order

$$Y_0+\frac{\partial\phi_1}{\partial z}=\tilde{d}X_{0D}, \tag{A25}$$

and by averaging

$$Y_0=0\Rightarrow\frac{\partial\phi_1}{\partial z}=\tilde{d}X_{0D}\Rightarrow\phi_1=\tilde{R}X_{0D}\sim O(\varepsilon^{1/2}). \tag{A26}$$

The slowly varying terms will again be assumed to be constant during the integration. In next order $\epsilon^{1/2}$ we find

$$Y_1=X_0\Rightarrow\frac{\partial\phi_2}{\partial z}=0\Rightarrow\phi_2=0. \tag{A27}$$

In first order we obtain

$$Y_2=0\Rightarrow\frac{\partial\phi_3}{\partial z}=\tilde{R}[X_{0D},X_0]\Rightarrow\phi_3=\tilde{S}[X_{0D},X_0]\sim O(\varepsilon^2). \tag{A28}$$

Order $\epsilon^{3/2}$ leads to

$$Y_3=-\frac{1}{2}\langle\tilde{d}\tilde{S}\rangle[X_{0D},[X_{0D},X_0]]\sim O(\varepsilon^2). \tag{A29}$$

Finally, we obtain the following equation for the averaged dynamics:

$$\frac{\partial v}{\partial x}=i\varepsilon^{1/2}\langle d\rangle v_{tt}+i\varepsilon^{1/2}|v|^2v-\frac{1}{2}$$
$$\times\langle\tilde{d}(S-\langle S\rangle)\rangle[X_{0D},[X_{0D},X_0]]+O(\varepsilon^2). \tag{A30}$$

Calculating the Lie brackets

$$[X_{0D},X_0]=2\varepsilon^{1/2}(v^*_{tt}v^2+v_t^2v^*+2vv_tv_t^*)=:\varepsilon^{1/2}N_1(v), \tag{A31}$$

$$[X_{0D},[X_{0D},X_0]]=-4i\varepsilon^{1/2}(v^*_{tttt}v^2+4v^*_{ttt}vv_t+4v^*_{tt}v_t^2$$
$$+2v|v_{tt}|^2+4v_{tt}|v_t|^2+v_{tt}^2v^*)$$
$$=:-i\varepsilon^{1/2}N_2(v), \tag{A32}$$

we obtain

$$\langle\tilde{d}\tilde{S}\rangle=-2M\overset{\langle R\rangle=0}{=}-\langle R^2\rangle\sim O(\varepsilon). \tag{A33}$$

Thereby, we can choose $\langle R\rangle=0$, since $R$ is only determined modulo a constant. The averaged equation is

$$iv_x+\varepsilon^{1/2}\langle d\rangle v_{tt}+\varepsilon^{1/2}|v|^2v=\tfrac{1}{2}\varepsilon^{1/2}\langle R^2\rangle N_2+O(\varepsilon^2). \tag{A34}$$

Next, we multiply with $\varepsilon^{1/2}$ to get

$$iv_z+\varepsilon\langle d\rangle v_{tt}+\varepsilon|v|^2v=\tfrac{1}{2}\varepsilon\langle R^2\rangle N_2+O(\varepsilon^{5/2}). \tag{A35}$$

The transformation is

$$u=v+iRv_{tt}-\tfrac{1}{2}R^2v_{tttt}-\tfrac{1}{6}iR^3v_{6t}+\varepsilon^{1/2}\tilde{S}N_1+O(\varepsilon^2). \tag{A36}$$

We can transform from Eq. (A13) to Eq. (A35) by

$$U=v-\frac{1}{2}\langle R^2\rangle v_{tttt}+\left(-\frac{i}{6}\langle R^3\rangle v_{6t}\right). \tag{A37}$$

The last term on the right-hand side is only necessary for the identification of the transformations, not for the averaged equations. Because of this, it follows that the different averaging methods lead to the same results.

[1] G.P. Agrawal, *Nonlinear Fiber Optics* (Academic Press, New York, 1989).
[2] A. Hasegawa and Y. Kodama, *Solitons in Optical Communications* (Clarendon Press, Oxford, 1995).
[3] A. Hasegawa and Y. Kodama, Opt. Lett. **16**, 1385 (1991).
[4] I. Gabitov, T. Schäfer, and S.K. Turitsyn, Phys. Lett. A **265**, 274 (2000).
[5] Y. Kodama, *Massive WDM and TDM Soliton Transmission Systems* (Kluwer Academic, Dordrecht, MA, 2000), p. 129.
[6] T.-S. Yang and W.L. Kath, Opt. Lett. **22**, 985 (1997).
[7] T.-S. Yang, W.L. Kath, and S.K. Turitsyn, Opt. Lett. **23**, 597 (1998).
[8] T. Yang, W.L. Kath, and S.K. Turitsyn, J. Eng. Math. **36**, 163 (1999).
[9] M.J. Ablowitz and G. Biondini, Opt. Lett. **23**, 1668 (1998).
[10] I. Gabitov and S.K. Turitsyn, Pıs'ma Zh. Eksp. Teor. Fiz. **63**, 814 (1996) [JETP Lett. **63**, 861 (1996)].
[11] K.H. Spatschek, S.K. Turitsyn, and Y.S. Kivshar, Phys. Lett. A **204**, 269 (1995).
[12] D.E. Pelinovsky and V. Zharnitsky, SIAM (Soc. Ind. Appl. Math.) J. Appl. Math. **63**, 745 (2003).
[13] N.N. Bogoliubov and Y.A. Mitropolski, *Asymptotic Methods in the Theory of Nonlinear Oscillations* (Gordon and Breach, New York, 1961).
[14] J.A. Sanders and F. Verhulst, *Averaging Methods in Nonlinear Dynamical Systems* (Springer, New York, 1985).
[15] P.L. Lushnikov, Opt. Lett. **25**, 1144 (2000).
[16] S.B. Medvedev, O.V. Styrina, S.L. Musher, and M.P. Fedoruk, Phys. Rev. E **66**, 066607 (2002).
[17] S.K. Turitsyn, T. Schäfer, K.H. Spatschek, and V.K. Mezentsev, Opt. Commun. **163**, 122 (1999).
[18] S.B. Medvedev and S.K. Turitsyn, Pıs'ma Zh. Eksp. Teor. Fiz. **69**, 465 (1999) [JETP Lett. **69**, 499 (1999)].
[19] S.K. Turitsyn, E.G. Turitsyna, S.B. Medvedev, and M.P. Fedoruk, Phys. Rev. E **61**, 3127 (2000).