31 January 2000

PHYSICS LETTERS A

Physics Letters A 265 (2000) 274–281

www.elsevier.nl/locate/physleta

Lie-transform averaging in nonlinear optical transmission systems with strong and rapid periodic dispersion variations

Ildar Gabitov [a,*], Tobias Schäfer [b], Sergei K. Turitsyn [c]

[a] Los Alamos National Laboratory, MS-B284, Los Alamos, NM 87545, USA

[b] Institut für Theoretische Physik I, Heinrich-Heine-Universitaet Duesseldorf, 40225 Duesseldorf, Germany

[c] Division of Electr. Eng. & Comp. Science, Aston University, Birmingham B4 7ET, UK

Received 16 November 1999; accepted 17 December 1999

Communicated by A.R. Bishop

Abstract

Using Lie-transform techniques, we derive higher-order corrections to the path-averaged model governing evolution of dispersion-managed solitons in the spectral domain. The result holds in the case of arbitrary (including moderate and strong) dispersion. The general theory is illustrated by deriving the exact formulas + for a specific symmetric dispersion map.

Development of ultrafast high-bit-rate optical communication lines is in the focus of extensive research because of the present growing demands on the capacity of transmission systems. High capacity communication system design has two objectives: (i) long-haul transmission systems using low dispersion fibers (dispersion shifted fibers) and (ii) upgrading existing fiber links based on highly dispersive (in the main fiber transparency window at 1.55 μm wavelength) standard telecommunication fibers.

The main factor that limits the bit-rate is pulse-broadening due to the chromatic dispersion of the optical fiber. This broadening is characterized by the dispersion length $Z_{\rm dis} \sim (d \times ({\rm BR})^2)^{-1}$. Here, d is the fiber chromatic dispersion and BR is the bitrate. The dispersion length $Z_{\rm dis}$ is the distance at which the pulse-width approximately doubles due to dispersive broadening. This distance decreases as inverse square of the bit-rate. Another important factor which limits capacity of the fiber links is the nonlinearity of the fiber refractive index (Kerr nonlinearity) $n = n_0 + \alpha I$ (where n_0 is linear part of the refractive index, I is pulse intensity, and α is coefficient of the Kerr nonlinearity). The spectrum of an optical pulse with characteristic power P_0 will experience noticeable nonlinear distortion at distances greater than the characteristic nonlinear length, $Z_{\rm nl} = (\alpha P_0)^{-1}$.

In traditional long-haul systems using low dispersion fibers the distance between optical amplifiers required for compensating fiber losses is considerably shorter than that of both the characteristic dispersion length $Z_{\rm dis}$ and the characteristic nonlinear length $Z_{\rm nl}$. In other words, both dispersion and nonlinearity can be treated as perturbations on the scale of the distance between amplifiers and, to first order, only the fiber losses and periodic amplification are

* Corresponding author. Fax: +1-505-665-5757; e-mail: ildar@t7.lanl.gov

PII: S0375-9601(99)00901-9

significant factors. These factors cause amplitude oscillations, but the shape of the pulse remains approximately unchanged. At the large-scales, pulse propagation in these communication systems is described by the well-established guiding-center (path-average) theory [1–4]. The nonlinear Schrödinger equation (NLSE) is the basis for path-averaged propagation model.

For the well-known case of upgrading existing optical links, based on standard mono-mode fiber (SMF), different physical scenarios and modeling equations are valid. SMF has rather high (approximately $\sim 17\,\mathrm{ps/nm\cdot km}$) dispersion in the $1.55\,\mu\mathrm{m}$ window of optical transparency. As a comparison, the dispersion value for dispersion shifted fiber used in long haul transmission links is $\sim 1\,\mathrm{ps/nm\cdot km}$. The negative impact of fiber chromatic dispersion on the data stream for the same value of bitrate is differs by more greater than order of magnitude for these two values of dispersion. For increasing bitrates system performance degrades as quadratic function of the bitrate value. For multigigabit transmission at $1.55\,\mu\mathrm{m}$, the corresponding dispersion length in SMF is approximately equal to the amplification distance in the existing networks. Consequently traditional guiding-center (path-averaged) soliton theory can not be applied. The limitations caused by fiber chromatic dispersion can be minimized by dispersion compensation, using pieces of fiber with high dispersion of the opposite sign to the dispersion of the original transmission fiber [5]. Dispersion compensation is an attractive technique to enhance the transmission capacity of fiber communication lines. The master equation for dispersion-managed systems is the perturbed NLSE with fast periodic dispersion variations. Stable propagation of dispersion-managed (DM) soliton is possible in such systems [6–10]. The true DM soliton is a periodic solution of the master equation that is entirely recovered after each compensation period [8–10]. In the case of weak dispersion, the Lie transform is a powerful method which can be used to average the basic equation directly [11]. For strong dispersion, however, direct averaging is not possible due to the large variations of the coefficient of dispersion. Different theoretical approaches have been developed to describe path-averaged propagation of DM soliton: multi-scale analysis [12,13]; different averaging methods [11,14,15], including averaging in the spectral domain [9,10,16,17] and an expansion of DM soliton in the basis of the chirped Gauss–Hermite functions [15,18].

Due to the practical importance of the problem, it is very useful to develop different analytical methods to describe the properties of DM soliton. A variety of complementary mathematical approaches can be advantageously exploited to find an optimal and economical description of any specific practical application. Note that the problem treated is not only of a high practical interest, but it is also an interesting fundamental mathematical problem. To determine the limitations of the averaged models it is important to further develop and extend averaging methods to derive first-order corrections to the averaged equations. The averaging approaches considered to date have much in common.

To make the averaging procedure possible, one should first eliminate the large variations of the coefficient of the dispersion. In other terms, it is necessary to rewrite the master equation in a different form before applying Lie techniques. This can be done in two ways. One possibility is to make the lens transform accounting for the dynamics of self-similar pulse core. The lens transformation was introduced for the first time for the description of light focusing in cubic media [19]. For optical telecommunication models it was independently applied in [20]. This approach is appropriate for describing the single optical pulse dynamics. A second method [9,10] is to apply a Fourier transform in order to remove rapid variations from the basic equation, thus preparing the equation for averaging. This approach, in addition to single pulse dynamics, can be applied for the analysis of pulse interactions for wavelength division multiplexing techniques (WDM), when several data streams with different carrier frequencies propagate through the same fiber. Presently, the WDM technique is the main approach for optical fiber communications.

In this Letter, we extend the analysis of Ref. [9,10] and using Lie-transforms, we derive higher-order corrections for the path-averaged model in the frequency domain [21]. We start our analysis from the master equation

$$iq_z + \tfrac{1}{2}d(z)\,q_{tt} + q^2 q^* = Rq\,. \tag{1}$$

Here, $d(z)$ measures dispersion variations, z is a distance normalized with respect to the soliton period found for the equivalent uniform-dispersion line with the same average dispersion as in the system under consideration. The right-hand-term, Rq, models fast pulse amplitude variations, due to the losses/gain, that could be present in addition to the dispersion variations, we assume that $z_a \ll 1$ is a fast scale for both dispersion and loss/gain variations.

When dispersion variations are moderate in size,

$$\int_0^{z_a} dz\, d(z) \ll 1, \tag{2}$$

then, the averaging method [11], based on the Lie transform, can be applied to (1) and the resulting 'averaged' system represents an almost 'pure' NLSE with small second order (in z_a) correction, provided resonances are absent. These results confirm the stability of solitons under small periodic amplitude perturbations. When dispersion variations are strong,

$$\int_0^{z_a} dz\, d(z) \geq 1, \tag{3}$$

then the averaging method [11] can not be applied directly to (1). Nevertheless, by applying simpler means, an averaged system was derived [9,10] for the case (3). This system, which, in general, is no longer a NLSE, was obtained [9,10] in the zeroth order (in z_a) only, because the calculation of important higher order corrections (in z_a) could not be done by simple means.

We present a generalization of the averaging method, based on the Lie transform, can be used for the general case (3) as well. Our method makes it possible to derive the averaged system to any order.

The problem of pulse propagation in an optical transmission line composed of optical fibers with alternating dispersion characteristics (dispersion management) is one important example of a system modeled by (1). Here

$$Rq = i\gamma q + iG(z)q, \tag{4}$$

where $\gamma < 0$ is the damping constant. The amplification necessary to compensate losses, is defined by $G(z)$

$$G(z) = G_0 \sum_{n=0}^{N} \delta(z - nz_a), \tag{5}$$

i.e. a periodic sequence of δ-functions. The small parameter z_a is the amplifier spacing. For many practical cases, damping term and dispersion terms are not small at all, while the nonlinear term, $q^2 q^*$, can be treated as a small perturbation. Before applying the averaging method one needs to prepare the system (1) by reducing it to the standard form with a small right hand side [22]

$$dx/dZ = \epsilon f(x,Z) \tag{6}$$

in two steps. First, by performing the transformation

$$q(t,z) = a(z)w(t,z), \quad ida/dz = Ra, \tag{7}$$

the large damping coefficient γ and the δ-function amplification term from the system (1) are removed. We then obtain a system with a variable nonlinearity coefficient $a^2(z)$

$$iw_z + \tfrac{1}{2}d(z)w_{tt} + a^2(z)w^2 w^* = 0. \tag{8}$$

The rapidly varying coefficients $d(z)$ and $a^2(z)$ can be split into two parts

$$d(z) = \langle d \rangle + \tilde{d}, \quad a^2(z) = I(z) = \langle I \rangle + \tilde{I}. \tag{9}$$

Here, the mean values $\langle d \rangle$, $\langle I \rangle$ and the periodic function $\tilde{I}$ are of order 1, while the variable part $\tilde{d}$ of the dispersion can be much greater than 1 for strong dispersion variations. To eliminate this large coefficient from the model system (8), we need to perform the second step of applying a Fourier transform – based reformulation:

$$w(t,z) = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\, u_\omega(\omega,z) \times \exp\left[-i\omega t - \frac{i}{2}\omega^2 \tilde{d}_1 + i\omega^2 \frac{\langle I\tilde{d}_1 \rangle}{\langle 2I \rangle}\right], \tag{10}$$

where

$$\frac{d}{dz}\tilde{d}_1 = \tilde{d} \to \tilde{d}_1 = \int_0^z dz\, \tilde{d}(z) + d_{10}, \tag{11}$$

and the integration constant (over z) d_{10} is fixed by the condition that the function $\tilde{d}_1$ has zero mean. As a result, we have an integro-differential equation in the frequency domain for u_ω

$$i\frac{du_\omega}{dz} - \tfrac{1}{2}\langle d\rangle\omega^2 u_\omega + J_\omega = 0\,, \qquad (12)$$

with the integral term J_ω given by

$$J_\omega(\omega,z) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty} d\omega_1 d\omega_2 d\omega_3 \times \delta(\omega_1+\omega_2-\omega_3-\omega) F u_{\omega_1}u_{\omega_2}u^*_{\omega_3}. \qquad (13)$$

The kernel F depends on all frequencies and the distance

$$F(\omega_1,\omega_2;\omega_3,\omega;z) = I(z)\exp\left[-\frac{i}{2}\beta\tilde{d}_1 + i\beta\frac{\langle I\tilde{d}_1\rangle}{\langle 2I\rangle}\right], \qquad (14)$$

$$\beta = \omega_1^2+\omega_2^2-\omega_3^2-\omega^2\,. \qquad (15)$$

Note that u_ω is not exactly the Fourier transform of w due to the presence of the terms $\omega^2\tilde{d}_1$ and $\omega^2\langle I\tilde{d}_1\rangle/\langle 2I\rangle$ in (10).

In the second step (whose aim is to eliminate the $\tilde{d}_1$ coefficient from (8)), we do not really need to use the integration constants d_{10} and $\langle I\tilde{d}_1\rangle/\langle 2I\rangle$ in (10). We use these constants only for the purpose of 'tuning up' the kernel F in (14), so that our final averaged system is consistent with the results of [23]. Consistency means that our results must agree with those of [23] in the limit of moderate dispersion variations. (See below). Finally, by scaling the distance z: $z = Z/z_a$ one can reduce (12) to the form (6) and prepare for the averaging method. To simplify notations we keep the 'old' variable z as the distance. To do the averaging based on the Lie-transform, we also have to rewrite (12) in the time-domain by taking the inverse Fourier transform

$$\frac{du}{dz} = X[u,u^*;z] = \frac{i}{2}\langle d\rangle u_{tt} + iJ\,, \qquad (16)$$

where u is the inverse Fourier transform of u_ω

$$u(t,z) = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\, u_\omega(\omega,z)\exp[-i\omega t]\,. \qquad (17)$$

The term J is given by

$$J = F\left(i\partial_{t_1}, i\partial_{t_2}; -i\partial_{t_3}, -i\partial_t = i\left(\partial_{t_1}+\partial_{t_2}+\partial_{t_3}\right); z\right) \times\{u(t_1)u(t_2)u^*(t_3)\}|_{t_i=t}\,, \qquad (18)$$

where the operator F is obtained by taking (14) and substituting $\omega \to i\partial_t$. The first derivative, $i\partial_{t_1}$, in the operator, F, acts on the first multiplier, $u(t_1)$, only, etc..... After performing all differentiations, we set $t_i = t, i = 1,2,3$. One can expand the exponential expression (14) into a Taylor series over β. As a result, one will have J as a differential series. Now we are ready to do the averaging.

The basic idea of the averaging method is to simplify a system, like (16), by eliminating the explicit dependence on z and expressing the simplified system in terms of a new variable v,

$$\frac{dv}{dz} = Y[v,v^*]\,, \qquad (19)$$

where the right hand side (rhs) Y is yet to be calculated. To reach this goal one uses a near-identity transformation: $u \to v$, generalized for the case of an infinite number of degrees of freedom [23,24] represented by v,v^* and all its derivatives $v_t, v_t^*, v_{tt}, v_{tt}^*, \cdots$. When this a transformation is written in exponential form (Lie-transform) [25]

$$u = e^{\phi\nabla}v = v + \phi + \tfrac{1}{2}\phi\nabla\phi + \cdots\,, \qquad (20)$$

all calculations can be elegantly performed. The rhs of (20) is obtained by a formal expansion of the exponent $e^{\phi\nabla}$ in a Taylor series over $\phi\nabla$. The directional derivative $\phi\nabla$ is defined as

$$\phi\nabla = \sum_{n=0}^{\infty}\left[\phi_{nt}\frac{\partial}{\partial v_{nt}} + \phi^*_{nt}\frac{\partial}{\partial v^*_{nt}}\right], \qquad (21)$$

where $\phi_{nt} = \partial^n\phi/\partial t^n, v_{nt} = \partial^n v/\partial t^n$, and the functional ϕ depends on v,v^*, all its t derivatives, and the distance z. Before applying these general ideas to our particular case (16), we would like to point out the first major difference between our problem of strong dispersion variations and that of moderate dispersion variations [23]. In a moderate case, the rhs of (16) is nothing but a differential polynomial, whereas in our case we have to deal with the differential series J. Luckily this complication does not prohibit the use of the 'old' definitions of the Lie-transform and of the directional derivative.

By inserting (20) into (16), one obtains in compact form the general rule for transformations of X (16) into Y (19) under a Lie-transformations [23]

$$Y\nabla + \left(\frac{\partial}{\partial z} e^{\phi\nabla}\right) e^{-\phi\nabla} = e^{\phi\nabla} X\nabla e^{-\phi\nabla}, \tag{22}$$

where according to the Campbell–Baker–Hausdorff formula

$$\begin{aligned} e^{\phi\nabla} X\nabla e^{-\phi\nabla} &= \left(\sum_{n=0}^{\infty} \frac{1}{n!}[\phi,[\phi,\cdots[\phi,X]]]\right)\nabla \\ &= \left(X + [\phi,X] + \frac{1}{2!}[\phi,[\phi,X]] + \cdots\right)\nabla. \end{aligned} \tag{23}$$

Also

$$\left(\frac{\partial}{\partial z} e^{\phi\nabla}\right) e^{-\phi\nabla} = \left(\phi_z + \tfrac{1}{2}[\phi,\phi_z] + \frac{1}{3!}[\phi,[\phi,\phi_z]] + \cdots\right)\nabla. \tag{24}$$

Eqs. (22)–(24) can be used to simplify greatly the following calculations. By expanding both Y and ϕ into a series in the small parameter, z_a: $Y = Y_0 + Y_1 + Y_2 + \cdots$, $\phi = \phi_1 + \phi_2 + \phi_3 + \cdots$; $Y_n, \phi_n = O(z_a^n)$, and collecting common terms in (22), one obtains expressions for the unknowns Y_n and ϕ_n. Note, differentiation by z lowers the order of ϕ_n by one: $\partial\phi_n/\partial z = O(z_a^{n-1})$. At the zeroth order, one gets

$$Y_0 + \frac{\partial\phi_1}{\partial z} = X, \tag{25}$$

where X (16) is of order 1. Note that all ϕ_n's are periodic functions of z. So by averaging (25) one gets

$$Y_0 = \langle X\rangle = \frac{i}{2}\langle d\rangle v_{tt} + i\langle J\rangle, \tag{26}$$

$$\langle J\rangle = \langle F\rangle\{v(t_1)v(t_2)v^*(t_3)\}|_{t_i=t}, \tag{27}$$

where the brackets denote averaging over the fast scale, z_a

$$\langle f\rangle = \frac{1}{z_a}\int_0^{z_a} dz f(z). \tag{28}$$

As a result

$$\frac{\partial\phi_1}{\partial z} = i\tilde{J} = iJ - i\langle J\rangle, \quad \phi_1 = \phi_{10} + iJ_{\tilde{1}}, \tag{29}$$

where

$$\frac{\partial}{\partial z} J_{\tilde{1}} = \tilde{J}, \quad \langle J_{\tilde{1}}\rangle = 0. \tag{30}$$

The integration constant $\phi_{10}[v,v^*]$ will be fixed in the next order. This concludes the zeroth order, Y_0, calculations. Note that the zeroth order term Y_0 was calculated in [9,10] by directly integrating (12) over the amplifier spacing z_a.

Before proceeding to the next order, we would like to point out a few important properties of the zeroth order system

$$\frac{dv}{dz} = Y_0, \tag{31}$$

which we would like to formulate in terms of v_ω, the Fourier transform of v. System (31) conserves the number of particles I_1 and the momentum I_2,

$$I_1 = \int_{-\infty}^{\infty} d\omega\, v_\omega v_\omega^*, \quad I_2 = \int_{-\infty}^{\infty} d\omega\, \omega\, v_\omega v_\omega^*. \tag{32}$$

It is also a Hamiltonian system

$$i\frac{dv_\omega}{dz} = \frac{\delta H}{\delta v_\omega^*}, \quad H = H_0 + H_{\rm int}, \tag{33}$$

$$H_0 = \frac{\langle d\rangle}{2}\int_{-\infty}^{\infty} d\omega\, \omega^2 v_\omega v_\omega^*, \tag{34}$$

$$H_{\rm int} = \frac{1}{2}\frac{1}{(2\pi)^2}\int_{-\infty}^{\infty} d\omega_1 d\omega_2 d\omega_3 d\omega_4 \times \delta(\omega_1+\omega_2-\omega_3-\omega_4), \tag{35}$$

$$\langle F\rangle(\omega_1,\omega_2;\omega_3,\omega_4)\, v_{\omega_1} v_{\omega_2} v_{\omega_3}^* v_{\omega_4}^*. \tag{36}$$

All these important properties are due to the fact that the averaged kernel $\langle F\rangle(\omega_1,\omega_2;\omega_3,\omega_4)$, obtained by integrating expression (14) over the amplifier spacing z_a, is symmetric with respect to two different types of permutations

$$\omega_1 \leftrightarrow \omega_2, \quad \omega_3 \leftrightarrow \omega_4, \tag{37}$$

$$\langle F\rangle(\omega_3,\omega_4;\omega_1,\omega_2) = \langle F\rangle^*(\omega_1,\omega_2;\omega_3,\omega_4). \tag{38}$$

Due to (38), the interaction term $H_{\rm int}$ can be rewritten in a symmetric way

$$H_{\rm int} = \frac{1}{4}\frac{1}{(2\pi)^2}\int_{-\infty}^{\infty} d\omega_1 d\omega_2 d\omega_3 d\omega_4 \times \delta(\omega_1+\omega_2-\omega_3-\omega_4), \quad (39)$$

$$\left[\langle F\rangle(\omega_1,\omega_2;\omega_3,\omega_4)\, v_{\omega_1} v_{\omega_2} v^*_{\omega_3} v^*_{\omega_4} + \text{c.c.}\right]. \quad (40)$$

The fact that system (31) is Hamiltonian is not surprising because the original, unaveraged system (8) is Hamiltonian as well. Path-averaged solitons of the form $v(z,\omega) = v_0(\omega)\exp(ikz)$ is an extremum of the H for fixed E

$$\delta(H+kE) = 0. \quad (41)$$

The results we presented are valid for any dispersion profile (including the case of large local dispersion variations) over the period $z_{\rm a}$.

Next, we calculate the first order term Y_1 in the averaged system (19)

$$Y_1 + \frac{\partial\phi_2}{\partial z} = \frac{1}{2}\left[\phi_1, \frac{\partial\phi_1}{\partial z}\right] + \left[\phi_1, \langle X\rangle\right]. \quad (42)$$

By averaging (42), we obtain an expression for the first order correction, Y_1,

$$Y_1 = \frac{1}{2}\langle\left[\phi_1, \frac{\partial\phi_1}{\partial z}\right]\rangle + \left[\langle\phi_1\rangle, \langle X\rangle\right]. \quad (43)$$

The only freedom we have in the expression (43) is ϕ_{10} (see (29)). In the case of weak dispersion variations [23], we can easily find ϕ_{10}, and it is possible to set Y_1 to be zero. This is the major practical difference between the more general case of strong dispersion variations and that of modest ones.

If dispersion variations are absent, then $\tilde{d}_1 = 0$, and the kernel F (14) is drastically simplified $F = I(z)$, its dependence on frequency is gone, and, by taking the inverse Fourier transform, we reduce (33) to the NLSE. The primary model system for nonlinear fiber optics, the NLSE, is a partial differential equation due to the simple physical fact [26] that the nonlinear part of the refractive index does not depend on the signal frequency (over reasonable frequency range.) Notice that in the averaged system (33) the effective nonlinear coefficient (i.e. the kernel F) does depend on the frequency due to strong dispersion variations. There are various physical systems which have an inherent frequency-dependent nonlinear response.

We would like to demonstrate that our results contain the known case [23] and, in the limit of modest dispersion variations (2), that one can eliminate the first order correction. We do this in two steps. Note, that due the fact that now $\tilde{d}_1$ is also a small parameter of the order of $z_{\rm a}$, all expressions are changed drastically. First, we need to verify that the first order correction (43) in our limiting case is reduced, in fact, to second order,

$$Y_1 = O(z_{\rm a}^2). \quad (44)$$

When $\tilde{d}_1$ is a small parameter, we can expand F (14) into a Taylor series in $\tilde{d}_1$ keeping only the zeroth and the first orders:

$$F = I\left(1 - \frac{i}{2}\beta\tilde{d}_1 + i\beta\frac{\langle I\tilde{d}_1\rangle}{\langle 2I\rangle}\right) + O(z_{\rm a}^2), \quad (45)$$

$$\langle F\rangle = \langle I\rangle + O(z_{\rm a}^2). \quad (46)$$

From the definition of ϕ_1, it follows that

$$\frac{\partial\phi_1}{\partial z} = i\tilde{I}v^2 v^* + O(z_{\rm a}),$$

$$\phi_1 = \phi_{10} + i\tilde{I}_1 v^2 v^* + O(z_{\rm a}^2), \quad (47)$$

where $d\tilde{I}_1/dz = \tilde{I}, \langle\tilde{I}_1\rangle = 0$. By using expressions (45) - (47) in the definition of Y_1 (43) and taking $\phi_{10} = 0$, we obtain the final result (44). Finally, we should also check that the term, Y_0, does not generate any first order $z_{\rm a}$-input when $\tilde{d}_1$ is a small parameter. This follows directly from (46) and is guaranteed by the special 'tuning' of the kernel (14) performed at the very beginning.

Let us now consider the general case of strong dispersion management when $\tilde{d}_1$ is not a small parameter anymore. We apply a specific two-step dispersion map and calculate ϕ_1 explicitly for this map. Without loss of generality, we choose a symmetric map $\tilde{d}(z) = d$ for $0 \le z \le \frac{1}{2}$ and $\tilde{d}(z) = -d$ for $\frac{1}{2} < z \le 1$. This immediately yields $\tilde{d}_1(z) = dz - \frac{d}{4}$ for $0 \le z \le \frac{1}{2}$ and $\tilde{d}_1(z) = -dz + \frac{3d}{4}$ for $\frac{1}{2} \le z \le 1$. To describe first order corrections to the averaged model, one should calculate the oscillating part of $\phi_1 = \phi_{10} + i\tilde{J}_1$. In order to carry out our calculations explicitly, we consider the special case (the so-called lossless model) $I(z) = const. = I_0$. The equation for $\tilde{J}_1$ is

$$\frac{\partial \tilde{J}_1}{\partial z} = \tilde{J} = J - \langle J \rangle, \quad \langle \tilde{J}_1 \rangle = 0. \tag{48}$$

In order to integrate this explicitly, we once again make use of the Fourier transform of $u(t_i)$ and write

$$J = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} d\omega_1 d\omega_2 \, d\omega_3 \times F(\omega_1, \omega_2, \omega_3, \omega_1 + \omega_2 - \omega_3)\, u_{\omega_1} u_{\omega_2} u^*_{\omega_3} \times \exp(-it(\omega_1 + \omega_2 - \omega_3)).$$

We shall look only at the parts depending on z. Since $I_0 = \langle I \rangle$ and $\langle \tilde{d}_1 \rangle = 0$, we find that

$$F(\omega_1, \omega_2, \omega_3, \omega_1 + \omega_2 - \omega_3) = I_0 \exp\left(iB\tilde{d}_1\right),$$
$$B = -(\omega_3 - \omega_1)(\omega_3 - \omega_2). \tag{49}$$

Let us determine the function $F_{\tilde{1}}$ with the properties $\partial F_{\tilde{1}} / \partial z = \tilde{F} = F - \langle F \rangle$ and $\langle F_{\tilde{1}} \rangle = 0$. Once $F_{\tilde{1}}$ is found, we obtain $\tilde{J}_1$ immediately from

$$\tilde{J}_1 = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} d\omega_1 d\omega_2 d\omega_3 \, F_{\tilde{1}} \, u_{\omega_1} u_{\omega_2} u^*_{\omega_3} \times \exp(-it(\omega_1 + \omega_2 - \omega_3)). \tag{50}$$

Since $\tilde{d}_1$ depends linearly of z, we can calculate all the integrations analytically. The results are:

$$\frac{\langle F \rangle}{I_0} = \langle e^{iB\tilde{d}_1} \rangle = 2\frac{e^{i\Theta} - e^{-i\Theta}}{iBd} = \frac{\sin\Theta}{\Theta}, \quad \Theta = \frac{Bd}{4},$$
$$\frac{\tilde{F}}{I_0} = \exp[i\,\Theta\,(4z - 1)] - \frac{\sin\Theta}{\Theta}, \quad z \le \tfrac{1}{2},$$
$$\frac{\tilde{F}}{I_0} = \exp[-i\,\Theta\,(4z - 3)] - \frac{\sin\Theta}{\Theta}, \quad z \ge \tfrac{1}{2}. \tag{51}$$

For $F_{\tilde{1}}$ we obtain

$$\frac{F_{\tilde{1}}}{I_0} = \frac{\exp[i\,\Theta\,(4z - 1)]}{4\,i\,\Theta} - \frac{\sin\Theta}{\Theta} z - \frac{\exp[-i\,\Theta\,]}{4\,i\,\Theta},$$
$$z \le \tfrac{1}{2}, \tag{52}$$
$$\frac{\tilde{F}_1}{I_0} = -\frac{\exp[-i\,\Theta\,(4z - 3)]}{4\,i\,\Theta} - \frac{\sin\Theta}{\Theta} z + \frac{2\exp[i\,\Theta\,] - \exp[-i\,\Theta\,]}{4\,i\,\Theta},$$
$$z \ge l\tfrac{1}{2}. \tag{53}$$

These formulas contain first-order corrections to the path-averaged model. Note that in contrast to the case of a weak dispersion management discussed above, we are now able to recognize that $F_{\tilde{1}}$ is zero at $z = \frac{1}{2}$ and therefore $\tilde{J}_1$ is zero also.

In summary, the equation describing in the leading-order slow evolution of $v_\omega(z)$ is

$$i\frac{\partial v_\omega(z)}{\partial z} = \omega^2 \frac{\langle d \rangle}{2} v_\omega - \frac{1}{(2\pi)^2} \times \int_{-\infty}^{+\infty} d\omega_1 d\omega_2 \, \langle F \rangle \, v_{\omega_1} v_{\omega_2} v^*_{\omega_1 + \omega_2 - \omega}.$$

The solution of these integro-differential equations permits to calculate the function $v_\omega(z)$ giving a zero-order path-averaged description for the dispersion-managed pulse. The physical interpretation of this averaging is rather transparent: we separate the rapid quasi-linear phase oscillations induced by large variations of the dispersion from the slow evolution of the pulse shape caused by nonlinearity and average dispersion. For the lossless two-step map, the kernel has a rather simple expression: $\langle F \rangle = I_0 \frac{\sin\Theta}{\Theta}$, with $4\Theta = -(\omega - \omega_1)(\omega - \omega_2)\, d$ [13,16,17]. Using this procedure, we obtain the first-order correction to the field $v_\omega(z)$

$$u_\omega = v_\omega + \frac{i}{(2\pi)^2} \times \int_{-\infty}^{+\infty} d\omega_1 d\omega_2 \times \tilde{F}_1(\omega_1, \omega_2, \omega_1 + \omega_2 - \omega, \omega)\, v_{\omega_1} v_{\omega_2} v^*_{\omega_1 + \omega_2 - \omega}, \tag{54}$$

where $\tilde{F}_1$ was derived above for the case of a two-step map.

In conclusion, we have developed Lie-transform averaging for the NLSE with varying periodic coefficients. Using the Lie-transform, we have derived the path-averaged equation and the first-order corrections to the averaged model in the spectral domain for the case of arbitrary (including large) local dispersion.

Acknowledgements

We are grateful to Sergei Burtsev for many critical suggestions and essential discussions. This work was supported in part (I.G.) by DOE Contract W-7-405-ENG-36 and by the DOE Program in Applied Mathematical Sciences, KJ-01-01.

References

[1] L.F. Mollenauer, S.G. Evangelides Jr., H.A. Haus, IEEE J. Lightwave Technol. 9 (1994) 194–201.
[2] A. Hasegawa, Y. Kodama, Opt. Lett. 15 (1990) 1443–1445.
[3] A. Hasegawa, Y. Kodama, Phys. Rev. Lett. 66 (1991) 161–164.
[4] K.J. Blow, N.J. Doran, IEEE Photon Technol. Lett. 3 (1991) 369–371.
[5] C. Lin, H. Kogelnik, L.G. Cohen, Opt. Lett. 5 (1980) 476.
[6] D.Le. Guen, S. Del Burgo, M.L. Moulinard, D. Grot, M. Henry, F. Favre, T. Georges, PD4-1, Postdeadline Papers, OFC, 1999.
[7] M. Nakazawa, H. Kubota, Electron. Lett. 31 (1995) 216–217.
[8] N. Smith, F.M. Knox, N.J. Doran, K.J. Blow, I. Bennion, Electron. Lett. 32 (1996) 54–55.
[9] I. Gabitov, S.K. Turitsyn, Opt. Lett. 21 (1996) 327–330.
[10] I. Gabitov, S.K. Turitsyn, JETP Lett. 63 (1996) 861–865.
[11] A. Hasegawa, Y. Kodama, A. Maruta, Opt. Fiber Technol. 3 (1997) 197–213.
[12] T.-S. Yang, W.L. Kath, Opt. Lett. 22 (1997) 985.
[13] M.J. Ablowitz, G. Biondini, Opt. Lett. 23 (1998) 1668.
[14] Y. Kodama, Physica D 123 (1998) 255.
[15] S.K. Turitsyn, V.K. Mezentsev, JETP Lett. 67 (1998) 640.
[16] I. Gabitov, E.G. Shapiro, S.K. Turitsyn, Opt. Commun. 134 (1996) 317.
[17] I. Gabitov, E.G. Shapiro, S.K. Turitsyn, Phys. Rev. E 55 (1997) 3624.
[18] T. Lakoba, D.J. Kaup, Electron. Lett. 34 (1998) 1124.
[19] V. Talanov, JETP Lett 11 (1970) 199.
[20] S. Kumar, A. Hasegawa, Opt. Lett. 22 (1997) 372.
[21] I. Gabitov, T. Schafer, S.K. Turitsyn, Preprint LAUR-99-4753, Los Alamos National Laboratory, 8 August 1999.
[22] J.A. Sanders, F. Verhulst, Averaging Methods in Nonlinear Dynamical Systems, 1985, Springer-Verlag, p. 21.
[23] A. Hasegawa, Y. Kodama, Solitons in Optical Communication, 1995, Oxford Univ. Press, p. 121.
[24] Y. Kodama, Phys. Lett. 112A (1985) 193.
[25] P.J. Olver, Applications of Lie Groups to Differential Equations (1986), Springer-Verlag.
[26] G.P. Agrawal, Nonlinear Fiber Optics, Acad. Press, 1994, 2nd edition.