The multiple-scale averaging and dynamics of dispersion-managed optical solitons

TIAN-SHIANG YANG[1], WILLIAM L. KATH[1] and SERGEI K. TURITSYN[2]
[1]*Department of Engineering Sciences and Applied Mathematics, McCormick School of Engineering, Northwestern University, 2145 Sheridan Road, Evanston, Illinois 60208–3125, U.S.A.*
[2]*Division of Electronic Engineering and Computer Science, Aston University, Birmingham B4 7ET, U.K.*

Abstract. Multiple-scale averaging is applied to the nonlinear Schrödinger equation with rapidly varying coefficients, the results are used to analyze pulse propagation in an optical fiber when a periodic dispersion map is employed. The effects of fiber loss and repeated amplification are taken into account by use of a coordinate transformation to relate the pulse dynamics in lossy fibers to that in equivalent lossless fibers. Second-order averaging leads to a general evolution equation that is applicable to both return-to-zero (soliton) and non-return-to-zero encoding schemes. The resulting equation is then applied to the specific case of solitons, and an asymptotic theory for the pulse dynamics is developed. Based upon the theory, a simple and effective design of two-step dispersion maps that are advantageous for wavelength-division-multiplexed soliton transmission is proposed. The use of these specifically designed dispersion maps allows simultaneous minimization of dispersive radiation in several different channels.

Key words: fiber optics, nonlinear Schrödinger equation, multiple scales, dispersion, management, solitons

1. Introduction

Optical solitons result when the nonlinear dependence of the index of refraction upon intensity balances linear chromatic dispersion ([1, Chapter 2] or [2, Chapter 3]). Under ideal conditions, soliton pulses can propagate for long distances in an optical fiber without substantial distortion. Under realistic conditions, solitons have been transmitted single-channel over a distance greater than 10,000 kilometers at a rate of 10 gigabits per second (Gb/s) or more [3]. When combined with wavelength-division-multiplexing (WDM), where N channels are transmitted simultaneously, soliton transmission with $N \times 10\,\text{Gb/s}$ aggregate bit rates over transoceanic distances have been achieved [4]. Generally, soliton transmission allows higher bit rates per channel (as high as 40–100 Gb/s over somewhat shorter distances) in comparison with other schemes, such as non-return-to-zero (NRZ) transmission [5].

Although these data rates are already quite high, the never-ending demand for increasingly higher-speed communication systems requires that the performance of fiber-optic transmission systems be substantially improved. Dispersion management is one method that has been demonstrated to do so [6–10]. The idea of dispersion management is to concatenate fibers of both normal and anomalous dispersion [2, Chapter 7] to form a transmission line having both a high local group-velocity dispersion (GVD) and a low path-averaged GVD. This is beneficial since high local dispersion significantly reduces the efficiency of four-wave mixing [7, 10] and decreases both the modulational instability gain and bandwidth [11, 12]. When used with

soliton systems, lowering the average dispersion also reduces the Gordon–Haus timing jitter [9, 13, 14].

In addition, dispersion management has been found to enhance the soliton energy [15, 16]; this further reduces the timing jitter below the amount that would be obtained in a system with constant dispersion equal to the path-averaged value. The power enhancement of dispersion-managed solitons can be easily understood on physical grounds. Within each dispersion-map period, a pulse undergoes significant broadening and re-compression due to the local dispersion imbalance. As the pulse broadens, of course, the peak power falls and nonlinear effects weaken. Compared to the case of constant dispersion, therefore, higher launch power (and hence higher pulse energy) is needed in order for nonlinearity to balance the net dispersion over each map period. For this reason, dispersion-managed solitons are also referred to as enhanced-power solitons.

Wavelength-division-multiplexed (WDM), *i.e.*, multi-channel, transmission [5, Chapter 7] of dispersion-managed solitons, therefore, is now believed to be a promising way to realize ultra-high capacity optical communication systems. Considerable work has thus been devoted to studying the dynamics of dispersion-managed solitons for various system designs and to optimizing dispersion management for such systems.

Optimization of dispersion-managed soliton systems involves many practical constraints and therefore is a rather complicated issue. Among other factors, however, it is important to diminish the shedding of energy from the input pulse into a dispersive pedestal, which can be achieved by launching properly shaped and chirped pulses with optimum power into the fiber [17]. In practice, typical optical sources generate unchirped pulses and input pulse chirping is realized by use of an additional piece of fiber preceding the line edge. It is therefore necessary to identify points in each dispersion-map period where pulses are unchirped, and use one such location as the launch point. The partial map period preceding the first complete map period then plays the role of the prechirping fiber.

Here we employ multiple-scale averaging to analyze the dynamics of dispersion-managed solitons. Variational approximations have also been used to accomplish this [18]. Here, the goals are to calculate the enhanced pulse energy and to locate the optimal (chirp-free) launch points in each dispersion-map period. In Section 2, we formulate the problem by modeling pulse propagation in a dispersion-managed optical fiber with loss and gain using a perturbed nonlinear Schrödinger (NLS) equation. Via a coordinate transformation, it is shown that for every lossy system with a finite number of amplifiers in each dispersion-map period to compensate fiber loss, there is an equivalent lossless system. Therefore, it suffices to develop a general theory for lossless systems. The effects of different amplifier arrangements on the pulse dynamics can then be studied simply by adapting the results.

In practice, the dispersion-map period is often shorter than the dispersion length (based upon the path-averaged GVD) of the fiber. Taking advantage of the length scale disparity, in Section 3 we use second-order averaging and a near-identity transformation to derive an evolution equation valid for general periodic dispersion maps and applicable to both return-to-zero (RZ, *e.g.*, soliton) and NRZ encoding schemes. The equation is then applied to the case of dispersion-managed solitons in Section 4, and analytic expressions for the evolution of various pulse characteristics are obtained.

In Section 4, based upon the asymptotic results, we discuss the effects of two-step dispersion maps on the pulse dynamics. The case of lossless fibers is analyzed in detail. As expected [15], the optimal (chirp-free) launch points are simply the fiber-segment midpoints when fiber

Table 1. Acronyms used in this paper

GVD	group-velocity dispersion
NLS	nonlinear Schrödinger
NRZ	non-return-to-zero
RZ	return-to-zero
WDM	wavelength-division multiplexing

loss is negligible. The prediction of the asymptotic theory for the power enhancement also compares favorably with an empirical correlation [15, 16].

When fiber loss is significant, however, the optimal launch point locations generally depend on the fiber-segments' dispersions in a complicated manner. An important factor affecting the performance of dispersion-managed WDM systems is the variation of the chromatic dispersion with wavelength (*i.e.*, the dispersion slope) [1, Section 1.2.3], which then necessitates prechirping each channel individually to accomodate the different dispersion values. In Section 5 we show that by choosing the relative lengths of the two fiber segments properly, there are optimal launch points whose locations relative to the amplifier are independent of the fibers' dispersions. This is a remarkable finding as it implies that optimization of dispersion-managed WDM systems can be realized simultaneously in different channels.

2. Perturbed nonlinear Schrödinger equation

Pulse propagation in a dispersion-managed optical fiber with an amplifier deployed in each dispersion-map period to compensate fiber loss is governed by the perturbed NLS equation [2, Chapter 7] (see [19, pp. 68–75] or [20, pp. 32–38] for general descriptions of the unperturbed NLS equation)

$$\mathrm{i}\frac{\partial v}{\partial z} - \tfrac{1}{2}\sigma\left(\frac{z}{\varepsilon}\right)\frac{\partial^2 v}{\partial t^2} + |v|^2 v$$

$$= -\mathrm{i}\left[\frac{\Lambda}{\varepsilon} - (\sqrt{G} - 1)\sum_{n=0}^{N-1}\delta(z - \varepsilon\zeta_A - \varepsilon n)\right]v. \tag{1}$$

We have scaled Equation (1) using a characteristic time τ_0 (to be identified later), the dispersion length based on the path-averaged GVD $L_D = \tau_0^2/|\langle\ddot{\beta}\rangle|$, and the characteristic power $P_0 = (\gamma L_D)^{-1}$, where γ is the fiber's nonlinear coefficient. The dimensionless GVD coefficient is then $\sigma = \ddot{\beta}/|\langle\ddot{\beta}\rangle|$ and is periodic in z with period ε, *i.e.*, $\sigma(z/\varepsilon) = \sigma(z/\varepsilon + 1)$. Here ε is the ratio of the dispersion-map period to the dispersion length, $\varepsilon = l/L_\mathrm{D}$. The dimensionless loss coefficient is $\Lambda = \alpha l/2$, where α is the dimensional loss, and the amplifier gain is $G = \exp(2\Lambda)$. The delta-function terms in (1) are compact representations of the jump conditions due to the N amplifiers deployed at $z = \varepsilon(\zeta_A + n), n = 0, 1, \ldots, N - 1$, with the understanding that the delta functions sample v on the left, namely

$$\int_{\varepsilon\zeta_A^-}^{\varepsilon\zeta_A^+}\delta(z - \varepsilon\zeta_A)v\,\mathrm{d}z \equiv v(\varepsilon\zeta_A^-, t).$$

The linear loss and gain term in (1) can be readily eliminated [21] by the transformation

$$v(z, t) = \tilde{a}(\zeta) u(z, t),\tag{2}$$

where $\zeta = z/\varepsilon$ and $\tilde{a}(\zeta)$ satisfies

$$\frac{d\tilde{a}}{d\zeta} + \left[\Lambda - (\sqrt{G} - 1) \sum_{n=0}^{N-1} \delta(\zeta - \zeta_A - n) \right] \tilde{a} = 0.$$

Equation (1) then becomes

$$i\frac{\partial u}{\partial z} - \tfrac{1}{2}\sigma(\zeta)\frac{\partial^2 u}{\partial t^2} + g(\zeta)|u|^2 u = 0,\tag{3}$$

where $g(\zeta) = \tilde{a}(\zeta)^2 = C \exp\{-2\Lambda[\zeta - \zeta_A + 1 - H(\zeta - \zeta_A)]\}(0 \leqslant \zeta \leqslant 1)$ and $g(\zeta) = g(\zeta+1)$. Here $H(\cdot)$ is the Heaviside step function and $C = G \log G/(G-1)$ if $g(\zeta)$ is normalized such that

$$\langle g \rangle = \int_0^1 g(\zeta)\, d\zeta = 1.$$

Note that in a lossless fiber $\Lambda = 0$ and hence $g(\zeta) \equiv 1$.

Since the pulse energy variation due to fiber loss and repeated amplification has been taken care of by the transformation (2), the equation for u now conserves 'energy', *i.e.*, $\int_{-\infty}^{\infty} |u|^2\, dt$ is a constant. Note the periodic coefficient of the self-phase modulation (nonlinear) term in (3), $g(\zeta)$, is a reflection of the true pulse energy variation (*i.e.*, in v). Note also that the special case where the amplifier spacing equals the dispersion-map period is used above for illustrative purpose. It is clear that the transformation (2) can also be applied to general cases where there are a finite number of amplifiers in each map period. The gains of the amplifiers may differ as long as they cooperatively compensate the fiber loss in each period. Different amplification schemes simply lead to different functional forms for $g(\zeta)$ appearing in (3).

It is interesting that for any fiber with periodic loss and amplification there is an equivalent lossless fiber. This can be seen by introducing the coordinate transformation

$$\xi = \xi(\zeta) = \int_0^\zeta g(s)\, ds.\tag{4}$$

Equation (3) then becomes

$$i\frac{\partial u}{\partial x} - \tfrac{1}{2}\hat{\sigma}(\xi)\frac{\partial^2 u}{\partial t^2} + |u|^2 u = 0,\tag{5}$$

where $\hat{\sigma} = \sigma/g$, $\hat{\sigma}(\xi) = \hat{\sigma}(\xi + 1)$, and $x = \varepsilon\xi$. The transformation (4) is one-to-one since g is always positive. Comparing (5) with (3), we clearly see that (5) now has the same form as the equation for pulse evolution in lossless fibers with dispersion management. The lossless Equation (5) is therefore directly relevant to the study of pulse evolution in systems with loss and gain. Furthermore, no restrictions have been placed on the pulse profile or the magnitude of ε, so (5) is applicable to both NRZ and RZ encoding schemes.

Equation (5) explains in a simple manner the benefits of using fibers with an exponentially decreasing dispersion to transmit solitons when fiber loss is significant, *e.g.*, when the amplifier spacing is large (which reduces system cost and complexity). Soliton transmission is the

result of a balance between nonlinearity and dispersion; when the pulse amplitude decreases due to loss, the fiber's dispersion must also decrease if that balance is to be maintained. In view of (5), ideal soliton transmission requires a constant (negative) effective GVD parameter $\hat{\sigma}$. Since the self-phase modulation effective strength g decreases exponentially, the two effects will balance if the fiber's GVD parameter σ also decreases exponentially.

Fibers with exponentially decreasing dispersion can be manufactured [22, 23]. Since the construction of a soliton transmission system employing such special fiber would be prohibitively expensive, however, a reasonable alternative is to concatenate a series of constant-dispersion fibers (with piecewise decreasing dispersion, of course) to approximate a fiber having exponentially decreasing dispersion. Clearly, the dispersion values of all the fiber segments in this dispersion management scheme have the same sign (*i.e.*, anomalous dispersion). We will see shortly, however, that a number of advantages can be realized if both signs are allowed. The asymptotic theory developed below, of course, is valid for general $g(\zeta)$ and $\sigma(\zeta)$ and therefore applies to both schemes.

3. Multiple-scale averaging

In practical dispersion-managed systems, the dispersion-map period is often shorter than the fiber's dispersion length. It is therefore useful to analyze the asymptotic limit where $0 < \varepsilon \ll 1$ by exploiting the small parameter. To do so, we expand the solution in a power series in ε,

$$u = U + \varepsilon u_1 + \varepsilon^2 u_2 + \cdots, \tag{6}$$

where the average of u_k ($k = 1, 2, \ldots$) with respect to ξ satisfies

$$\langle u_k \rangle_\xi = \int_0^1 u_k \, \mathrm{d}\xi = 0,$$

so that U represents the average of u with respect to ξ over the dispersion map period. Since $\langle u \rangle_\xi = \int_0^1 ug(\zeta) \, \mathrm{d}\zeta = \langle gu \rangle$, U may also be interpreted as the average of u with respect to ζ weighted by the self-phase modulation strength g.

We take Equation (6) to be a multiple-scale expansion [24, Chapter 3], assuming that the fluctuating u_k depend upon U and the 'fast' variable $\xi = x/\varepsilon$, while the average U depends upon 'slow' variables $x_k = \varepsilon^k x$, $k = 0, 1, 2, \ldots$. This means that

$$\frac{\partial U}{\partial x} = \frac{\partial U}{\partial x_0} + \varepsilon \frac{\partial U}{\partial x_1} + \varepsilon^2 \frac{\partial U}{\partial x_2} + \cdots. \tag{7}$$

We can then separate fast oscillations produced by the dispersion map from longer-term effects.

Upon substitution of (6) and (7) in (5), we find at leading order, ε^0, that

$$\mathrm{i}\frac{\partial u_1}{\partial \xi} + \mathrm{i}\frac{\partial U}{\partial x_0} - \tfrac{1}{2}\hat{\sigma}(\xi)\frac{\partial^2 U}{\partial t^2} + |U|^2 U = 0. \tag{8}$$

In order to avoid terms which grow linearly with ξ in u_1, *i.e.*, secular terms [24, p. 107], we must have

$$\mathrm{i}\frac{\partial U}{\partial x_0} = \tfrac{1}{2}\langle \sigma \rangle \frac{\partial^2 U}{\partial t^2} - |U|^2 U \tag{9}$$

and

$$u_1 = -\frac{\mathrm{i}}{2}[\hat{\mu}_1(\xi) - \langle\hat{\mu}_1\rangle_\xi]\frac{\partial^2 U}{\partial t^2}. \tag{10}$$

Here $\langle\sigma\rangle = \langle\hat{\sigma}\rangle_\xi$ is the average GVD coefficient, and the fluctuations about the mean depend on $\hat{\mu}_1(\xi) = \int_0^\xi[\hat{\sigma}(s) - \langle\sigma\rangle]\,\mathrm{d}s = \mu_1(\zeta) - \langle\sigma\rangle\nu_1(\zeta)$, with $\mu_1(\zeta) = \int_0^\zeta[\sigma(s) - \langle\sigma\rangle]\,\mathrm{d}s$ and $\nu_1(\zeta) = \int_0^\zeta[g(s) - 1]\,\mathrm{d}s$.

At the next order, ε^1, we find $\partial U/\partial x_1 = 0$ and

$$\begin{aligned}
u_2 &= -[\hat{\mu}_2(\xi) - \langle\hat{\mu}_2\rangle_\xi](U^2 U_{tt}^* + 2U|U_t|^2 + U_t^2 U^*) \\
&\quad -\tfrac{1}{8}\{[\hat{\mu}_1(\xi) - \langle\hat{\mu}_1\rangle]_\xi^2 - 2\mathcal{M}\}U_{tttt},
\end{aligned} \tag{11}$$

where $\hat{\mu}_2(\xi) = \int_0^\xi[\hat{\mu}_1(s) - \langle\hat{\mu}_1\rangle_\xi]\,\mathrm{d}s$, and $\mathcal{M} = \frac{1}{2}(\langle\hat{\mu}_1^2\rangle_\xi - \langle\hat{\mu}_1\rangle_\xi^2)$ is the variance of $\hat{\mu}_1$. For lossless fibers, $\Lambda = 0$ and $\mathcal{M} = \frac{1}{2}(\langle\mu_1^2\rangle - \langle\mu_1\rangle^2)$ is the map's dispersion variance, *i.e.*, the variance of the local net accumulated dispersion surplus or deficit in the map relative to the path-averaged case. When the pulse amplitude variations due to fiber loss and amplification are significant, of course, the parameter $\mathcal{M}$ depends upon the dispersion map, the pulse energy variation, and their interactions.

Finally, at second order we find an expression for u_3 and also

$$\begin{aligned}
\mathrm{i}\frac{\partial U}{\partial x_2} &= \mathcal{M}[U^2 U_{tttt}^* + 6U U_t U_{ttt}^* + 2(|U|^2)_t U_{ttt} + 5U|U_{tt}|^2 \\
&\quad + 7U_t^2 U_{tt}^* + 10|U_t|^2 U_{tt} + \tfrac{5}{2}U_{tt}^2 U^*].
\end{aligned} \tag{12}$$

Substituting (9), $\partial U/\partial x_1 = 0$, and (12) in (7), we then obtain an averaged evolution equation for U, the mean of u, correct to $O(\varepsilon^2)$,

$$\begin{aligned}
&\mathrm{i}U_x - \tfrac{1}{2}\langle\sigma\rangle U_{tt} + |U|^2 U \\
&\quad = \varepsilon^2 \mathcal{M}[U^2 U_{tttt}^* + 6U U_t U_{ttt}^* + 2(|U|^2)_t U_{ttt} + 5U|U_{tt}|^2 \\
&\qquad + 7U_t^2 U_{tt}^* + 10|U_t|^2 U_{tt} + \tfrac{5}{2}U_{tt}^2 U^*] + \cdots.
\end{aligned} \tag{13}$$

Equation (13) is the same as what one obtains by performing 'guiding center averaging' using Lie transforms [25, 43, 44]. Not all terms on the right-hand side of (13) are important for determining the long-term behavior of U, however. Some of the perturbing terms produce relatively trivial effects upon the pulse evolution. For example, if there were a perturbing term of the form $\varepsilon^2|U|^2 U$, then it is clear that the perturbation could be eliminated merely by grouping this with the self-phase modulation term on the left-hand side and then renormalizing the pulse amplitude.

It is not possible to eliminate all types of perturbations, of course. A general method for determining which are significant and which are not is available [26, 27]. This method involves examining general classes of renormalizations or corrections to the solution and identifying which resulting perturbations appear on the right-hand side of (13). The details of the application of the method to (13) are rather involved. We found, however, a simple interpretation,

described in the Appendix, that elucidates the essence of the method. In any event, it can be verified by direct substitution that the near-identity transformation

$$q = U + \varepsilon^2 \mathcal{M} \left\{ \tfrac{1}{4} U_{tttt} + \frac{1}{\langle \sigma \rangle} [U_t(|U|^2)_t + U(UU_t^*)_t] \right\} + \cdots \tag{14}$$

converts (13) to

$$i\frac{\partial q}{\partial x} - \tfrac{1}{2}\langle \sigma \rangle \frac{\partial^2 q}{\partial t^2} + |q|^2 q$$

$$= \varepsilon^2 \frac{\mathcal{M}}{\langle \sigma \rangle} \left\{ |q|^2 q \frac{\partial^2}{\partial t^2}(|q|^2) + \tfrac{1}{2} q \frac{\partial^2}{\partial t^2}(|q|^4) \right\} + \cdots. \tag{15}$$

As is shown in the Appendix, the transformation (14) is unique if one requires it to eliminate all cubic perturbations from the right-hand-side of (15). Of course, it is straightforward to revert (14) to provide a similar transformation from q back to U.

Note that to this point (15) is completely general, and thus applicable to both NRZ and RZ encoding schemes. To be valid, however, it is clear that we must require the path-averaged dispersion, $\langle \sigma \rangle$, to be nonzero. This implies that significant differences in the propagation behavior should be expected when $\langle \sigma \rangle = 0$. Note also that the perturbing terms in the transformed Equation (15) are now independent of the derivatives of the solution's phase, implying that the effects of the pulse phase variation have already been absorbed into the transformation (14). As a result, since the pulse phase dynamics is usually more complicated than that of the pulse power, the transformed Equation (15) is advantageous both for further analytical and numerical treatment.

4. Analysis of enhanced power solitons

For the remainder of the discussion we will apply (15) to the case of dispersion-managed solitons. We will assume that $\langle \ddot{\beta} \rangle$ is negative, corresponding to anomalous path-averaged dispersion, as is necessary for solitons when the strength of the dispersion map is either weak or moderately strong. Note that for sufficiently strong dispersion maps it is possible to have dispersion-managed solitons when the path-averaged GVD is zero or even positive (*i.e.*, normal) [28–37]. It is not possible to analyze such strong maps with the present theory, however, since the averaging method produces results that must be valid in the limit of weak dispersion variation, and the behavior observed for strong maps does not persist in this limit. Nevertheless, a substantial advantage of the averaging method is that closed-form results are obtained that allow simple and direct exploration of various dispersion map designs; unfortunately, such considerations for strong dispersion maps are much more difficult to analyze.

For the analysis, we first normalize the path-averaged dispersion so that $\langle \sigma \rangle = -1$. Using a regular perturbation expansion, we find that (15) has the solution

$$q = e^{ix/2}\{\text{sech}\, t + \tfrac{4}{3}\varepsilon^2 \mathcal{M}(2\,\text{sech}\, t - \text{sech}^3 t - \text{sech}^5 t) + \cdots\}.$$

Transforming q back to u using (14) and (6), we find $u = A\exp(i\phi)$, where

$$\phi = \tfrac{1}{2}x - \tfrac{1}{2}\varepsilon[\hat{\mu}_1(\xi) - \langle\hat{\mu}_1\rangle_\xi](1 - 2\,\text{sech}^2 t) + \cdots \tag{16}$$

and

$$
\begin{aligned}
A ={} & \operatorname{sech} t\{1 + \tfrac{1}{3}\varepsilon^2 \mathcal{M}(8 + 8\operatorname{sech}^2 t - 19\operatorname{sech}^4 t) \\
& + \varepsilon^2(\tfrac{1}{2}[\hat{\mu}_1(\xi) - \langle\hat{\mu}_1\rangle_\xi]^2 - [\hat{\mu}_2(\xi) - \langle\hat{\mu}_2\rangle_\xi]) \\
& \times (4\operatorname{sech}^2 t - 5\operatorname{sech}^4 t) + \cdots\}.
\end{aligned}
\tag{17}
$$

The asymptotic results (16) and (17) indicate that the pulse's profile and phase deviate from the results for a hyperbolic secant-shaped soliton in a uniform fiber, as observed numerically [15, 16]. Note, however, that the pulse shape variations are $O(\varepsilon^2)$, where the $O(\varepsilon)$ correction is entirely a phase variation in (16).

One way to reduce this phase variation to a single number is to define the phase chirp by

$$
b = \frac{\int_{-\infty}^{\infty} t\,\Im\{u^* u_t\}\,\mathrm{d}t}{\int_{-\infty}^{\infty} t^2 |u|^2\,\mathrm{d}t}
$$

(if the perturbations are small and the phase varies quadratically in t, the parameter b is simply the coefficient; note that it is not related to one of the conserved quantities of the NLS equation, however, and therefore its value is expected to be more strongly affected by pulse shape). We then find directly from (16) and (17) that

$$
b = -\frac{4}{\pi^2}\varepsilon[\hat{\mu}_1(\xi) - \langle\hat{\mu}_1\rangle_\xi].
\tag{18}
$$

Therefore, the pulse is unchirped at points $\xi = \xi_0$, where $\hat{\mu}_1(\xi_0) = \langle\hat{\mu}_1\rangle_\xi$.

We can also show, using (17), that the minimum pulse width occurs at points where the pulse is unchirped and is therefore transform-limited [1, Section 3.2.2]. At such a point we can thus determine the relationship between the characteristic time τ_0 and the minimum full-width-at-half-maximum (FWHM) pulse width τ_{FWHM} [1, p. 65], which is

$$
\frac{\tau_{\mathrm{FWHM}}}{\tau_0} = 2\log(1 + \sqrt{2}) + \frac{\sqrt{2}}{6}\varepsilon^2(41\mathcal{M} - 21\mathcal{N}) + \cdots,
\tag{19}
$$

where $\mathcal{N} = \hat{\mu}_2(\xi_0) - \langle\hat{\mu}_2\rangle_\xi$. Alternatively, τ_0 is also related to the minimum root-mean-square (RMS) pulse width τ_{RMS} [1, pp. 71–73] by

$$
\frac{\tau_{\mathrm{RMS}}}{\tau_0} = \frac{\pi}{\sqrt{12}}\left[1 + \frac{4}{3\pi^2}\varepsilon^2(11\mathcal{M} - 3\mathcal{N}) + \cdots\right].
\tag{20}
$$

Also, as expected, (17) gives the constant pulse 'energy'

$$
\int_{-\infty}^{\infty} |u|^2\,\mathrm{d}t = 2\left(1 + \frac{32}{15}\varepsilon^2\mathcal{M} + \cdots\right).
\tag{21}
$$

So far, the dispersion map and the amplifier arrangement have not been specified. The asymptotic results above therefore can be specialized to predict the dynamics of dispersion-managed solitons for various system designs and then used to address the optimization issue, as long as the dispersion-map period is shorter than the fiber's dispersion length.

A simple implementation of dispersion maps is to concatenate two fibers, one having anomalous dispersion and the other normal dispersion. This has been demonstrated experimentally to be an effective means for improving the performance of soliton transmission systems [9, 38]. Here we use the asymptotic results obtained above, and compare them with numerical results for the perturbed NLS Equation (1) to discuss the effects of two-step dispersion maps on the dispersion-managed soliton pulse shape and power. The dimensional GVD coefficients of the two fiber segments will be denoted by $\ddot{\beta}_{1,2}$ and their lengths by $l_{1,2}$. The path-averaged GVD then is $\langle\ddot{\beta}\rangle = \zeta_1\ddot{\beta}_1 + \zeta_2\ddot{\beta}_2$, where $\zeta_1 = l_1/l$ and $\zeta_2 = l_2/l = 1 - \zeta_1$.

First we consider the idealized case where fiber loss is negligible, *i.e.*, $\Lambda = 0$, so that $g \equiv 1$ and hence $\xi = \zeta$, $v = u$. It can then be readily deduced that $\hat{\mu}_1 = \langle\hat{\mu}_1\rangle_\xi$ at the midpoints of the fiber segments. In other words, dispersion-managed solitons in lossless fibers are unchirped at the segment midpoints, consistent with the numerical results of Smith *et al.* [15]. An empirical formula for the enhanced pulse energy, denoted by E below, that fits the numerical results is [15]

$$E = -\frac{3 \cdot 52\langle\ddot{\beta}\rangle}{\gamma\,\tau_{\mathrm{FWHM}}}\left[1 + 0{\cdot}7\left(\frac{(\ddot{\beta}_1 - \langle\ddot{\beta}\rangle)l_1 - (\ddot{\beta}_2 - \langle\ddot{\beta}\rangle)l_2}{\tau_{\mathrm{FWHM}}^2}\right)^2\right]. \tag{22}$$

We wish to compare the asymptotic result (21) with this empirical formula.

To do so, the parameters $\mathcal{M}$ and $\mathcal{N}$ must be evaluated. Expression (19) or (20) then indicates that the (locally) minimum pulse widths at two consecutive segment midpoints actually differ slightly in general, since the value of $\mathcal{N}$ depends on ξ_0. To avoid ambiguity, we choose the point $\xi = \xi_0$ to be the midpoint of the segment whose GVD coefficient is $\ddot{\beta}_1$. Also, since $\hat{\mu}_1(0) = \hat{\mu}_2(0) = 0$ by definition, it is convenient to start the dispersion map at the midpoint so that $\langle\hat{\mu}_1\rangle_\xi = \langle\mu_1\rangle = 0$. By direct calculation, it then follows that

$$\mathcal{N} = -\tfrac{1}{48}(\zeta_1 + 2\zeta_2)[\zeta_1(\sigma_1 + 1) - \zeta_2(\sigma_2 + 1)]$$

and

$$\mathcal{M} = \tfrac{1}{96}[\zeta_1(\sigma_1 + 1) - \zeta_2(\sigma_2 + 1)]^2,$$

where $\sigma_1 = \ddot{\beta}_1/|\langle\ddot{\beta}\rangle|$ and $\sigma_2 = \ddot{\beta}_2/|\langle\ddot{\beta}\rangle| = -(1 + \zeta_1\sigma_1)/\zeta_2$ are the dimensionless GVD coefficients of the fiber segments.

Note that interchanging the subscripts '1' and '2' does not affect the expression for $\mathcal{M}$. This reflects the fact that the parameter $\mathcal{M}$ is related to the pulse energy; its value, therefore, is independent of how the segments are labeled and where the starting point of the dispersion map is chosen. Also, using the above expressions for $\mathcal{M}$ and $\mathcal{N}$, it is straightforward for us to determine the segment midpoint where the pulse is truly narrowest by comparing the pulse width that (19) or (20) predicts for two consecutive segment and midpoints.

Combining the expressions for $\mathcal{M}$ and $\mathcal{N}$ with (19) and (21), we then obtain the pulse energy, in terms of dimensional quantities [39],

$$E = -\frac{3{\cdot}52\langle\ddot{\beta}\rangle}{\gamma\,\tau_{\mathrm{FWHM}}}\left[1 + 0{\cdot}766\,S^2 - 0{\cdot}564\frac{l_1 + 2l_2}{\tau_{\mathrm{FWHM}}^2}\langle\ddot{\beta}\rangle S + \cdots\right], \tag{23}$$

where $S = [l_1(\ddot{\beta}_1 - \langle\ddot{\beta}\rangle) - l_2(\ddot{\beta}_2 - \langle\ddot{\beta}\rangle)]/\tau_{\mathrm{FWHM}}^2$ measures the strength of the dispersion map and is a scaled version of the map's dispersion variance ($S^2 = 96\varepsilon^2\mathcal{M}$). The first term is

precisely that discovered via numerical simulations [15], but with a coefficient (0·766) that is slightly greater than that determined numerically (0·7). The second term is basically similar to the first, but it is proportional to the product of the dispersion map strength S and the path-averaged dispersion $\langle \ddot{\beta} \rangle$, rather than S^2. Since a goal of using dispersion maps is to reduce $\langle \ddot{\beta} \rangle$, this second term is expected to be negligible in most cases.

Figure 1 compares the analytic expression for the enhanced pulse energy (23) with the empirical result (22). Here $\langle \ddot{\beta} \rangle = -0 \cdot 1 \, \text{ps}^2/\text{km}$, $\gamma = 2 \cdot 65 \, \text{W}^{-1}\text{km}^{-1}$, $\tau_{\text{FWHM}} = 20 \, \text{ps}$, $l_1 = l_2 = 100 \, \text{km}$, and $\ddot{\beta}_2 - \ddot{\beta}_1$ varies from 0 to $12 \, \text{ps}^2/\text{km}$ (the range over which the empirical best fit was made). The largest difference between the two curves arises from the different coefficients, 0·766 *vs.* 0·7, but some inaccuracy should be anticipated in the analytic result when the perturbations are strong, since it is a 'best fit' only for small $\ddot{\beta}_2 - \ddot{\beta}_1$. Nevertheless, considering that the curves span a relatively wide range, the agreement is much better than one would expect *a priori*.
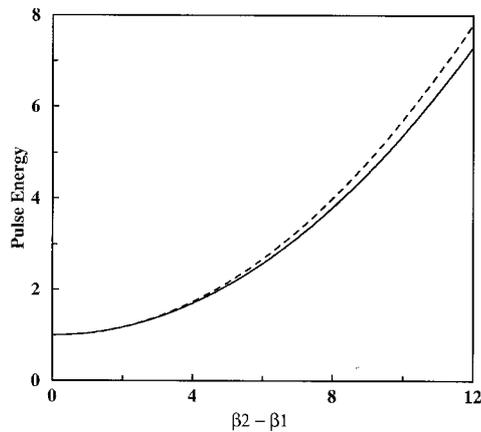
Figure 1. Comparison between the numerical (solid curve) and analytical (dashed curve) results for the soliton power enhancement due to a two-step dispersion map. The parameters are as described in the text, and the pulse energy has been normalized by $-3 \cdot 52 \langle \ddot{\beta} \rangle / (\gamma \tau_{\text{FWHM}})$.
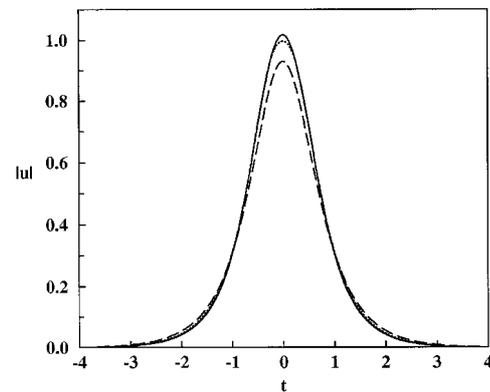
Figure 2. Comparison of pulse shapes obtained with the numerical (solid curve), analytical (dotted curve) and path-averaged NLS (dashed curve) solutions. Here $\ddot{\beta}_2 - \ddot{\beta}_1 = 1 \cdot 82 \, \text{ps}^2/\text{km}$.

In Figure 2 we compare the pulse shapes obtained from the numerical solution, the analytic solution (17), and what one obtains using the path-averaged NLS equation. The parameters are as before, and $\ddot{\beta}_2 - \ddot{\beta}_1 = 1 \cdot 82 \, \text{ps}^2/\text{km}$. The perturbation expansion does a good job of predicting the pulse shape for small dispersion map strengths. For values of $\ddot{\beta}_2 - \ddot{\beta}_1$ above about $2 \cdot 9 \, \text{ps}^2/\text{km}$, however, the perturbation solution becomes double-peaked and is no longer a good pointwise approximation to the solution. This suggests that the pulse energy, being an integrated quantity, is to some extent tolerant of inaccuracies in the pulse shape.

More recent results [40] have observed discrepancies between additional simulations and the empirically determined curve, (22), particularly when the dispersion map strength is strong. In this limit a saturation effect as been observed which limits the growth of the power enhancement factor. We conjecture that these discrepancies may be due in part to difficulties in the measurement of the FWHM pulse width, since the pulse shape can become quite complicated in this limit [41]. It is possible that use of another method to calculate the pulse width (such as

the minimum RMS width) may yield better agreement, but eventually, of course, one should expect all analytic predictions based upon averaging to break down for very strong maps.

5. Optimal launch points for WDM soliton transmission

We now turn to the more realistic case where fiber loss is significant and is compensated by deploying an amplifier at a segment junction per dispersion-map period. As mentioned in Section 1, an objective is to locate the points in each map period where dispersion-managed solitons are unchirped, so that pulses can be launched at one such point to minimize the shedding of dispersive radiation. The locations of these optimal launch points now are affected by fiber loss and repeated power amplification, however, and in general no longer occur at the midpoints of the fiber segments.

Nevertheless, the general results obtained in Section 4 can be readily adapted to study the effects of loss and gain on the locations of the optimal launch points. Here we shall denote the dimensionless GVD coefficient of the fiber segment after (before) the amplifier by σ_1 (σ_2) and the fraction it occupies the map period by ζ_1 (ζ_2), and again take $\langle \sigma \rangle = -1$. According to (18), the optimal launch points occur when $\hat{\mu} = \langle \hat{\mu} \rangle_\xi$. With some straightforward algebra, we find that the distance between the amplifier and a chirp-free point on the fiber segment after the amplifier, ζ_a, can be determined from [42, 45].

$$\sigma_1 + 1 = \frac{2\Lambda G \exp(-2\Lambda \zeta_1 a) + (2\Lambda \zeta_1 a - 1 - \Lambda)(G - 1)}{2\Lambda \zeta_1 a(G - 1) - G + [\exp(2\Lambda \zeta_2) - \zeta_1]/\zeta_2}, \tag{24}$$

where $a = \zeta_a/\zeta_1 (0 < a < 1)$. Similarly, the distance between the amplifier and a chirp-free point on the segment before the amplifier, ζ_b, is given by

$$\sigma_1 + 1 = \frac{2\Lambda \exp(-2\Lambda \zeta_2 a) + (2\Lambda \zeta_2 a - 1 + \Lambda)(G - 1)}{-2\Lambda \zeta_1 a(G - 1) - G + [\exp(2\Lambda \zeta_2) - \zeta_1]/\zeta_2}, \tag{25}$$

where $a = -\zeta_b/\zeta_2 (-1 < a < 0)$.

For given values of Λ, ζ_1, and a, the dimensionless GVD coefficient σ_1 can be calculated from (24) and (25). Several solution curves are plotted in Figure 3, where $\Lambda = 2\cdot763$ (corresponding, for example, to $\alpha = 0\cdot04605\,\mathrm{km}^{-1}$ and $l - 120\,\mathrm{km}$). It appears that for a given dispersion map (characterized by ζ_1 and σ_1), there are in general two optimal launch points. Three or four optimal launch points are also possible for weaker dispersion maps where both fiber segments have anomalous dispersion.

When $\zeta_1 = 0\cdot1762$ or $0\cdot2650$, however, the location of a chirp-free point is independent of the GVD coefficients of the fiber segments. These 'magic' length fractions correspond to special cases where the numerator and the denominator of the right-hand side of (24) or (25) are both zero, and, as can be readily deduced from (24) and (25), only depend on the loss coefficient Λ. We calculate numerically the magic fractions as functions of Λ and plot the results in Figure 4. Note the magic fraction ζ_1 tends to decrease (hence $\zeta_2 = 1 - \zeta_1$ increases) with increasing Λ. In the lossless limit ($\Lambda = 0$), of course, the chirp-free points are simply the midpoints of the fiber segments.

The observation that special dispersion maps exist having GVD-independent optimal launch point locations is of crucial importance for WDM soliton transmission. As noted in Section 1, due to the dispersion slope different channels have different dispersion characteristics and therefore in general need to be prechirped differently. The existence of the special
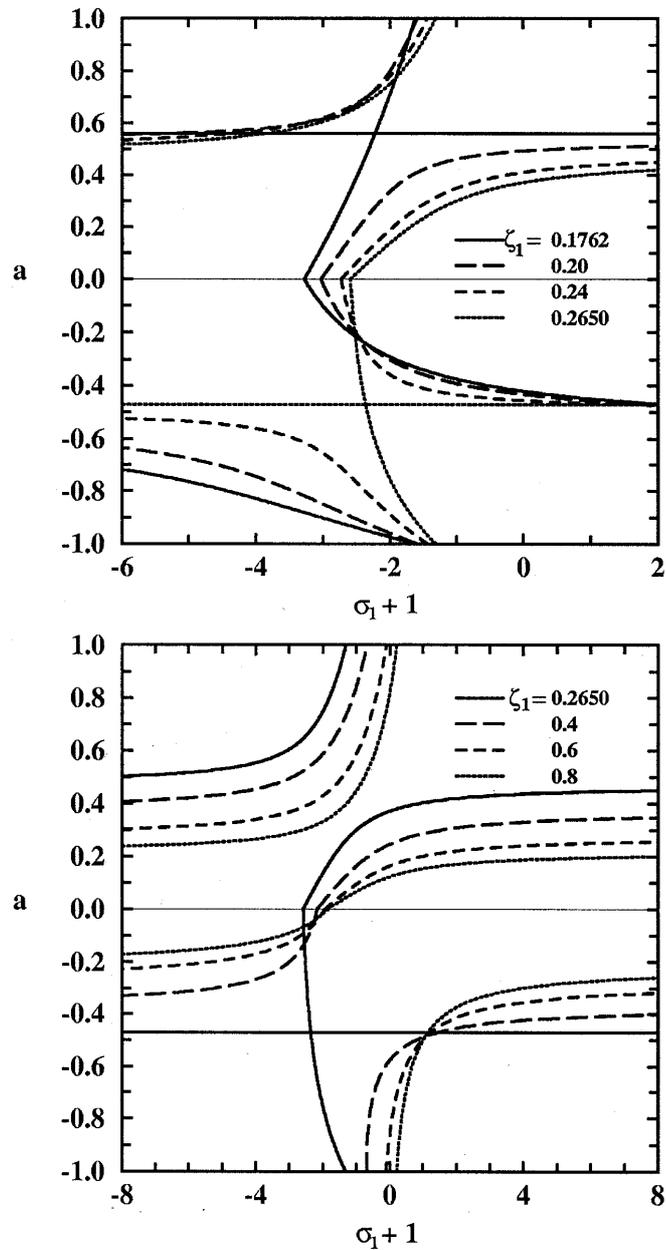
Figure 3. Locations of the chirp-free points within a dispersion map period. The distance between a chirp-free point and the amplifier is normalized by the length of the segment on which the chirp-free point lies and positive (negative) values of the normalized distance *a* correspond to chirp-free points on the segment after (before) the amplifier. Here the loss parameter is $\Lambda = 2 \cdot 763$.
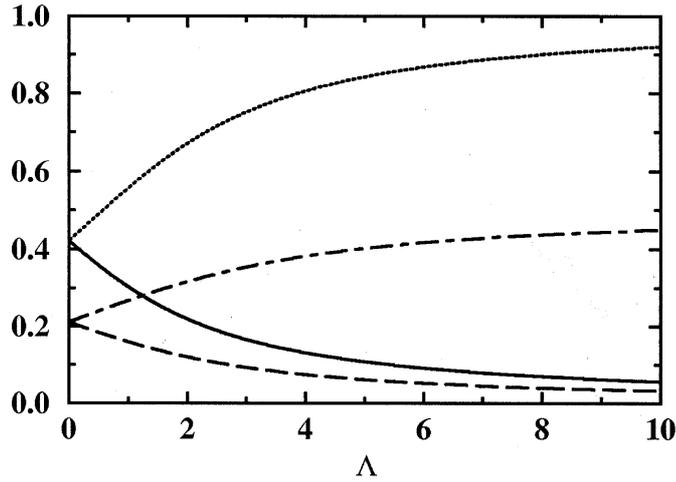
Figure 4. The magic fractions and the locations of the corresponding GVD-independent chirp-free points as functions of the loss coefficient Λ. The fraction ζ_1 (solid line) and the distance ζ_a (dashed line) amplifier, while ζ_2 (dotted line) and ζ_b (dot-dashed line) are plotted when the chirp-free point is on the fiber segment before the amplifier.

dispersion maps, however, implies that optimization of dispersion-managed WDM systems can be realized simultaneously in different channels.

We now compare the predictions of the asymptotic theory with numerical results based on the perturbed NLS Equation (1). We shall take the attenuation constant and the dispersion-map period to be $\alpha = 0.04605 \, \text{km}^{-1}$ (corresponding to a $0.2 \, \text{dB/km}$ power loss) and $l = 120 \, \text{km}$, respectively; the dimensionless loss coefficient is then $\Lambda = 2.763$.

As an example, first consider a low-strength dispersion map (in the sense that $\varepsilon^2 \mathcal{M} \ll 1$) with $\zeta_1 = 0.225$ and $\sigma_1 = -3.33$; accordingly, $\mathcal{M} = 6.29 \times 10^{-4}$. It can be readily checked that this map consists of two anomalous-dispersion segments and therefore is effectively a simple approximation of a fiber having exponentially decreasing dispersion [see the discussion following (5)]. For this map, (24) and (25) predict that there are four optimal launch points corresponding to $a = 0.7099, 0.1442, -0.2587$, and -0.8688, respectively. Figure 5 shows the evolution of the peak pulse amplitude (occurring at $t = 0$) and the chirp parameter b after an unchirped hyperbolic-secant pulse is launched at the optimal point corresponding to $a = 0.1442$, taking $\varepsilon = 0.5$. Both numerical and asymptotic results are plotted in Figure 5 for comparison. Clearly, the asymptotic theory provides a good prediction for this map.

We next confirm numerically the existence of special dispersion maps having GVD-independent optimal launch points. To demonstrate the effectiveness of the theory, for illustrative purposes we choose maps with significantly different dispersion characteristics (in practice, different WDM channels have much closer dispersion parameters). In case I, the dimensional GVD coefficients of the fiber segments are taken to be $\ddot{\beta}_1^{\text{I}} = -20.1$ and $\ddot{\beta}_2^{\text{I}} = 2.76 \, \text{ps}^2/\text{km}$, while in case II $\ddot{\beta}_1^{\text{II}} = -22.7$ and $\ddot{\beta}_2^{\text{II}} = 3.76 \, \text{ps}^2/\text{km}$. Here it is assumed that the amplifier is deployed at the beginning of the anomalous dispersion segment in both cases. For $\Lambda = 2.763$, we find $\zeta_1 = 0.1762$ is a fraction with a GVD-independent optimal launch point corresponding to $a = 0.5593$. The path-averaged GVD coefficients in the two cases are then $\langle \ddot{\beta} \rangle^{\text{I}} = -1.268$ and $\langle \ddot{\beta} \rangle^{\text{II}} = -0.902 \, \text{ps}^2/\text{km}$, respectively. For initially 35 ps FWHM pulses, $\varepsilon^{\text{I}} = 0.486$ and $\varepsilon^{\text{II}} = 0.417$. Note that for strong dispersion maps, to determine the
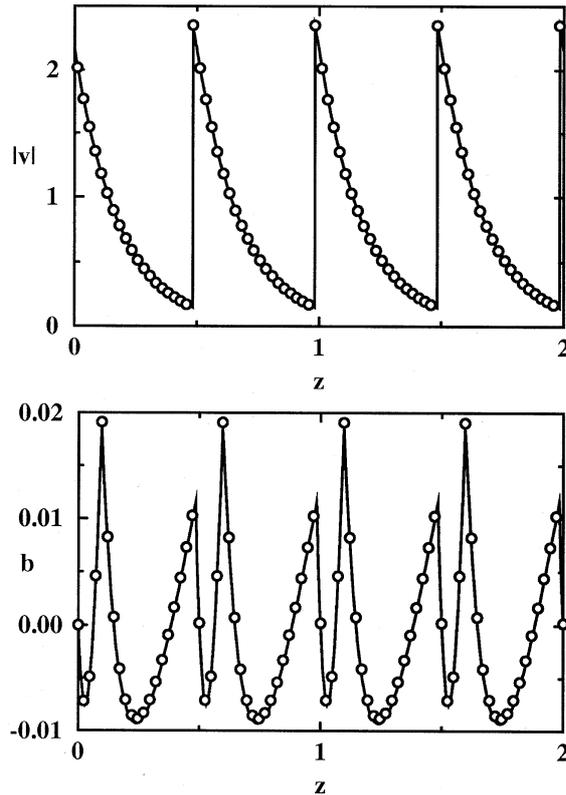
Figure 5. Evolution of the peak pulse amplitude $|v(t = 0)|$ and the chirp parameter b (circles: numerical results; solid lines: asymptotic predictions). The parameters are described in the text.

theoretical value of ε when the FWHM pulse width is given one must simultaneously solve $\varepsilon = l/L_{\mathrm{D}}$, $L_{\mathrm{D}} = \tau_0^2/|\langle\ddot{\beta}\rangle|$, and Equation (19).

Figure 6 shows the evolution of the phase chirp in both cases after unchirped hyperbolic-secant pulses are launched at the optimal launch point corresponding to $a = 0.5593$ with power enhancements of 11% and 21%, respectively. It is clearly seen that the asymptotic theory provides good predictions for the locations of the chirp-free points in both cases (even though pointwise asymptotic estimates for the chirp parameter b differ from the numerical results by roughly a factor of two because the maps are quite strong). Furthermore, despite the dispersion characteristics being significantly different, use of the GVD-independent dispersion map produces pulses in both cases that are unchirped at practically the same period points. This confirms the validity of the asymptotic predictions regarding the zero-chirp points of such special dispersion maps.

In Figure 7 we show the stroboscopic evolution of the peak pulse amplitude and the chirp parameter b at successive GVD-independent chirp-free points corresponding to $a = 0.5593$ in Cases I and II when unchirped enhanced-power hyperbolic-secant pulses are launched at either the optimal point or the beginning of the normal dispersion segment. Clearly, the peak pulse amplitude and chirp oscillate much less significantly when pulses are launched
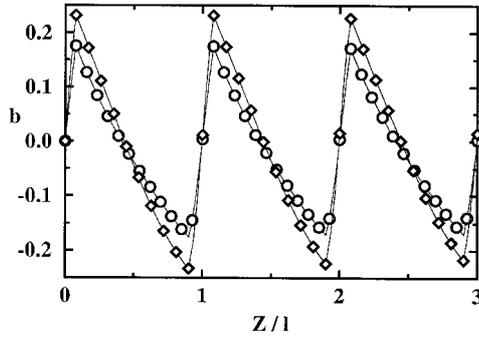
Figure 6. Evolution of the chirp parameter *b* in Cases I (circles) and II (diamonds) (with parameters described in the text) after unchirped power-enhanced hyperbolic-secant pulses are launched at a GVD-independent optimal launch point.
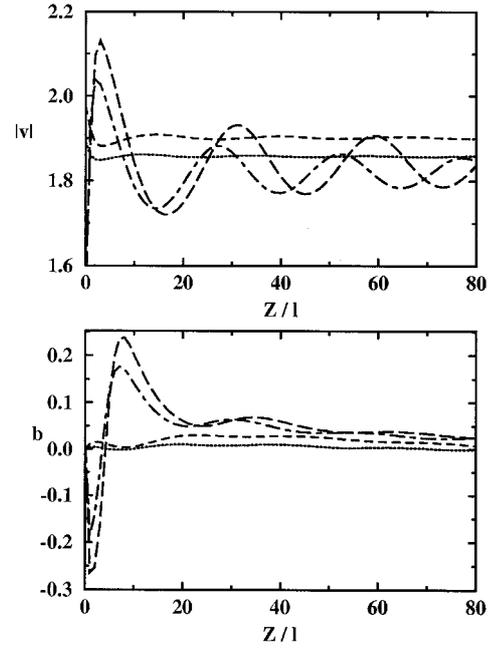
Figure 7. Graphs of the peak pulse amplitude $|v(t = 0)|$ and the chirp parameter *b* measured stroboscopically at the analyticaly-predicted GVD-independent chirp-free points. The launch point is either a chirp-free point (Case I: dotted lines; Case II: dashed lines) or the beginning of the normal dispersion segment (Case I: dot-dashed lines; Case II: long dashed lines). Here $Z/l = z/\varepsilon$. The parameters for Cases I and II are described in the text.

at the optimal point. This is because when pulses are not launched at an optimal point they ideally should be prechirped; initially unchirped pulses, therefore, must adjust themselves to become truly periodic dispersion-managed solitons, resulting in the transient peak-amplitude and chirp oscillations. Moreover, as shown in Figure 8, the amount of dispersive radiation generated is also significantly reduced in both cases when the pulses are launched at the optimal point. We emphasize that the improvement is achieved in *both* channels simultaneously using the *same* optimal launch point. Thus the optimization can be realized using the same means for different channels in dispersion-managed WDM soliton transmission.

In the demonstrations above we have been using optimal (chirp-free) launch points predicted by the asymptotic results (24) and (25). Launching unchirped hyperbolic-secant pulses with optimum power at one such point we significantly reduce the transient pulse oscillations and the accompanying shedding of dispersive radiation. It is seen in Figure 5 that pointwise agreement between the asymptotic and numerical results for the pulse chirp is excellent when the dispersion-map strength is small. For stronger maps, however, such as those considered in Figure 6, good pointwise agreement is no longer possible. Nevertheless, the asymptotic theory still provides good predictions for the *locations* of the chirp-free points, even when the map strength is large.

To show explicitly that the chirp-free locations of magic maps are truly independent of the dispersion map strength, we have calculated them for a range of dispersion values numerically. Clearly, due to the asymptotic nature of the theory, the pulse shape given by (17), or, alternatively, an initially power-enhanced hyperbolic-secant shape (which we have used as an input pulse) is not exact. However, as pointed out earlier, reasonable input pulses eventually settle into dispersion-managed solitons via the shedding of dispersive radiation. The chirp-free points of dispersion-managed solitons therefore can be easily located after the initial transient evolution has been completed. (Technically, though, the initial transient typically dies out slowly, and it is desirable to shorten artificially the transient period; this can be achieved using an averaging procedure originally proposed by Nijhof *et al.* [28].)
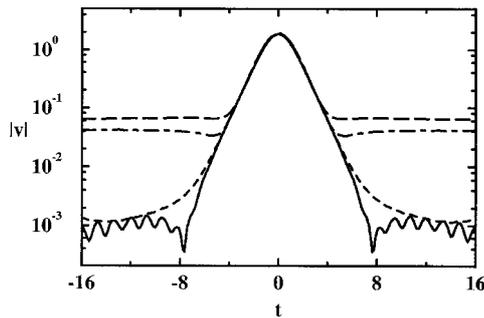
Figure 8. Pulse profiles after propagation over 82 dispersion map periods (9,840 km). The launch point is either a chirp-free point (Case I: solid lines; Case II: dashed lines) or the beginning of the normal dispersion segment (Case I: dot-dashed lines; Case II: long dashed lines). The parameters for Cases I and II are described in the text.
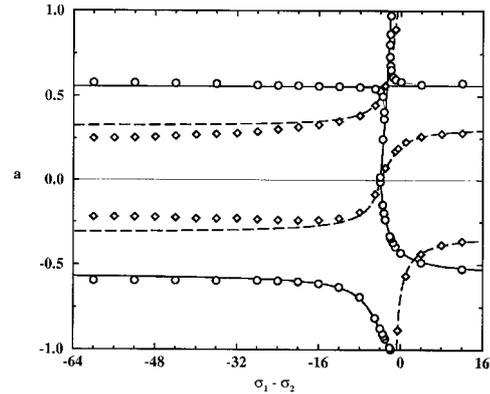
Figure 9. Comparison of numerically determined chirp-free point locations (circles: $\zeta_1 = 0.1762$; diamonds: $\zeta_1 = 0.5$) for dispersion-managed solitons with the asymptotic predictions (solid lines: $\zeta_1 = 0.1762$; dashed lines; $\zeta_1 = 0.5$). Here the normalized dispersion-map period is $\varepsilon = 0.4$ and the loss parameter is $\Lambda = 2.763$. The parameter a specifies the chirp-free point locations and is defined in the text.

A comparison between the numerically determined chirp-free point locations and the asymptotic predictions is given in Figure 9. Here the normalized dispersion-map period is taken to be $\varepsilon = 0.4$ and two map constructions (a map with $\zeta_1 = 0.1762$ and, for comparison purposes, a map with $\zeta_1 = 0.5$) with various dispersion differences $\sigma_1 - \sigma_2$ are considered. Recall that, according to the asymptotic theory, $\zeta_1 = 0.1762$ is a 'magic' fraction producing a GVD-independent chirp-free point (corresponding to $a = 0.5593$) for dispersion-managed solitons. The numerical results for $\zeta_1 = 0.1762$ plotted in Figure 9 again confirm the GVD independence of the special chirp-free point, and show that this remains valid even for stronger dispersion maps. Note the theory is not fully accurate near the point where the two branches of chirp-free locations intersect around $\sigma_1 - \sigma_2 \approx -2.7$; it is expected that such an intersection point is likely to be sensitive to perturbations, though, and the map period used here, $\varepsilon = 0.4$, is large enough that higher-order terms in the perturbation expansion may be significant.

For $\zeta_1 = 0.1762$, the dispersion difference $\sigma_1 - \sigma_2 = -60$ corresponds to a map strength of $\varepsilon^2 \mathcal{M} \approx 0.52$ and a power-enhancement factor of 73%. The agreement between the numer-

ically calculated chirp-free point locations and the asymptotically predicted ones is excellent for this map strength. For $\zeta_1 = 0.5$, the same dispersion difference corresponds to a larger map strength of $\varepsilon^2 \mathcal{M} \approx 1.1$ and a power-enhancement factor of 97%. As is clearly seen in Figure 9, the numerically calculated chirp-free point locations still agree quite well with the asymptotic predictions (and the agreement is better for the 'magic' map). The comparison above confirms the accuracy of the asymptotic predictions for the optimal (chirp-free) launch point locations for dispersion-managed solitons. Note, of course, that for very large map strengths the zero-chirp locations asymptotically become independent of the dispersion difference, because the local dispersion becomes the dominant physical effect within any map period. Only for the 'magic' map fractions, however, is the location constant for all values of the dispersion different.

6. Concluding remarks

We have used multiple-scale averaging to analyze the dynamics of dispersion-managed solitons, an expansion formally valid when the dispersion-map period is shorter than the fiber's dispersion length. The predictions of the asymptotic theory compare favorably with numerical results, oftentimes even when the strength of the dispersion map is so large that the theory would not be expected *a priori* to be valid. The simple asymptotic results are useful for providing valuable insights into the issue of optimizing dispersion-managed soliton systems.

In particular, using the asymptotic theory, we have discovered (and confirmed numerically) a class of two-step dispersion maps that render the location of the optimal (chirp-free) launch points independent of the fibers' dispersions. These GVD-independent dispersion maps are easy to implement as their design only involves choosing the ratio between the two segment lengths properly. It is demonstrated that the amount of energy shed into a dispersive pedestal can be substantially reduced by using such launch points, and that the result obtained provides a simple method for optimizing WDM soliton transmission simultaneously in several different channels.

Appendix: Normal form of perturbed NLS equation

Here we give the specific details as applied to Equation (13) of a general method for identifying nonessential perturbing terms in perturbed NLS equations [26, 27]. The method accomplishes this by using a near-identity transformation (14) to reduce (13) to a simpler form, or, 'normal form'.

We can gain an intuitive understanding of the method by seeking a solution of (13) in the form of a perturbation expansion

$$U = U_0 + \varepsilon^2 U_1 + \cdots.$$

Upon substitution of the expansion in (13), it is readily seen that U_0 satisfies the NLS equation

$$iU_{0x} - \tfrac{1}{2}\langle\sigma\rangle U_{0tt} + |U_0|^2 U_0 = 0, \tag{A.1}$$

and the correction U_1 is governed by the forced linearized NLS equation

$$\mathcal{L}\{U_1\} = iU_{1x} - \tfrac{1}{2}\langle\sigma\rangle U_{1tt} + U_0^2 U_1^* + 2|U_0|^2 U_1$$

$$= \mathcal{M}[U_0^2 U_{0tttt}^* + 6U_0 U_{0t} U_{0ttt}^* + 2(|U_0|^2)_t U_{0ttt} + 5U_0|U_{0tt}|^2$$

$$+7U_{0t}^2 U_{0tt}^* + 10|U_{0t}|^2 U_{0tt} + \tfrac{5}{2}U_{0tt}^2 U_0^*] + \cdots . \qquad (A.2)$$

Note that all of the forcing terms on the right-hand side of (A.2) are cubic in U_0 and involve four differentiations with respect to t. These are 'order-7' terms [26, 27], in the sense that the number of U_0's and t- differentiations in these terms add up to seven. Also, since the linearized NLS operator $\mathcal{L}\{\cdot\}$ involves multiplication of two U_0's and two differentiations with respect to t (in view of Equation (A.1), a differentiation with respect to x is the same order as two t derivatives), for U_1 to have a closed-form solution it must consist of terms that are order-5.

We can readily construct some order-5 terms by permuting the numbers of U_0's and differentiations with respect to t. There are six linearly independent terms of this form:

$$F_1 = U_{0tttt}, \qquad F_2 = U_{0t}(|U_0|^2)_t, \qquad F_3 = U_0(U_0 U_{0t}^*)_t,$$

$$F_4 = U_0(U_0^* U_{0t})_t, \qquad F_5 = |U_0|^2 U_{0tt}, \qquad F_6 = |U_0|^4 U_0.$$

Other order-5 terms exist and are connected with the conserved quantities of the NLS equation [26]. These are

$$F_7 = (-6|U_0|^2 U_{0t} + \langle\sigma\rangle U_{0ttt}) \int_{-\infty}^{t} |U_0|^2 \, dt,$$

$$F_8 = (-2|U_0|^2 U_0 + \langle\sigma\rangle U_{0tt}) \int_{-\infty}^{t} U_0 U_{0t}^* \, dt,$$

$$F_9 = U_{0t} \int_{-\infty}^{t} (|U_0|^4 + \langle\sigma\rangle|U_{0t}|^2) \, dt,$$

$$F_{10} = U_0 \int_{-\infty}^{t} (3|U_0|^2 U_0 U_{0t}^* + \langle\sigma\rangle U_{0t} U_{0tt}^*) \, dt.$$

[Integration is the inverse operation of differentiation, of course, and therefore contributes -1 to the 'order' of these terms.] Note that when we complete the integrals above by letting $t \to \infty$, they become the conserved quantities of the NLS equation. Meanwhile, the factors multiplying the integral terms in $F_k (k = 7, 8, 9, 10)$ are simply the functional derivatives of the integrands in F_{17-k} with respect to U_0^*. Order-7 terms involving integrals also exist, of course, but are not needed here.

Solution of (A.2) amounts to calculating $\mathcal{L}\{F_k\} (k = 1, 2, \ldots, 10)$ and then constructing a linear combination of $\mathcal{L}\{F_k\}$ that produces the right-hand side of (A.2). Although tedious, it is straightforward to show that

$$\mathcal{L}\{\mathbf{F}\} = -\mathbf{BS}, \qquad (A.3)$$

where

$$\mathbf{F} = \left[F_1, \frac{F_2}{\langle\sigma\rangle}, \frac{F_3}{\langle\sigma\rangle}, \frac{F_4}{\langle\sigma\rangle}, \frac{F_5}{\langle\sigma\rangle}, \frac{F_6}{\langle\sigma\rangle^2}, \frac{F_7}{\langle\sigma\rangle^2}, \frac{F_8}{\langle\sigma\rangle^2}, \frac{F_9}{\langle\sigma\rangle^2}, \frac{F_{10}}{\langle\sigma\rangle^2} \right]^T,$$

$$\mathbf{S} = \left[S_1, S_2, \ldots, S_9, \frac{S_{10}}{4\langle\sigma\rangle}, \frac{S_{11}}{4\langle\sigma\rangle}, \frac{S_{12}}{4\langle\sigma\rangle}, \frac{S_{13}}{4\langle\sigma\rangle}, \frac{S_{14}}{4\langle\sigma\rangle}, \frac{S_{15}}{4\langle\sigma\rangle^2} \right]^T,$$

with

$$
\begin{aligned}
S_1 &= U_0(|U_0|^2)_{tttt}, \quad S_2 = U_0 U_{0t} U_{0ttt}^*, \quad S_3 = U_0|U_{0tt}|^2, \\
S_4 &= U_{0t}^2 U_{0tt}^*, \quad S_5 = U_0 U_{0t}^* U_{0ttt}, \quad S_6 = |U_{0t}|^2 U_{0tt}, \\
S_7 &= |U_0|^2 U_{0tttt}, \quad S_8 = U_0^* U_{0t} U_{0ttt}, \quad S_9 = U_0^* U_{0tt}^2, \\
S_{10} &= U_{0t}(|U_0|^4)_t, \quad S_{11} = |U_0|^4 U_{0tt}, \quad S_{12} = U_0^3 U_{0t}^{*2}, \\
S_{13} &= U_0(|U_0|^4)_{tt}, \quad S_{14} = |U_0|^2 U_0(|U_0|^2)_{tt}, \quad S_{15} = |U_0|^6 U_0,
\end{aligned}
$$

and

$$
\mathbf{B} = \begin{pmatrix}
0 & 8 & 12 & 12 & 8 & 24 & 0 & 8 & 6 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 2 & 0 & 3 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & -1 & -5 & 2 & -4 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & -2 & -4 & 0 \\
0 & 1 & 2 & 1 & 1 & 2 & 0 & 1 & 0 & 0 & 0 & 0 & 2 & -4 & 0 \\
0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & -8 & 4 & -8 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8 & -8 & 8 & -2 & 12 & -8 \\
0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & -24 & 24 & -12 & 24 & 0 \\
0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & -4 & 10 & -16 & 4 & -20 & 12 \\
0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & -4 & 4 & -8 & 4 & -8 & 0 \\
0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 6 & -12 & 22 & -7 & 26 & -10
\end{pmatrix}.
$$

As mentioned above, there are more possible order-7 terms than those in $\mathbf{S}$, but they are not relevant here. Note also that, in terms of $\mathbf{S}$, (A.2) can be written as

$$
\mathcal{L}\{U_1\} = \mathcal{M}\mathbf{n}^T\mathbf{S},
$$

where $\mathbf{n} = [1, 2, -1, 7, -2, 10, -1, 2, \frac{5}{2}, 0, 0, 0, 0, 0, 0,]^T$.

For convenience, we express a general linear combination of $F_k (k = 1, 2, \ldots, 10)$ as

$$
f = -\mathcal{M}\mathbf{m}^T\mathbf{F} \quad \text{where } \mathbf{m} = [m_1, m_2, \ldots, m_{10}]^T,
$$

which produces $\mathcal{L}\{f\} = \mathcal{M}\mathbf{m}^T\mathbf{B}\mathbf{S}$. Complete solution of (A.2) then requires $\mathbf{m}^T\mathbf{B}\mathbf{S} = \mathbf{n}^T\mathbf{S}$, or $\mathbf{B}^T\mathbf{m} = \mathbf{n}$, which is a 15×10 system of linear equations. However, as can be checked, the linear system is over-determined and has no solutions for the coefficient vector $\mathbf{m}$. Otherwise, we would have found a solution in terms of U_0, correct to $O(\varepsilon^2)$, of the perturbed NLS Equation (1) for U_1. Since U_0 is governed by the NLS Equation (A.1), which is integrable by use of the inverse scattering transform, it could have then been said that the perturbed NLS Equation (1) would be 'integrable up to $O(\varepsilon^2)$' [26].

Although complete solution of (A.2) is not possible, the linear system (A.3) can still be used to simplify (A.2). The idea is to solve (A.2) 'partially' so that the difference between the chosen solution and the exact solution satisfies the same equation, but with simpler forcing terms. We can do this by introducing $U_1 = f + V_1$ to transform (A.2) to

$$
\mathcal{L}\{V_1\} = \mathcal{M}(\mathbf{n}^T\mathbf{S} - \mathbf{m}^T\mathbf{B}\mathbf{S}) \equiv \mathcal{M}\mathbf{n}'^T\mathbf{S}, \tag{A.4}
$$

and then trying to make the new forcing terms $\mathcal{M}\mathbf{n}'^T\mathbf{S}$ as simple as possible. Of course, what constitutes a simplification criteria may not be the same for all circumstances. Once the criteria are specified, however, $\mathbf{n}'$ is then fixed and the coefficient vector $\mathbf{m}$ can be determined from

$$\mathbf{B}^T\mathbf{m} = \mathbf{n} - \mathbf{n}'. \tag{A.5}$$

A justifiable criterion is motivated by the observation that in the order-7 terms S_1, S_{13}, S_{14} and S_{15}, differentiations only apply to the pulse intensity $|U_0|^2$. From a computational point of view, at least, these terms are relatively benign, since the pulse phase usually oscillates much more rapidly in time, while the pulse intensity varies much more smoothly. Derivatives of the pulse power, therefore, are less likely to cause numerical difficulties than those of the phase.

We therefore allow the forcing terms in (A.4) to be

$$\mathbf{n}'^T\mathbf{S} = p_1 S_1 + p_{13} S_{13} + p_{14} S_{14} + p_{15} S_{15}.$$

where p_1, p_{13}, p_{14} and p_{15} can be arbitrary constants. This effectively relaxes four constraints and hence reduces (A.5) to an 11×10 linear system. The reduced system then has the solution

$$\mathbf{m} = [\tfrac{1}{4}, 1, 1 + \alpha, 2\alpha, -2\alpha, \tfrac{3}{4}\alpha, \alpha, 3\alpha, -3\alpha, -\alpha]^T,$$

where α is a free parameter, and, accordingly,

$$p_1 = -\alpha, \qquad p_{13} = (\tfrac{1}{2} + \tfrac{25}{8}\alpha)/\langle\sigma\rangle, \qquad p_{14} = (1 + \tfrac{25}{4}\alpha)/\langle\sigma\rangle, \qquad p_{15} = -10\alpha/\langle\sigma\rangle^2.$$

Combining then (A.4) with (A.1), one finds that the new variable

$$q = U_0 + \varepsilon^2 V_1 = U + \varepsilon^2 \mathcal{M}\mathbf{m}^T\mathbf{F} + \cdots \tag{A.6}$$

satisfies

$$iq_x - \tfrac{1}{2}\langle\sigma\rangle q_{tt} + |q|^2 q = \varepsilon^2 \mathcal{M}\mathbf{m}'^T\mathbf{S} + \cdots. \tag{A.7}$$

Compared with (13), the transformed Equation (A.7) now has 'simpler' perturbing terms. Furthermore, the particular choice $\alpha = 0$ avoids including integrated terms (F_7 through F_{10}) in the near-identity transformation (A.6); these non-local terms are clearly less convenient from a computational point of view. Finally, then, with $\alpha = 0$ Equations (A.6) and (A.7) reduce to (14) and (15) used in Section 3.

Acknowledgements

conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Office of Scientific Research or the US Government.

References

1. G. P. Agrawal, *Nonlinear Fiber Optics* (2nd ed.). San Diego: Academic Press (1995) 592pp.
2. A. Hasegawa and Y. Kodama, *Solitons in Optical Communications*. Oxford: Oxford University Press (1995) 320pp.
3. L. F. Mollenauer, E. Lichtman, M. J. Neubelt and G. T. Harvey, Demonstration, using sliding-frequency guiding filters, of error-free soliton transmission over more than 20 Mm at 10 Gbit/s, single channel, and over more than 13 Mm at 20 Gbit/s in a two-channel WDM. *Electron. Lett.* 29 (1993) 910–911.
4. L. F. Mollenauer, P. V. Manyshev and M. J. Neubelt, Demonstration of soliton WDM transmission at up to 8 × 10 Gbit/s error-free over transoceanic distances. In: *Optical Fiber Communications '96 Technical Digest*. Washington: Optical Society of America (1996) p. PD22.
5. G. P. Agrawal, *Fiber-Optic Communication Systems*. New York: Wiley (1992) 445pp.
6. I. R. Gabitov and S. K. Turitsyn, Averaged pulse dynamics in a cascaded transmission system with passive dispersion compensation. *Opt. Lett.* 21 (1996) 327–329.
7. C. Kurtzke, Suppression of fiber nonlinearities by appropriate dispersion management. *IEEE Photonics Technol. Lett.* 5 (1993) 1250–1253.
8. M. Nakazawa and H. Kubota, Construction of a dispersion-allocated soliton transmission-line using conventional dispersion-shifted nonsoliton fibers. *Jpn. J. Appl. Phys. Lett.* 34 (1995) L681–L683.
9. M. Suzuki, I. Morita, N. Edgawa, S. Yamamoto, H. Taga and S. Akiba, Reduction of Gordon–Haus timing jitter by periodic dispersion compensation in soliton transmission. *Electron. Lett.* 31 (1995) 2027–2029.
10. R. W. Tkach, A. R. Chraplyvy, F. Forghieri, A. H. Gnauck and R. M. Derosier, Four-photon mixing and high-speed WDM systems. *J. Lightwave Technol.* 13 (1995) 841–849.
11. J. C. Bronski and J. N. Kutz, Modulational stability of plane waves in nonreturn-to-zero communications systems with dispersion management. *Opt. Lett.* 21 (1996) 937–939.
12. N. J. Smith and N. J. Doran, Modulational instabilities in fibers with periodic dispersion management. *Opt. Lett.* 21 (1996) 570–572.
13. W. Forysiak, K. J. Blow, and N. J. Doran, Reduction of Gordon–Haus jitter by post-transmission dispersion compensation. *Electron. Lett.* 29 (1993) 1225–1226.
14. J. P. Gordon and H. A. Haus, Random walk of coherently amplified solitons in optical fiber transmission. *Opt. Lett.* 11 (1986) 665–667.
15. N. J. Smith, N. J. Doran, F. M. Knox and W. Forysiak, Energy-scaling characteristics of solitons in strongly dispersion-managed solitons. *Opt. Lett.* 21 (1996) 1981–1983.
16. N. J. Smith, F. M. Knox, N. J. Doran, K. J. Blow and I. Bennion, Enhanced power solitons in optical fibers with periodic dispersion management. *Electron. Lett.* 32 (1996) 54–55.
17. T. Georges and B. Chabonnier, Reduction of the dispersive wave in periodically amplified links with initially chirped solitons. *IEEE Photonics Tech. Lett.* 9 (1997) 127–129.
18. I. Gabitov, E. G. Shapiro and S. K. Turitsyn, Optical pulse dynamics in fiber links with dispersion compensation. *Optics Comm.* 134 (1997) 317–329.
19. M. J. Ablowitz and H. Segur, *Solitons and the Inverse Scattering Transform*. Philadelphia: SIAM (1981) 425pp.
20. A. C. Newell, *Solitons in Mathematics and Physics*, Philadelphia: SIAM (1985) 244pp.
21. L. F. Mollenauer, S. G. Evangelides and J. P. Gordon, Wavelength division multiplexing with solitons in ultra-long distance transmission using lumped amplifiers. *J. Lightwave Technol.* 9 (1991) 362–367.
22. D. J. Richardson, R. P. Chamberlin, L. Dong and D. N. Payne, High quality soliton loss-compensation in 38 km dispersion-decreasing fibre. *Elect. Lett.* 31 (1995) 1681–1682.
23. A. J. Stentz, R. W. Boyd and A. F. Evans, Dramatically improved transmission of ultrashort solitons through 40 km of dispersion-decreasing fiber. *Opt. Lett.* 20 (1995) 1770–1772.
24. J. Kevorkian and J. D. Cole, *Perturbation Methods in Applied Mathematics*. New York: Springer (1981) 558pp.

25. A. Hasegawa and Y. Kodama, Guiding-center soliton in fibers with periodically varying dispersion. *Opt. Lett.* 16 (1991) 1385–1387.

26. T. Kano, Normal form of nonlinear Schrödinger equation. *Phys. Soc. Japan* 58 (1989) 4322–4328.

27. Y. Kodama, Optical solitons in a monomode fiber. *J. Stat. Physics* 39 (1985) 597–614.

28. J. H. B. Nijhof, N. J. Doran, W. Forysiak and F. M. Knox, Stable soliton-like propagation in dispersion managed systems with net anomalous, zero and normal dispersion. *Electron. Lett.* 33 (1997) 1726–1727.

29. V. S. Grigoryan and C. R. Menyuk, Dispersion-managed solitons at normal average dispersion. *Opt. Lett.* 23 (1998) 609–611.

30. J. N. Kutz and S. G. Evangelides, Dispersion-managed breathers with average normal dispersion. *Opt. Lett.* 23 (1998) 685–687.

31. S. K. Turitsyn and E. G. Shapiro, Dispersion-managed solitons in optical amplifier transmission systems with zero average dispersion. *Opt. Lett.* 23 (1998) 682–684.

32. T. I. Lakoba, J. Yang, D. J. Kaup and B. A. Malomed, Conditions for stationary pulse propagation in the strong dispersion management regime. *Opt. Commun.* 149 (1998) 366–375.

33. S. K. Turitsyn, N. F. Smyth and E. G. Turitsyna, Solitary waves in nonlinear dispersive systems with zero average dispersion. *Phys. Rev. E* 58 (1998) R44-R47.

34. S. K. Turitsyn, I. Gabitov, E. W. Laedke, V. K. Mezentsev, S. L. Musher, E. G. Shapiro and K. H. Spatschek, Variational approach to optical pulse propagation in dispersion compensated transmission systems. *Optics Comm.* 151 (1998) 117–135.

35. Y. Chen and H. A. Haus, Dispersion-managed solitons with net positive dispersion. *Opt. Lett.* 23 (1998) 1013–1015.

36. Y. Kodama, Nonlinear chirped RZ and NRZ pulses in optical transmission lines. In: A. Hasegawa (ed.), *New Trends in Optical Soliton Transmission Systems*. Dordrecht: Kluwer Academic (1998) 502pp.

37. T. Georges and F. Favre, Transmission systems based on dispersion managed solitons: Theory and experiment. In: A. Hasegawa (ed.), *New Trends in Optical Soliton Transmission Systems.* Dordrecht: Kluwer Academic (1998) 502pp.

38. De. Le Guen, F. Favre, M. L. Moulinard, M. Henry, G. Michaud, L. Macé, F. Devaux, B. Charbonnier and T. Georges, 200 Gbit/s 100 km-span soliton WDM transmission over 1000 km of standard fiber with dispersion compensation and pre-chirping. In: *Optical Fiber Communications '97 Technical Digest.* Washington: Optical Society of America (1997) p. PD17.

39. T. S. Yang and W. L. Kath, Analysis of enhanced-power solitons in dispersion-managed optical fibers. *Opt. Lett.* 22 (1997) 985–987.

40. V. S. Grigoryan, T. Yu, E. A. Golovchenko, C. R. Menyuk and A. N. Pilipetskii, Dispersion-managed soliton dynamics. *Opt. Lett.* 22 (1997) 1609–1611.

41. T. Yu, E. A. Golovchenko, A. N. Pilipetskii and C. R. Menyuk, Dispersion-managed soliton interactions in optical fibers. *Opt. Lett.* 22 (1997) 793–795.

42. T.-S. Yang, W. L. Kath and S. K. Turitsyn, Optimal dispersion maps for wavelength-division-multiplexed soliton transmission. *Opt. Lett.* 23 (1998) 597–599.

43. Y. Kodama and A. Maruta, Optimal design of dispersion management for a soliton-wavelength-division-multiplexed system, *Opt. Lett.* 22 (1997) 1692–1694.

44. Y. Kodama, Nonlinear pulse propagation in dispersion managed system, *Physica D* 123 (1998) 255–266.

45. M. Wald, I. M. Uzunov, F. Lederer and S. Wabnitz, Optimization of soliton transmissions in dispersion-managed fiber links, *Opt. Comm.* 145 (1998) 48–52.