PHYSICAL REVIEW A VOLUME 44, NUMBER 7 1 OCTOBER 1991

Second-order perturbations for solitons in optical fibers

D.J. Kaup

Department of Mathematics and Computer Science
and Institute for Nonlinear Studies, Clarkson University, Potsdam, New York 13699-5815

(Received 9 November 1990)

A singular perturbation theory for nonlinear Schrödinger solitons is extended into second order. An application is the calculation of the contribution of the radiation to the phase-independent interaction between soliton pairs, as suggested by Smith and Mollenauer [Opt. Lett. **14**, 1284 (1989)].

PACS number(s): 42.50.Qg, 42.81.Dp

I. INTRODUCTION

Recently, this author has presented a singular perturbation theory for soliton propagation in optical fibers. This work was an extension of previous soliton perturbation theories [2–5], and gave explicit formulas for calculating first-order effects. In fact, the analytic results were shown to very nicely compliment known numerical calculations [1].

Also recently Smith and Mollenauer have experimentally observed a long-range phase-independent interaction occurring between soliton pairs [6]. One suggestion for a possible source of this interaction was the continuous spectrum (radiation) emitted by the perturbed solitons. This would be a second-order interaction, where the nonlinearity couples the emitted radiation from one soliton to the other soliton. Estimates of such an effect require an estimate of the size and nature of the emitted continuous spectrum. The purpose of this paper is to provide such a calculation and to give numerical estimates relevant to the experiment.

In addition, we shall also calculate and numerically estimate the first-order contributions due to various perturbations, such as higher-order dispersion [7,8], the delayed Raman effect [8,9], damping [1,2,4,8–10], and the soliton self-frequency [11–14,9] shift. Of these mechanisms, the dominant one for optical fibers is damping. Although one can partially compensate for this damping by using the Raman effect to pump the soliton [15], one is then left with a net effective damping which is spatially oscillating and periodic. The difficulty with this is that this periodic damping can then pump the continuous spectrum, generating radiation of this periodicity which radiates away from the soliton. This damping-generated radiation travels freely until it collides with the next soliton. As it collides with and passes through the next soliton, it gives an attractive impulse to this second soliton, causing it to slightly shift its position toward the original soliton. Consequently, an array of solitons in an optical fiber will be unstable due to this long-range and phase-independent interaction, since it will pull initially well-separated solitons towards one another and into collision. In particular, a periodic array of solitons is unstable to this interaction. Any slight deviation of one soliton in such an array from its equilibrium position will grow exponentially. However, for the single-mode quartz fiber proposed in Ref. [7], a shift of one soliton width would require a cable length of about 400 000 km.

In Sec. II we present the general second-order equations for the perturbed nonlinear Schrödinger equation (NLS). In Sec. III, we use appropriate numerical values to estimate the magnitude of the major perturbations. Then using these values, in Sec. IV, we are able to select out the dominant second-order terms for consideration. As we shall see, the dominant second-order term is the nonlinear coupling between the Raman-compensated pump-generated radiation and the adjacent solitons. Our final result can be interpreted in terms of the phase shift that occurs to a soliton when a small packet of radiation passes through it. Thus the results of Alonso [16,17] can also be used to estimate the size of this effect.

II. THE SECOND-ORDER PERTURBED NLS

As shown in Ref. [1], one may expand solutions of the perturbed NLS

$$i\partial_t q = -\partial_x^2 q - 2q^* q^2 + \epsilon R[q, q^*] \tag{1}$$

in the singular perturbation expansion about a single soliton as

$$q = q_0 + \epsilon q_1 + \epsilon^2 q_2 + \cdots , \tag{2}$$

where

$$q_0 = \frac{A e^{i\alpha}}{\cosh\theta} , \tag{3}$$

$$\theta = 2\eta(x - \bar{x}) , \tag{4}$$

$$\alpha = -2\xi(x - \bar{x}) + \bar{\alpha} . \tag{5}$$

In the above, A is the amplitude of the soliton, η is the inverse of the width, $\bar{x}$ is its position, and $\bar{\alpha}$ is its phase at its center. Also q is the scaled electric-field envelope, t is actually the scaled position along the fiber while x is the scaled comoving temporal coordinate. As in Ref. [1], we allow the soliton amplitude and width to be independent variables. However, if the perturbation is small, we do expect A to be very close to the value of 2η, which is the pure soliton value.

We also introduce the multiple time scales, whence

$$\tau_n = \epsilon^n t \ , \tag{6a}$$

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial \tau_0} + \epsilon \frac{\partial}{\partial \tau_1} + \epsilon^2 \frac{\partial}{\partial \tau_2} + \cdots \ . \tag{6b}$$

We take $A = A(\tau_1, \tau_2, \ldots)$ and $\eta = \eta(\tau_1, \tau_2, \ldots)$. For simplicity, and to avoid certain secular terms, we also take $\bar{x}$ and $\bar{\alpha}$ to be independent of τ_0 $(= t)$ and to be dependent only on the slow time scales. Consequently, in order to have the proper zeroth-order behavior, we expand $\bar{x}$ as

$$\bar{x} = \bar{x}_{-1}\frac{1}{\epsilon} + \bar{x}_0 + \bar{x}_1 \epsilon + \cdots \ , \tag{7}$$

and similarly for $\bar{\alpha}$.

For the higher-order parts of the solution, we define the column matrix

$$v = \begin{pmatrix} e^{-i\alpha}(q_1 + \epsilon q_2 + \cdots) \\ e^{i\alpha}(q_1^* + \epsilon q_2^* + \cdots) \end{pmatrix} \ ; \tag{8}$$

then the evolution of v is given by

$$i\partial_t v + 4\eta^2 L v = F \ , \tag{9}$$

where the operator L is

$$L = \sigma_3(\partial_\theta^2 - 1) + \frac{2}{\cosh^2\theta}(2\sigma_3 + i\sigma_2) \ , \tag{10}$$

with σ_1, σ_2, and σ_3 being the Pauli spin matrices. The eigenstates and closure of L are summarized in Appendix A.

The source F is of the form

$$F = \begin{pmatrix} \mathcal{R} \\ -\mathcal{R}^* \end{pmatrix} \ , \tag{11}$$

where from all of the above, we obtain

$$\begin{aligned} \epsilon\mathcal{R} = {} & \epsilon e^{-i\alpha}R - \frac{i}{\cosh\theta}\partial_t A - \left(\frac{A\theta}{\eta\cosh\theta} + \frac{\epsilon\theta}{\eta}\mu\right)\partial_t\xi + (4\eta^2 - A^2)\left(\frac{2A}{\cosh^3\theta} + \frac{2\epsilon}{\cosh^2\theta}(2\mu + \mu^*)\right) \\ & + i\theta\left(\frac{A\sinh\theta}{\eta\cosh^2\theta} - \frac{\epsilon}{\eta}\partial_\theta\mu\right)\partial_t\eta + \left(\frac{2\xi A}{\cosh\theta} - 2iA\eta\frac{\sinh\theta}{\cosh^2\theta} + 2i\eta\epsilon\partial_\theta\mu + 2\epsilon\xi\mu\right)(\partial_t\bar{x} + 4\xi) \\ & + \left(\frac{A}{\cosh\theta} + \epsilon\mu\right)(\partial_t\bar{\alpha} - 4\eta^2 - 4\xi^2) - 2\epsilon^3(\mu^2\mu^*) - \frac{2A\epsilon^2}{\cosh\theta}(\mu)^2 - \frac{4\epsilon^2 A}{\cosh\theta}(\mu^*\mu) \ , \end{aligned} \tag{12}$$

where μ is the first component of v, which by (8) is

$$\mu = e^{-i\alpha}(q_1 + q_2\epsilon + q_3\epsilon^2 + \cdots) \ . \tag{13}$$

The terms of order ϵ in (12) were given in Ref. [1]. Here we have given the general expression, which can be expanded to any order desired. One would only have to use (6), (7), and (13) in order to expand (12) to the order desired.

One should note that the coefficients of $\partial_t\xi$ and $\partial_t\eta$ contain terms secular in θ at second order since, in general, μ will approach a plane wave for large θ. The origin of these terms and their handling is described in Appendix B. We shall not need to be concerned with them here.

The last two terms in (12) give the second-order nonlinear coupling between the soliton and the continuous spectrum (radiation). In these terms we shall find the phase-independent long-range interaction.

As discussed in Appendix A (and B), the correct expansion for v is in terms of the continuous eigenfunctions of L. Thus we take

$$v = \eta\int_{-\infty}^{\infty} dk[g(\eta k + \xi, t)|\psi, k> + \bar{g}(\eta k - \xi, t)|\bar{\psi}, k>] \ . \tag{14}$$

Consequently v, to all orders, is orthogonal to the four states $|\phi_e>$, $|\phi_0>$, $|\chi>$, and $|\theta\phi_e>$. Inner products of (12) with these states determine the evolution of the soliton parameters. To first order [1], these are

$$\partial_t\xi = \frac{\epsilon\eta}{2A} < \phi_0|\sigma_3|F_{\text{ext}}> \ , \tag{15}$$

$$4\partial_t A - 2\frac{A}{\eta}\partial_t\eta = -i\epsilon < \phi_e|\sigma_3|F_{\text{ext}}> \ , \tag{16}$$

$$\partial_t\bar{x} + 4\xi = -\frac{i\epsilon}{4A\eta} < \theta\phi_e|\sigma_3|F_{\text{ext}}> \ , \tag{17}$$

$$\partial_t\bar{\alpha} - 4\eta^2 - 4\xi^2 = \frac{\epsilon}{2A\eta} < \chi|\sigma_3|F_{\text{ext}}> \ , \tag{18}$$

where

$$|F_{\text{ext}}> = \begin{pmatrix} Re^{-i\alpha} \\ -R^* e^{i\alpha} \end{pmatrix} \tag{19}$$

is the external forcing. The densities of the radiation, g and $\bar{g}$, are given to first order by

$$\partial_t g - 4i\eta^2(k^2 + 1)g = \frac{1}{2\pi\eta a^2} < \bar{\phi}|\sigma_3|F_{\text{ext}}> \ , \tag{20}$$

$$\partial_t \bar{g} + 4i\eta^2(k^2+1)\bar{g} = \frac{-1}{2\pi\eta a^2} < \phi|\sigma_3|F_{\rm ext} > \ , \quad (21)$$

where a is given by (A19) and $< \phi|$ and $< \bar{\phi}|$ by (A11) and (A12).

Let us now consider the evolution under the general perturbation

$$R = -i\gamma q + iC_1\partial_x^3 q + iC_2\partial_x^2 q + C_3 q\partial_x(q^* q) \ , \quad (22)$$

where γ is the coefficient of linear damping [1,2,4,8–10], C_1 is the coefficient of the third-order dispersion [7,8], C_2 is the coefficient of the second-order delayed Raman effect [8,9] (the first-order delayed Raman effect is just a shift in the group velocity), and C_3 is the coefficient of the self-frequency shift [9,11–14]. Using (3) to evaluate (22) in lowest order, we have

$$\begin{aligned}|F_{\rm ext} > &= 4\eta A(12i\eta^2 C_1 - A^2 C_3\sigma_3)\frac{\tanh\theta}{\cosh^3\theta}\begin{pmatrix}1\\1\end{pmatrix}\\ &+8\eta A(3i\xi^2 C_1 - C_2\xi\sigma_3 - i\eta^2 C_1)\frac{\tanh\theta}{\cosh\theta}\begin{pmatrix}1\\1\end{pmatrix}\\ &+4\eta^2 A(iC_2 + 6C_1\xi\sigma_3)\left(\frac{1}{\cosh\theta} - \frac{2}{\cosh^3\theta}\right)\begin{pmatrix}1\\1\end{pmatrix}\\ &-A(i\gamma + 4i\xi^2 C_2 + 8\xi^3 C_1\sigma_3)\frac{1}{\cosh\theta}\begin{pmatrix}1\\1\end{pmatrix} \ .\end{aligned} \quad (23)$$

Evaluating the inner products in (15)–(18) gives

$$\partial_t\xi + \frac{16}{3}\epsilon\eta^2 C_2\xi = -\frac{16}{15}\epsilon A^2\eta^2 C_3 + \cdots \ , \quad (24)$$

$$\frac{4}{A}\partial_t A - \frac{2}{\eta}\partial_t\eta = -4\epsilon\gamma - \frac{16}{3}\epsilon C_2(\eta^2 + 3\xi^2) + \cdots \ , \quad (25)$$

$$\partial_t\bar{x} + 4\xi = 4\epsilon C_1(\eta^2 + 4\xi^2) + \cdots \ , \quad (26)$$

$$\partial_t\bar{\alpha} - 4(\eta^2+\xi^2) = 2(A^2 - 4\eta^2) + 16\epsilon\xi C_1(\eta^2 - \xi^2) + \cdots \ . \quad (27)$$

Of course, the first two results are exactly the same as those derived from the conservation laws [8,9], since the conservation laws are a consequence of the equations of motion [4,18]. However, the evolution of the phases and the radiation densities g and $\bar{g}$ cannot be obtained from the conservation laws.

The soliton self-frequency shift is given by (24), where $-\xi$ is the relative frequency. Since C_2 and C_3 are normally positive, the frequency always downshifts, but is stabilized by a nonzero C_2 [9]. From (25), once a relationship between A and η is chosen [19], one can determine how the soliton's amplitude will vary. It is driven only by the damping and the delayed Raman effect. Equation (26) gives how the soliton's center will evolve. In zeroth order, it is driven only by the relative frequency $-\xi$, which is normally zero or at least very small. Thus small corrections can be important in the evolution of $\bar{x}$, such as the higher-order dispersion in (26). This correction is nothing more than what one would obtain from a linear theory. Equation (27) gives the evolution of the soliton's phase. Note that it has a zeroth-order part of $4\eta^2$ (since ξ is normally zero or small and $A \approx 2\eta$ to within an order of ϵ).

Next we shall obtain numerical estimates for the size of the effect of the various phenomena in (22)–(27), and then calculate the second-order effects.

III. QUANTITATIVE ESTIMATES

Before giving quantitative estimates of the constants, first we need to determine the units in which (1) is written. The procedure for obtaining these has been given in other references (see, for example, the appendix of Ref. [15]). Here we shall simply summarize the results. First, one has one arbitrary unit which one could choose to be the pulse width, or some convenient unit of time on the order of the pulse width. This unit of time t_c converts time quantities into the unitless coordinate x, where

$$x = t_{\rm lab}/t_c \quad (28)$$

and $t_{\rm lab}$ is the time coordinate in the laboratory frame. Once t_c is fixed, then the characteristic length z_c is determined by the condition that the coefficient of $\partial_x^2 q$ be unity. This leads to [15]

$$z_c = \frac{4\pi c}{D\lambda^2}t_c^2 \ , \quad (29)$$

where D is the dispersion coefficient and λ is the wavelength. Note that (29) differs by a factor of 2 from that in Ref. [15]. This is because of the different normalizations used in (1). The only real difference is that our z_c will be twice that of Ref. [15]. One uses z_c to convert spatial distances into the unitless coordinate t via

$$t = z_{\rm lab}/z_c \ , \quad (30)$$

where $z_{\rm lab}$ is the spatial distance in the laboratory.

For convenience, we shall choose t_c=50 ps, since this is the typical soliton width in optical fibers. And we also take D=17 ps/nm km and λ=1.56 μm as typical values [6], although a value for D as small as 2 ps/nm km seem to be feasible [20]. Once these parameters are specified, then the characteristic length is calculated to be z_c=228 km. We shall refer to these parameters as the "standard case." Note that for lower D values, z_c increases. Thus if D=2 ps/nm km, then z_c increases by a factor of about 10 to around 2000 km. Similarly, if t_c is increased, z_c increases quadratically.

With these values now specified, we may now estimate the (unitless) size of the various perturbations. Starting with the damping, from Ref. [15] we have that the Raman gain is typically $\alpha_g \approx$0.07/km, whereas the net gain, which is periodic, has a maximum amplitude of $\alpha_g - \alpha_s \approx$0.02/km. To determine the damping constant γ in (22), we simply multiply by the value of z_c. Thus the amplitude of the periodic damping (or gain) is $\gamma_{\rm max} \approx 4.56$. Now this is larger than unity and therefore is not a small perturbation. In fact, it suggests that the damping will be just as important, if not more important, than the dispersion or the nonlinearity. That is certainly the case if one would use a lower value for D, such as 2ps/nm km. In this case, the larger value of z_c will increase the unitless amplitude of the periodic damping to approximately 40. Now clearly the periodic damping dominates both the dispersion and nonlinearity, and attemping to describe this system with a perturbation theory based on a soliton would not be advisable. Rather the linear theory would dominate. Still one could recover the soliton case by reducing the solitons's width until z_c was sufficiently small. However, for D=2 ps/nm km, one would require a pulse width of less than 1 ps in order to reduce the value of the unitless periodic damping below unity. We also wish to point out that many of the features of linear damping can be understood by constructing adiabatic invariants [21].

Next let us look at the coefficient of the soliton self-frequency shift C_3. This coefficient is basically the delay time of the Stokes response, which is of the order of 5 fs [9]. To obtain the unitless constant C_3, we divide this delay time by the characteristic time t_c. Whence $C_3 \approx 10^{-4}$ for the standard case and is very small. However, according to (24), this small value can have a cumulative effect if C_2 is not too large. Thus we also need to know this value as well. It can be obtained by expanding the Raman gain α_g in a Taylor series about a zero time delay. Assuming a simple exponential decay, one has $C_2 \approx \frac{1}{2}(\alpha_g z_c)(t_d/t_c)^2$, where t_d is the time delay of around 5 fs. One can obtain the same form from a bandwidth-limited-amplification argument also [9]. Thus the unitless value of C_2 is of the order of 8×10^{-8} for the standard case. Consequently, for realistic values of $\eta(\sim 1)$ and $\xi(|\xi| << 1), C_2$ is never important and C_3 is only important over extremely long distances. For the standard case of a 4000-km cable, one would have t in (24) ranging from 0 to 18, giving a net change in ξ of only 7×10^{-3}. By (26), this would allow a maximum shift in $\bar{x}$ of around 0.02, which is only a small fraction of the width of a soliton. Thus for these effects to be important at all, narrower pulses or longer distances or higher dispersion would have to be used.

The size of the higher-order dispersion coefficient C_1 can be estimated from experimental dispersion curves [20]. The typical slope in Fig. 3 of Ref. [20] is 40 ps/km nm μm. Since $D = -2\pi c k''/\lambda^2$, this gives $|\frac{1}{6}k'''| \approx 10^{-2}$ ps^3/km, which upon using the characteristic lengths of the standard case gives $C_1 = \frac{1}{6} z_c k'''/t_c^3 \approx 2 \times 10^{-5}$. Even with the lower dispersion of 2 ps/km nm, one can increase this part by only a factor of 10. And since z_c is proportional to t_c^2, to dramatically increase C_1, one would require pulse widths on the order of 10 fs. So higher-order dispersion effects are also very small.

In conclusion, the dominant perturbation is the periodic gain (damping). In fact, it is so large that it is almost *not* a perturbation. Instead, it is a dominant feature of the evolution and is as important as the dispersion or nonlinearity. The one saving feature in the optical-fiber case is the short periodicity of the periodic gain. Because of this, the average or first-order effects are smaller than one would normally expect. Thus one still can use a soliton based perturbation theory.

To see this, let us model the Raman-compensated case with

$$\gamma = \gamma_m \cos\left(\frac{2\pi t}{l}\right) , \tag{31}$$

where $l = L/z_c$ and L is the laboratory distance between repeaters. The important quantity is the integral of (31),

$$\Gamma = \int_0^t \gamma\, dt = \frac{\gamma_m l}{2\pi} \sin\left(\frac{2\pi t}{l}\right) , \tag{32}$$

which, for small l, does give Γ as being small. In fact, for the standard case and repeaters spaced 40 km apart, the maximum amplitude of Γ is 0.13, which is indeed small. (However, for the low-dispersion case, the maximum amplitude of Γ is 1.10. So for this case, the damping is just as important as the nonlinearity.)

IV. THE GENERATED RADIATION

Let us look at the size of the radiation generated by the periodic damping. From (21)–(23), we have

$$\partial_t \bar{g} + 4i\eta^2(k^2+1)\bar{g} = \frac{\bar{S}}{2\pi(k-i)^2} - \epsilon\gamma\bar{g} , \tag{33}$$

where

$$\bar{S} = (k^2+1)\left(\frac{A}{\eta}\gamma + \frac{1}{\epsilon}\partial_t(A/\eta)\right)\frac{\pi}{\cosh[(\pi/2)k]} - \frac{i}{2\epsilon}(4\eta^2 - A^2)\frac{A}{\eta}(k^2+1)^2\frac{\pi}{\cosh[(\pi/2)k]} . \tag{34}$$

Let us use the constraint $\partial_t\eta$=0 [1]. Then (25) gives

$$A = 2\eta \exp(-\epsilon\Gamma) . \tag{35}$$

With this, upon expanding, (33) and (34), (35) reduces to

$$\partial_t \bar{g} + 4i\eta^2(k^2+1)\bar{g} = -4i\eta^2\frac{(k+i)^2\Gamma}{\cosh[(\pi/2)k]} + \cdots . \tag{36}$$

If there is no resonance, then the amplitude of $\bar{g}$ will be on the order of $\Gamma/\cosh[(\pi/2)k]$, and thereby small if Γ is small. In this case, $\bar{g}$ will be composed mostly of long-wavelengh radiation where $|k| \lesssim 1$ (and therefore slowly moving). The higher values of k will be present, but their amplitudes will be seriously reduced due to the exponential nature of the $\cosh[(\pi/2)k]$ factor.

However, when Γ is periodic, then a resonance at certain values of k can occur whereby $\bar{g}$ will grow secularly in t. Using the model (31), this is at

$$k_0^2 = \frac{\pi}{2\eta^2 l} - 1 . \tag{37}$$

For the standard case, a solution always exists for (37) for real k whenever $L < 716$ km. However, for large k_0, the amplitude of the radiation density will be reduced because of the factor $1/\cosh[(\pi/2)k]$ in (36). In the standard case, this resonance occurs at $k_0 = 2.82$ for which $\cosh[(\pi/2)k_0] = 42$. Thus the amount of radiation produced should be small. Still, the secular nature of the resonance could allow a growth to occur.

To analyze this, we integrate (36) using the model (31). The solution is

$$\bar{g} = \bar{g}_s + \bar{g}_r + \bar{g}_p \ , \tag{38}$$

where

$$\bar{g}_s = \frac{i\eta^2(k+i)^2\gamma_m l}{\pi \cosh[(\pi/2)k]} \frac{e^{2\pi i t/l} - 1}{4\eta^2(k^2+1) + 2\pi/l} \ , \tag{39}$$

$$\bar{g}_r = \frac{-i\eta^2(k+i)^2\gamma_m l}{\pi \cosh[(\pi/2)k]} \frac{e^{-4i\eta^2(k^2+1)t} - 1}{4\eta^2(k^2+1) + 2\pi/l} \ , \tag{40}$$

$$\bar{g}_p = \frac{i\eta^2(k+i)^2\gamma_m l}{\pi \cosh[(\pi/2)k]} \frac{e^{-4i\eta^2(k^2+1)t} - e^{-2\pi i t/l}}{4\eta^2(k^2+1) - 2\pi/l} \ . \tag{41}$$

The term $\bar{g}_s$ represents that part of the continuous spectrum which moves with the soliton, since it has no dispersive time dependence. This part contains the reshaping of the soliton due to the perturbation. The part given by $\bar{g}_r$ represents the transient continuous spectrum which is generated because we started with a soliton that was initially unperturbed. Over a long time, this part evolves like the solution of the free Schrödinger equation: phase mixing to zero with an amplitude vanishing like $t^{-1/2}$.

The only part that does not vanish and does travel away from the soliton is the last term in (38), $\bar{g}_p$. Due to the resonance at k given by (37), this part of the radiation grows spatially while at the same time maintaining a fairly uniform amplitude. Thus it has a structure like a shelf which creeps outward from the soliton, in both directions, and the front of the shelf moving with the group velocity of the resonance. Of course, there will also be a much smaller amount of high-k radiation out in front, as well as some slower moving low-k radiation in the rear. Using the stationary phase on (14) to estimate the amplitude behind the front, one finds

$$\mu \approx \frac{\eta\gamma_m l(k_0+i)^2 e\sqrt{\pi}}{8\pi k_0 \cosh[(\pi/2)k_0]} e^{-2\pi i t/l} \times \left[e^{ik_0\theta} H(\theta) H\left(k_0 - \frac{\theta}{8\eta^2 t}\right) + e^{-ik_0\theta} H(-\theta) H\left(k_0 + \frac{\theta}{8\eta^2 t}\right) \right] \ , \tag{42}$$

where k_0 is the positive root of (37) and $H(x)$ is the Heaviside function. We emphasize that (42) is only an approximation in that the front of the shelf does have a structure whose width varies as $\sqrt{8\eta^2 t k_0}$. When the front is several soliton widths away from the source, this width broadens out to be larger than a soliton width and the front can be considered to be slowly varying with respect to a soliton width. Thus the Heaviside function in (42) is deceptive and only indicates the order of magnitude of the amplitude. It would be therefore useless to attempt to estimate gradients of μ from (42). However, we shall apply (42) only when the width of the front is wide compared to a soliton's width such that gradients would be expected to be of less importance.

By (8) and considering the phases, the right-going part of the first-order solution [see (2)] is

$$q_{1+} = e^{i(\alpha + k_0\theta)} \mathcal{A} H\left(k_0 - \frac{\theta}{8\eta^2 t}\right) \tag{43}$$

and the left-going part is

$$q_{1-} = e^{i(\alpha - k_0\theta)} \mathcal{A} H\left(k_0 + \frac{\theta}{8\eta^2 t}\right) \ , \tag{44}$$

where the amplitude is

$$\mathcal{A} \approx \frac{\eta\gamma_m l e\sqrt{\pi}(k_0+i)^2}{8\pi k_0 \cosh[(\pi/2)k_0]} e^{-2\pi i t/l} \ . \tag{45}$$

Let us consider the right-going radiation. As it approaches the next soliton, its effect on that soliton will be that of an almost plane wave of wave vector k_0 with an adiabatically varying amplitude. Of course, as the radiation passes through the soliton, it undergoes a phase shift in accordance with (A4). From (8), (14), and (A4), in the vicinity of the next soliton, we have

$$\mu = \mathcal{B} e^{ik_0\theta_+} \left(1 + \frac{2ik_0 e^{\theta_+}}{(k_0-i)^2 \cosh\theta_+} + \frac{1}{(k_0-i)^2 \cosh^2\theta_+}\right) \times H\left(k_0 - \frac{\theta_0}{8\eta^2 \tau}\right) + \frac{\mathcal{B}^* e^{-ik_0\theta_+}}{(k_0+i)^2 \cosh^2\theta_+} H\left(k_0 - \frac{\theta_0}{8\eta^2 t}\right) \ , \tag{46}$$

where θ_0 is θ measured from the first soliton, θ_+ is θ measured from the next one, and

$$\mathcal{B} = e^{ik_0(\theta_0 - \theta_+)} e^{i(\alpha_0 - \alpha_+)} \mathcal{A} \ , \tag{47}$$

where α_0 is the phase of the first soliton and α_+ is the phase of the next one.

Let us now ask what the dominant terms in (12) will be when a single soliton is perturbed by the above radiation field. First, from (46) the terms linear in μ would have oscillations like $e^{-2\pi i t/l} e^{ik_0\theta}$. Since k_0 is large, the integral over θ would reduce the amplitude by a factor of order $1/k_0$, and then the forcing would also be rapidly oscillatory in τ. Thus time-integrated quantities would have their amplitudes reduced to the order of $l\mathcal{B}/(2\pi k_0)$ and would remain oscillatory and noncumulative.

Now we consider the second-order (in μ) terms in (12). Some of these terms will go as $\mathcal{B}^*\mathcal{B}$ and therefore lack the oscillatory structure present in the first order. These terms are also phase independent and are the last two terms in (12). Evaluating only them, we have

$$\epsilon\mathcal{R} \approx -\frac{4A\epsilon^2 \mathcal{B}^*\mathcal{B}}{(k_0^2+1)^2 \cosh\theta_+} H^2\left(k_0 - \frac{\theta_0}{8\eta^2 t}\right) \times \left((k_0^2+1)^2 - \frac{k_0^2+3}{\cosh^2\theta_+} + \frac{3}{\cosh^4\theta_+} + \frac{2ik_0 \tanh\theta_+}{\cosh^2\theta_+}\right) . \tag{48}$$

To determine how this perturbation affects this next soliton, we evaluate the inner products with F as in (15)–(18). One finds that (48) has no effect on the evolution of the eigenvalue parameters η and ξ, but the position and phase are affected. In particular, (26) and (27) are replaced by

$$\partial_\tau \bar{x} + 4\xi = 4\epsilon C_1(\eta^2 + 4\xi^2) - \frac{4\epsilon^2 k_0 \mathcal{B}^* \mathcal{B}}{3\eta(k_0^2+1)^2} H^2 \left(k_0 - \frac{\Delta\theta}{8\eta^2 t}\right) , \quad (49)$$

$$\partial_\tau \bar{\alpha} - 4(\eta^2 + \xi^2) = 2(A^2 - 4\eta^2) + 16\epsilon\xi C_1(\eta^2 - \xi^2) + \frac{4}{\eta}\epsilon^2 \mathcal{B}^* \mathcal{B} \frac{k_0^4 + k_0^2 + \frac{2}{3}}{(k_0^2+1)^2} H^2 \left(k_0 - \frac{\Delta\theta}{8\eta^2 t}\right) , \quad (50)$$

where $\Delta\theta$ is the θ difference between the adjacent solitons. It should be noted that only the imaginary part of (48) survived in (49). This part is smaller than the real part of (48) by a factor of about k_0^{-3} ($\simeq 22$). Thus it is possible that neglected phases and gradients in (42) could be important to (49). But in any case one would not expect them to be larger than a factor of k_0^3. The estimate of (50) is not affected by this argument since it is the dominant part of (48) that survives in (50).

For the standard case, $\mathcal{B}^*\mathcal{B} = \mathcal{A}^*\mathcal{A} \approx 10^{-4}$ whence (49) and (50) give

$$\partial_\tau \bar{x}_+ \approx -6 \times 10^{-6} \frac{1}{\eta} H^2 \left(k_0 - \frac{\Delta\theta}{8\eta^2 t}\right) , \quad (51)$$

$$\partial_\tau \bar{\alpha}_+ \approx 4 \times 10^{-4} H^2 \left(k_0 - \frac{\Delta\theta}{8\eta^2 t}\right) . \quad (52)$$

This is a small effect, and for a distance of t=20, $\bar{x}_+$ would change only by 1×10^{-4}, an effect which would be undectectable. To shift by one soliton width in the standard case, one would require a 4×10^7-km cable.

Note that the interaction is attractive. The second soliton has been shifted toward the first one by the integrated amount in (51). Thus the interaction is an unstabilizing one. Consider an array of equal spaced solitons and of equal amplitudes. In this case the Raman-compensating-generated radiation will reach the next soliton at the exact same time as that from the soliton on the other side. Thus no net shift occurs. Now slightly increase the position of one. Its damping-generated radiation will reach the one just in front of it a little earlier than that from the one on the other side. Now a slight shift will occur and the shortest separation distance will shorten even more. Thus the array will be unstable.

Finally, we give a second argument for the attractive nature of this interaction. This is based on the observation that the physics of this phase-independent interaction is simply a packet of radiation passing through a soliton, with the soliton then undergoing a phase shift. To do this, we only require the following simple basics of the inverse scattering transform [22,23]. Alonso [16,17] has treated the general case. However, for our purposes, it is possible to use a very simple model and obtain the same result quite easily and directly. Consider the general solution for the scattering coefficients $(a, \bar{a}, b, \bar{b})$ when one has exactly one soliton and some radiation present. Due to the analytical properties of the scattering coefficients, one may show that for ζ in the upper-half complex plane

$$a(\zeta) = \frac{\zeta - \zeta_1}{\zeta - \zeta_1^*} \exp\left(\frac{1}{2\pi i}\int_{-\infty}^{\infty} d\zeta' \frac{1}{\zeta' - \zeta} \ln(\bar{a}a)'\right) , \quad (53)$$

where ζ_1^* is the complex conjugate of ζ_1, with ζ_1 having a positive imaginary part. As a consequence of (53), we have

$$\frac{b_k}{a_k'} = b_k(\zeta_k - \zeta_k^*) \exp\left(-\frac{1}{2\pi i}\int_{-\infty}^{\infty} d\zeta' \frac{1}{\zeta' - \zeta_k} \ln(\bar{a}a)'\right) , \quad (54)$$

$$\frac{\bar{b}_k}{a_k'} = \frac{1}{b_k}(\zeta_k - \zeta_k^*) \exp\left(-\frac{1}{2\pi i}\int_{-\infty}^{\infty} d\zeta' \frac{1}{\zeta' - \zeta_k} \ln(\bar{a}a)'\right) , \quad (55)$$

where b_k and $\bar{b}_k$ are the scattering coefficients b and $\bar{b}$ evaluated at the bound state $\zeta = \zeta_k$. In (55), we have also used the relation $\bar{b}_k b_k = 1$. We define Z_0 by

$$| b_k | = e^{2\eta Z_0} , \quad (56)$$

where we have set $\zeta_k = \xi + i\eta$. Note that $Z_0(t)$ is the position the soliton would have in the absence of any radiation. We also define

$$\Delta = \frac{1}{4\pi}\int_{-\infty}^{\infty} d\zeta \frac{[-\ln(\bar{a}a)]}{(\zeta - \xi)^2 + \eta^2} , \quad (57)$$

which is positive definite. Then (54)–(57) give

$$| b_k/a_k' | = 2\eta e^{2\eta(Z_0 - \Delta)} , \quad (58)$$

$$| \bar{b}_k/a_k' | = 2\eta e^{2\eta(Z_0 + \Delta)} . \quad (59)$$

Now with (58) and (59), and the fact that $\Delta > 0$, we may argue the sign of the interaction. Consider $t \to -\infty$ and the radiation well to the left of the soliton. In this case, we should use inversion about $x = +\infty$ for constructing the solution. [Due to the assumed localization of the radiation to the left of the soliton, the radiation part of the kernel, $(1/2\pi)\int_{-\infty}^{\infty} d\zeta (b/a) e^{2i\zeta z}$, for the Marchenko equations will be very small in the region near the soliton.] Here the kernel in the Marchenko equations effectively only contains the soliton part, and is therefore effectively a pure one-soliton solution, the position of which is determined by (58). Thus

$$x_s(t \to -\infty) = Z_0 + \Delta . \quad (60)$$

Now, let the radiation pass through the soliton and $t \to +\infty$. Clearly we again will have a pure one-soliton solution but in the region to the *left* of the radiation. And if we now invert about $x \to -\infty$, we can determine the soliton's position from the kernel for the left Marchenko equations, whose amplitude will be (59). Thus

$$x_s(t \to +\infty) = Z_0 - \Delta \ . \tag{61}$$

Comparing (60) and (61) shows that the soliton has been pulled back by an amount of 2Δ as the radiation passed through it. This is the same sign as (51) and we may use (57) to check the validity of (49) and (50).

Since $\bar{a}a = 1-|b|^2$ and the reflection coefficient is small, for small $|b|^2$, (57) reduces to

$$\Delta = \frac{1}{4\pi}\int_{-\infty}^{\infty} d\zeta \frac{1}{\zeta^2+\eta^2}|b|^2 \ . \tag{62}$$

Now we need an expression for the reflection coefficient of a soliton with a small packet of radiation to its left. From Eqs. (18) and (19) in Ref. [24], we have

$$b = -\epsilon\frac{\zeta - i\eta}{\zeta + i\eta}\int_{-\infty}^{\infty} dx e^{-2i\zeta x} q_1^*(x) \ , \tag{63}$$

where $q_1(x)$ is the packet of radiation, which can be estimated from (13), (14), (41), and (A4). Care must be taken to only use the part of $\bar{g}_p$ that is responsible for the right-going radiation. The left-going part never crosses the path of the next soliton on the right and therefore does not contribute to the phase shift of this latter soliton. Since these two parts originate from the poles at $k = \pm k_0$, it is rather easy to separate them out. The result is

$$(\bar{g}_p)_r \simeq \frac{i\mathcal{A}}{e\sqrt{\pi}\eta}\frac{e^{-8i\eta^2 k_0(k-k_0)\tau}-1}{k-k_0} \ , \tag{64}$$

where the right-going part of q_1 is then

$$(q_1)_r = e^{i\alpha}\eta\int_{-\infty}^{\infty} dk\, e^{ik\theta}\left(\frac{k-i}{k+i}\right)^2 (\bar{g}_p)_r \ . \tag{65}$$

Now evaluating (62) by using (63) and (65), one obtains

$$\Delta = \frac{2\pi}{\eta}\int_{-\infty}^{\infty} dk\frac{1}{k^2+4}|(\bar{g}_p)_r|^2 \ , \tag{66}$$

and using (64) in the limit of $t \to \infty$

$$\Delta \sim \frac{32\pi\epsilon^2\mathcal{A}^*\mathcal{A}k_0\tau}{\eta e^2(k_0^2+4)} \ . \tag{67}$$

Since -2Δ is the shift, (67) is equivalent to

$$\partial_\tau \bar{x} \sim -\frac{64\pi\mathcal{A}^*\mathcal{A}k_0}{\eta e^2(k_0^2+4)} \ . \tag{68}$$

Note that (68) is larger than (49) by a factor of about $20k_0^2$, so as already noted after (50), (49) is indeed an underestimate.

Let us note that (67) is also only an estimate even though (62) and (63) are exact to first order. This is because the radiation is being continually produced. Thus (66) is only accurate if $(\bar{g}_p)_r$ only includes that part of the radiation that has already passed through the next soliton on the right. As it stands, it includes all right-going radiation produced up to time τ, including that part which is still to cross the next soliton on the right (and therefore yet has *not* phase shifted the latter soliton). Consequently, (67) is probably an overestimate, but in the limit of $t \to \infty$, (68) should be the limiting value.

For the standard case, (68) gives

$$\partial_t \bar{x} \simeq -\frac{6\times 10^{-4}}{\eta} \ , \tag{69}$$

which would still require at least a 400 000-km cable in order to obtain a shift of only one soliton width.

APPENDIX A

The operator L was discussed in Ref. [1] and has the following eigenstates:

$$L|\psi,k> = (k^2+1)|\psi,k> \ , \tag{A1}$$

$$L|\bar{\psi},k> = -(k^2+1)|\bar{\psi},k> \ , \tag{A2}$$

$$L|\phi_0> = 0 = L|\phi_e> \ , \tag{A3}$$

where

$$|\psi,k> = e^{ik\theta}\left(1 - \frac{2ike^{-\theta}}{(k+i)^2\cosh\theta}\right)\begin{pmatrix}0\\1\end{pmatrix} + \frac{e^{ik\theta}}{(k+i)^2\cosh^2\theta}\begin{pmatrix}1\\1\end{pmatrix} \ , \tag{A4}$$

$$|\bar{\psi},k> = \sigma_1|\psi,k> \ , \tag{A5}$$

$$|\phi_e> = \frac{1}{\cosh\theta}\begin{pmatrix}1\\-1\end{pmatrix} \ , \quad |\phi_0> = \frac{\sinh\theta}{\cosh^2\theta}\begin{pmatrix}1\\1\end{pmatrix} \ . \tag{A6}$$

The closure of these states requires two more states

$$L|\chi> = -|\phi_e> \ , \tag{A7}$$

$$L|\theta\phi_e> = -|\phi_0> \ , \tag{A8}$$

where

$$|\chi> = \frac{\theta\tanh\theta - 1}{\cosh\theta}\begin{pmatrix}1\\1\end{pmatrix} \ , \tag{A9}$$

$$|\theta\phi_e> = \theta|\phi_e> \ . \tag{A10}$$

Although L is self-adjoint, instead of using (A1) and (A2) as solutions of the adjoint problem, it is best to use the states

$$<\phi,k| = e^{-ik\theta}\left(1 - \frac{2ike^{\theta}}{(k+i)^2\cosh\theta}\right)\begin{pmatrix}1\\0\end{pmatrix}^T + \frac{e^{-ik\theta}}{(k+i)^2\cosh^2\theta}\begin{pmatrix}1\\1\end{pmatrix}^T \ , \tag{A11}$$

$$<\bar{\phi},k| = <\phi,k|\sigma_1 \ , \tag{A12}$$

where the superscript T indicates the matrix transpose. Then

$$< \phi, k|L^A = -(k^2+1) < \phi, k| \ , \tag{A13}$$

$$< \bar{\phi}, k|L^A = +(k^2+1) < \bar{\phi}, k| \ . \tag{A14}$$

For the adjoint bound states and closure states, we use the matrix transpose of (A6), (A9), and (A10).

The only nonzero inner products are

$$< \phi, k'|\sigma_3|\bar{\psi}, k >= 2\pi a^2 \delta(k-k') \ , \tag{A15}$$

$$< \bar{\phi}, k'|\sigma_3|\psi, k >= -2\pi a^2 \delta(k-k') \ , \tag{A16}$$

$$< \theta\phi_e|\sigma_3|\phi_0 >= 2 =< \phi_0|\sigma_3|\theta\phi^e > \ , \tag{A17}$$

$$< \phi_e|\sigma_3|\chi >= -2 =< \chi|\sigma_3|\phi_e > \ , \tag{A18}$$

where

$$a = \frac{k-i}{k+i} \ , \tag{A19}$$

and the inner product is defined as

$$< u|\sigma_3|v >= \int_{-\infty}^{\infty} d\theta [u(\theta)]^T \sigma_3 v(\theta) \ , \tag{A20}$$

with no complex conjugations involved.

APPENDIX B

Here we will discuss the origin of the secular (in θ) terms in (12), and describe how one can handle them such that the solution for v will have no terms secular in θ.

If one is far away from the soliton (θ large), then (12) reduces to

$$\epsilon\mathcal{R} = \epsilon e^{-i\alpha} R - \frac{\epsilon\theta}{\eta}\mu\partial_t\xi - i\frac{\epsilon\theta}{\eta}(\partial_\theta\mu)\partial_t\eta + \cdots \ . \tag{B1}$$

Of course the second and third terms are secular in θ, but these terms are also proportional to either $\partial_t\xi$ or $\partial_t\eta$. Consequently, when η and ξ are stationary, these secularities do not occur. It is not surprising then that these secularities are related to the evolution of η and ξ. In fact, if one traces back to the origin of these terms, one will find that the first one originated from the phase factor $e^{-i\alpha}$ in (13) and the second one from taking the independent spatial variable to be θ instead of x. These choices were determined in first order. By introducing the phase factor in (13) and by using θ as the independent spatial variable, then a first-order solution is easily obtained by using separation of variables. But the price of this is seemingly secularity in second order.

To see how to handle this, first consider only one small-amplitude, simple plane wave emitted by a perturbed soliton. After the emission, it propagates essentially linearly with an unchanging wavelength. However, if we represent our radiation solution as a linear integral over the fundamental solutions (A4), then since θ is given by (4), if η changes in time, it follows that k must also change in time in exactly the opposite manner so as to maintain the constancy of the wavelength of the emitted radiation. Once emitted, the radiation is naturally independent of the future history of the soliton. But θ is completely dependent on this history. In other words, by using θ as the independent spatial variable, we are using a variable standard of length which is tied to the evolution of the soliton. To factor out this dependency, we must maintain the product ηk as constant. Since k is really a dummy variable (we shall always integrate over it to obtain physical quantities), it is not very feasible to have k evolve directly. Rather, it is easiest to compensate for this by allowing the density of radiation g to evolve appropriately. One can show that these secular terms can be absorbed by taking

$$v = \eta \int_{-\infty}^{\infty} dk [g(\eta k + \xi, t)|\psi, k > + \bar{g}(\eta k - \xi, t)|\bar{\psi}, k >] \ , \tag{B2}$$

where ψ and $\bar{\psi}$ are given by (A4) and (A5). By direct calculation, one has

$$\begin{aligned} -i\partial_t v - \frac{\theta}{\eta}(\partial_t\xi)\sigma_3 v - i\frac{\theta}{\eta}(\partial_t\eta)\partial_t v &= -i\eta \int_{-\infty}^{\infty} dk[(\partial_t g)|\psi, k > + (\partial_t \bar{g})|\bar{\psi}, k >] \\ &\quad + (\partial_t\eta)\int_{-\infty}^{\infty} dk (g|\eta, k > + \bar{g}|\bar{\eta}, k >) \\ &\quad + (\partial_t\xi)\int_{-\infty}^{\infty} dk (g|\xi, k > + \bar{g}|\bar{\xi}, k >) \ , \end{aligned} \tag{B3}$$

provided $|kg|$ and $|k\bar{g}|$ vanish as $|k| \to \infty$. The states $|\eta, k >$, $|\bar{\eta}, k >$, $|\xi, k >$, and $|\bar{\xi}, k >$ are given by

$$\begin{aligned} |\eta, k > = & \frac{2k\theta e^{ik\theta}}{(k+i)^2\cosh^2\theta}\begin{pmatrix}0\\1\end{pmatrix} \\ & - \frac{2ka}{(k+i)^2}(1-\tanh\theta)e^{ik\theta}\begin{pmatrix}0\\1\end{pmatrix} \\ & + \frac{2i\theta\tanh\theta e^{ik\theta}}{(k+i)^2\cosh^2\theta}\begin{pmatrix}1\\1\end{pmatrix} \\ & - \frac{2ike^{ik\theta}}{(k+i)^3\cosh^2\theta}\begin{pmatrix}1\\1\end{pmatrix} \ , \end{aligned} \tag{B4}$$

$$|\bar{\eta}, k >= \sigma_1|\eta, k > \ , \tag{B5}$$

$$\begin{aligned} |\xi, k >= & -\frac{2a}{(k+i)^2}(1-\tanh\theta)e^{ik\theta}\begin{pmatrix}0\\1\end{pmatrix} \\ & - \frac{2\theta e^{ik\theta}}{(k+i)^2\cosh^2\theta}\begin{pmatrix}1\\0\end{pmatrix} \\ & - \frac{2ie^{ik\theta}}{(k+i)^3\cosh^2\theta}\begin{pmatrix}1\\1\end{pmatrix} \ , \end{aligned} \tag{B6}$$

$$|\bar{\xi},k>=-\sigma_1|\xi,k> \quad . \tag{B7}$$

Note that these states are nonsecular in θ. Whence the θ-secular terms in (12) can be successfully and effectively summed by taking the k and time dependence of g and $\bar{g}$ to be of the form (B2).

To lowest order, these terms have no effect on the long-range phase-independent interaction.

[1] D.J. Kaup, Phys. Rev. A **42**, 5689 (1990).
[2] D.J. Kaup, SIAM J. Appl. Math. **31**, 121 (1976).
[3] K.J. Blow, N.J. Doran, and David Wood, J. Opt. Soc. Am. B **5**, 1301 (1988).
[4] Y. Kodama and K. Nozaki, Opt. Lett. **12**, 1038 (1987).
[5] Keith J. Blow and David Wood, IEEE J. Quantum Electron. **QE-25**, 2665 (1989).
[6] K. Smith and L.F. Mollenauer, Opt. Lett. **14**, 1284 (1989).
[7] Yuji Kodama and Akira Hasegawa, Opt. Lett. **7**, 339 (1982).
[8] K.J. Blow, N.J. Doran, and David Wood, Opt. Lett. **12**, 1011 (1987).
[9] K.J. Blow, N.J. Doran, and David Wood, J. Opt. Soc. Am. **5**, 1301 (1988).
[10] Boris A. Malomed, Opt. Commun. **61**, 192 (1987).
[11] F.M. Mitschke and L.F. Mollenauer, Opt. Lett. **11**, 659 (1986).
[12] J.P. Gordan, Opt. Lett. **11**, 662 (1986).
[13] Y. Kodama and K. Nozaki, Opt. Lett. **12**, 1038 (1987).
[14] Keith J. Blow and David Wood, IEEE J. Quantum Electron. **QE-25**, 2665 (1989).
[15] L.F. Mollenauer, J.P. Gordan, and M.N. Islam, IEEE J. Quantum Electron. **QE-22**, 157 (1986).
[16] L. Martinez Alonso, Phys. Rev. Lett. **54**, 499 (1985).
[17] L. Martinez Alonso, Phys. Rev. A **32**, 1459 (1985).
[18] D.J. Kaup and A.C. Newell, Phys. Rev. B **18**, 5162 (1978).
[19] See Ref. [1] for a discussion of choosing this relationship.
[20] L.G. Cohen and W.L. Mammel, Electron. Lett. **18**, 1023 (1982).
[21] R.K. Bullough, A.P. Fordy, and S.V. Manakov, Phys. Lett. **91A**, 98 (1982).
[22] M.J. Ablowitz, D.J. Kaup, A.C. Newell, and H. Segur, Stud. Appl. Math. **53**, 249 (1974).
[23] V.E. Zakharov and A.B. Shabat, Zh. Eksp. Teor. Fiz. **61**, 118 (1971) [Sov. Phys.—JETP **34**, 62 (1972)].
[24] D.J. Kaup, SIAM J. Appl. Math. **31**, 121 (1976).