Optical and Quantum Electronics **33:** 19–54, 2001.

Linear guidance properties of solitonic Y-junction waveguides

J.A. BESLEY[1,3], P.D. MILLER[2]* AND N.N. AKHMEDIEV[1]
[1]*Optical Sciences Centre, Australian National University, Canberra ACT 0200 Australia (E-mail: nna124@rsphysse.anu.edu.au)*
[2]*Department of Mathematics and Statistics, Monash University, Clayton, VIC 3168 Australia*
[3]*Present address: Racal Research Ltd., Worton Drive, Worton Grange Industrial Estate, Reading Berkshire RG2 0SB UK (E-mail: James.Besley@rrl.co.uk)*
*(*author for correspondence: E-mail: millerpd@mail.maths.monash.edu.au)*

Received 7 August 1999; accepted 15 December 1999

Abstract. A linear Y-junction waveguide device is designed using a generalization of the theory of solitonic potentials for the linear Schrödinger equation. This Y-junction device, unlike other adiabatic Y-junctions, has the advantage that it may be directly written into a planar medium with homogeneous saturable nonlinearity by a strong light beam. The generalized theory provides the error terms that are introduced when the parameters of a solitonic potential are allowed to vary in the propagation direction, and shows that under certain adiabaticity conditions the error is small although the deformation of the potential is significant. At the operating wavelength for which the device is designed to function optimally, the Y-junction has two approximate bound modes that we find explicitly. Each mode has the property that when it is excited at the neck of the junction, it exits in only one of the two output ports. In this way, the device functions like a standard modal splitter in a multimode slab waveguide. When the wavelength is detuned, modal beating is introduced that degrades the optimal switching characteristics. We describe this effect in terms of four universal coupling functions using perturbation theory.

Key words: planar waveguides, spatial solitons, Y-junctions

1. Introduction

Passive dielectric linear Y-junctions are fundamental elementary components of planar lightwave integrated circuits used for signal demultiplexing applications. A junction designed for such a purpose[1] consists of a single input port or 'trunk' and two (or more) output ports or 'branches'. In idealized operation, a linear superposition of signals propagating into the device along the trunk is divided without loss such that each signal exits through only one branch. There are two types of demultiplexing Y-junction of interest here:

[1]Other design purposes include power division. In this case, the goal is simpler: simply the predictable splitting of input power among the output branches with minimal loss. There is no need for the device to differentiate among multiple input signals. See Weissman *et al.* (1989), Shirafuji and Kurazono (1991), Ni *et al.* (1995) and Hsu and Lee (1998) for some applications.

wavelength demultiplexers and mode demultiplexers. Wavelength demultiplexing Y-junctions have a trunk that is a single-moded waveguide over a range of wavelengths of interest, and the coupling of this mode into the branches depends on the wavelength; if designed perfectly for operation at two different wavelengths λ_L and λ_R, say, the trunk mode at wavelength λ_L (respectively λ_R) excites only the mode of the left (respectively right) branch. Thus, the two wavelength components present in an input signal are separated at the output. See Negami *et al.* (1989) and Weissman (1995) for design examples. On the other hand, mode demultiplexing Y-junctions are designed for operation at a fixed wavelength (although they tend to have robust performance over a range of wavelengths). The trunk is multi-moded at the operating wavelength, and ideally, each of the trunk modes excited at the input is channeled to a distinguished branch upon output. Some design examples can be found in Kapon and Thurston (1987), Yuan *et al.* (1994), Love *et al.* (1996), Henry and Love (1997), Besley *et al.* (1998b).

The choice of geometry for the junction waveguides proposed in the literature often appears to be driven, understandably, by issues of practicality of fabrication by known techniques. Thus, most proposed designs are combinations of elementary components like straight channels and fixed-angle tapers, and virtually all designs have step-index profiles, so that only a few different materials, or a few steps of an exposure process, are required for fabrication.

On the other hand, many of well-established rules that hold in these relatively simple situations can be violated in a favourable way if one drops these design restrictions by allowing arbitrary reasonable spatial variations in the refractive index in the planar waveguide. For example, exploiting the separability of the wave equation in unusual coordinate systems, Marcatili (1985) showed that as long as one is willing to admit smooth spatial variations in the refractive index, one can build a taper with no losses whatsoever, either in terms of power transfer among bound modes or coupling to radiation. This result clearly indicates that the common rules of thumb concerning power loss in tapers hold only under the fabrication-favorable assumptions about the refractive index of the sort described above.

There are similar results for certain kinds of junction devices for which the distribution of refractive index in the planar waveguide is prescribed by a mathematical machine intimately connected with soliton theory (Miller 1996; Miller and Akhmediev 1996, 1998). Using these techniques, one can design multiport junction waveguides with N input and N output ports, all single-moded, that function simultaneously as N absolutely lossless power dividers, one for each input port. The power transfer characteristics can be prescribed fairly arbitrarily in the design process. Not surprisingly, the lossless behavior of these devices requires their operation at a particular wavelength.

Because these 'solitonic' waveguides, like the simpler waveguides proposed by Marcatili, would require precise and spatially detailed control on the refractive index during fabrication, it seems these designs must wait in the knowledge bank for further advances in fabrication technology. On the other hand, the coincidental mathematical connection of the design of the solitonic waveguides with soliton theory actually has far-reaching implications for device fabrication using existing technology.

The idea is simply that a usefully large number of the solitonic waveguides correspond to planar refractive index distributions that are identical to those self-induced in a homogeneous planar (focusing or defocusing) Kerr medium by a monochromatic linearly polarized optical field. In the simplest cases, the field inducing the index change is just a nonlinear superposition of spatial Kerr solitons. The corresponding index profiles are the multiport junction solitonic waveguides (Miller 1996; Miller and Akhmediev 1996). If the solitons are co-propagating, that is, as 'higher-order' solitons, then the field induces a periodic index distribution. This periodic index distribution has multiple exact bound modes (Floquet or Bloch modes) at the appropriate wavelength (Besley *et al*. 1997, 1998a).

Therefore, fabricating a solitonic waveguide can be as easy as launching a system of pump beams into a homogeneous Kerr medium. As they provide the solitonic waveguide, the pump beams must be maintained while the device is being used to guide weaker signal beams.[2] This means that the signal beams must differ in either wavelength or polarization from the pump, simply in order to be distinguished at the output. If one uses a different wavelength for the signal, then in the absence of material resonances and unusually phase-matched parametric processes, the signal beam simply propagates in the (fixed) potential induced by the pump. Unfortunately, the detuning of the wavelength of the signal from that of the pump slightly perturbs the diffraction coefficients of the linear propagation so that the solitonic waveguide theory does not hold in the strict sense. The losses introduced by the detuning are, however, small, and may be taken into account perturbatively (Besley *et al*. 1998a; Miller *et al*. 2000). On the other hand, if one launches the signal beam at the same wavelength as the pump but in the orthogonal linear polarization, then the signal is influenced not only by the induced index of the pump, but also by four-wave mixing terms which again ruin the solitonic waveguide theory in its idealized form. However, it has recently been shown (Kang *et al*. 1996) that if the pump and signal polarizations are mutually incoherent, then the four-wave mixing terms may be neglected and the solitonic waveguide theory again holds. For both methods of separating the pump from the signal, laboratory experiments have shown that signal beams

[2]Weakness of the signals is generally needed so that they do not affect the refractive index themselves, but see also Chen *et al*. (1996).

can indeed be guided by the pump. In connection with the scheme of using a signal of a different wavelength from the pump, De la Fuente *et al.* (1991) used bright spatial solitons to induce waveguides in homogeneous materials and observed guidance of signals of other wavelengths, Luther-Davies and Xiaoping (1992a, b) used dark spatial solitons in organic media to do the same, and Chen *et al.* (1996) observed similar effects in photorefractive media. The AlGaAs experiments reported by Kang *et al.* (1996) show that the phase-incoherence method of suppressing four-wave mixing effects in the scheme of polarization-orthogonal pump and signal actually works well into the nonlinear regime for the signal.

Many of the materials in which these experiments have been performed are highly nonlinear and consequently the Kerr nonlinearity is at best a crude approximation for the true response of the medium. Thus, the authors of these papers report seeing self-guided beam structures that are unfamiliar within the context of cubic Schrödinger solitons. In particular, beams that split into two as they propagate were reported in Chen *et al.* (1996) and Luther-Davies and Xiaoping (1992a). In reference to guiding weak signals with these strong pumps, the authors identified these structures as nonlinearly induced Y-junctions. It is possible to analyze the splitting phenomenon in more detail if one posits a mathematical model for the nonlinear response of the medium that is slightly more complicated that the Kerr effect, but still contains new physics. A reasonable correction to a focusing Kerr nonlinearity that models saturation of the refractive index for high intensities is a small defocusing quintic term (that is, a negative correction to the refractive index proportional to the square of the intensity). Such a simple model is in fact very good for some experimental materials, like the polymer polydiacetylene para-toulene sulfonate (PTS) (Lawrence *et al.* 1994). As has been observed numerically (Artigas *et al.* 1997) and analyzed using multiscale asymptotics (Besley *et al.* 1999), this small non-Kerr correction destabilizes the periodic double-soliton and causes it to split asymmetrically into two solitons with different propagation directions accompanied by only a small amount of radiation.

Because the non-Kerr correction is small, the splitting process can be considered as an adiabatic one that is 'locally Kerr' and can be described approximately by multiscale perturbation theory based on integrable soliton mathematics (Besley *et al.* 2000). Analogously, the propagation of signal beams in a solitonic Y-junction induced in a material like PTS can be studied by an adiabatic deformation of the theory already in place, with losses of similar magnitude to the usual adiabatic losses for waveguide junctions, tapers, and so on.

In this paper, we study in detail the linear guidance properties of these adiabatic solitonic Y-junctions. We begin by reviewing the design and properties of solitonic waveguides, a class of lossless structures indexed by

several parameters. For some values of the parameters, one can obtain multimode periodic waveguides that can serve as Y-junction trunks, while for other values of the parameters one can find X-junctions whose output arms might serve as Y-junction branches. We propose a Y-junction waveguide by letting the parameters depend on z, the coordinate in the propagation direction, in such a way that the multimode trunk gradually opens up into the two branches of the X-junction. We show that, when operated at a particular wavelength (a design parameter) the junctions behave as modal splitters in the usual sense, guiding each of two independent Floquet modes of the periodic trunk into a unique output branch. As a particular example, we consider the specific adiabatic deformation that would be self-induced by a double-soliton pump beam launched into saturating material like PTS. In order to permit the possible use of the device at different wavelengths (so that if the device is written with light, the signal may be differentiated from the pump without using two polarization components, for example) we study the way the mode switching characteristics depend on the operating wavelength. Although we imagine that these solitonic junctions would be easiest to implement all-optically by using a nonlinear process to write the waveguide, the adiabatic theory that we develop applies in more general circumstances, when the adiabatic deformation of the solitonic waveguide parameters is specified arbitrarily, assuming that the resulting structure can be fabricated by more conventional means.

2. Solitonic planar refractive index distributions

As is well-known, the paraxial approximation applied to Maxwell's equations for monochromatic fields in a weakly inhomogeneous medium leads promptly to the linear Schrödinger equation as a model for beam propagation. Assume light of vacuum wavelength λ_0 is propagating in a two-dimensional medium having laboratory coordinates x_{lab} and z_{lab}. Suppose the refractive index in the medium is given by $n(x_{\text{lab}}, z_{\text{lab}})$ with background value n_0. Choose a convenient length scale L_0, and a small dimensionless scaling parameter δ. Introducing dimensionless coordinates $z = \delta^2 z_{\text{lab}}/L_0$ and $x = \delta x_{\text{lab}}/L_0$, the dimensionless frequency parameter $\beta = 2\pi L_0 n_0/\lambda_0$, and the representation of the electric field vector as $\vec{E} = f(x,z)\exp(\mathrm{i}\beta z/\delta^2)\vec{e}$ for some unit vector $\vec{e}$, one finds

$$\mathrm{i}\beta\frac{\partial f}{\partial z} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2} - \beta^2 V_0(x,z) f = 0, \tag{1}$$

up to terms of order $O(\delta^2)$ where

$$V_0(x,z) = -\frac{1}{2\delta^2}\left[\frac{n(xL_0\delta^{-1}, zL_0\delta^{-2})^2}{n_0^2} - 1\right]. \tag{2}$$

For a given function $V_0(x,z)$, the paraxial approximation becomes as accurate as desired if it is possible to deform the index distribution $n(x_{\text{lab}}, z_{\text{lab}})$ as δ goes to zero in such a way that the function $V_0(x,z)$ remains fixed. In this limit, the physical index profile becomes wide (of size $L_0\delta^{-1}$), long (of size $L_0\delta^{-2}$), and a shallow deformation of the background value (so that $n_{\max} - n_0 = n_0 O(\delta)$). In most applications, we will consider the scales to be chosen so that at the operating wavelength λ_0 of interest, $\beta = 1$. Varying the optical wavelength for a given physical index profile then means changing the value of β.

We now recall the construction of solitonic potentials[3] for the linear Schrödinger equation. These are special cases of $V_0(x,z)$ with qualitative properties that are very useful for integrated optics applications, and moreover for which the linear beam propagation problem can be solved exactly in terms of explicit formulas for all initial conditions. More details can be found in Miller and Akhmediev (1998) and in the Appendix of Miller *et al.* (2000). We emphasize that, as our ultimate interest is one of device design for integrated planar waveguide optics, where the linear Schrödinger equation is a basic model for beam propagation, the design of a potential function representing inhomogeneities in the medium to achieve specific results is of equal interest to the study of the solutions of the equation for a given potential. With this in mind, let us begin.

Regarding notation, we will use the inner product in $L^2(\mathbb{R})$ defined by

$$\langle f(\cdot), g(\cdot)\rangle \doteq \int_{-\infty}^{\infty} f(x)^* g(x)\mathrm{d}x. \tag{3}$$

Here and throughout, the star denotes complex conjugation. We use $\Re(w)$ and $\Im(w)$ respectively for the real and imaginary parts of w.

2.1. ALGEBRAIC CONSTRUCTION OF EXACT SOLITONIC POTENTIALS

We are going to specify a waveguide consisting of a number of linear channels oriented at various angles in a planar medium. Let N and M be independent positive integers. M corresponds to the number of linear channels. N is, roughly speaking, the number of internal degrees of freedom available for the design of each channel; these become important only where the channels are

[3]In Miller and Akhmediev (1996) and Miller *et al.* (2000) these are called 'separable potentials'.

overlapping. Let M complex numbers ζ_k, $\Im(\zeta_k) > 0$ be given along with M constant vectors $\vec{g}_0^{(1)}, \ldots, \vec{g}_0^{(M)}$ in $\mathbb{C}^N$. Introduce the scalar expression

$$a(x,z,\zeta) = \left(\zeta^M + \sum_{p=0}^{M-1} \zeta^p a^{(p)}(x,z)\right) \exp(-2\mathrm{i}\zeta x), \tag{4}$$

and the N-component vector expression

$$\vec{b}(x,z,\zeta) = \sum_{p=0}^{M-1} \zeta^p \vec{b}^{(p)}(x,z). \tag{5}$$

Here, $a^{(p)}(x,z)$ and $b^{(p)}(x,z)$ are unknown coefficient functions. We now determine them from the given data by insisting that the following relations hold for all $k = 1, \ldots, M$:

$$\begin{aligned} a(x,z,\zeta_k) &= \vec{g}^{(k)\dagger}\vec{b}(x,z,\zeta_k), \\ \vec{b}(x,z,\zeta_k^*) &= -a(x,z,\zeta_k^*)\vec{g}^{(k)}, \end{aligned} \tag{6}$$

where

$$\vec{g}^{(k)} = \vec{g}_0^{(k)} \exp(-2\mathrm{i}\zeta_k^{*2} z). \tag{7}$$

It is easy to check that these relations imply, for each fixed x and z, a square $M \times (N+1)$ linear system for the unknowns $a^{(k)}(x,z)$ and the components of $\vec{b}^{(k)}(x,z)$; therefore one can write down exact formulas for these in terms of determinants.

From the solution of this linear algebraic system, define the solitonic potential function

$$V_0(x,z) \doteq -4 \sum_{n=1}^{N} \left| b_n^{(M-1)}(x,z) \right|^2, \tag{8}$$

and the radiation mode function

$$\Psi_{\mathrm{r}}(x,z,\zeta) \doteq \left(\pi \prod_{k=1}^{M} |\zeta - \zeta_k|^2\right)^{-1/2} a(x,z,\zeta) \exp(-2\mathrm{i}\zeta^2 z). \tag{9}$$

Finally, define the bound modes $\Psi_{\mathrm{b},k}(x,z)$ for $k = 1, \ldots, M$ to be linear combinations of the functions $a(x,z,\zeta_j^*)\exp(-2\mathrm{i}\zeta_j^{*2} z)$ for $j = 1, \ldots, M$ that are chosen to be orthonormal with respect to L^2 in x for some fixed z.

The importance of these steps for optical applications is that the following two things can be proved:

1. Each function $\Psi_{\mathrm{r}}(x,z,\zeta)$ for $\zeta \in \mathbb{R}$ and each function $\Psi_{\mathrm{b},k}(x,z)$ is a solution of the linear Schrödinger equation

$$\mathrm{i}\frac{\partial f}{\partial z} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2} - V_0(x,z)f = 0. \tag{10}$$

2. For each fixed z, these functions form a complete orthonormal basis of the function space $L^2(\mathbb{R})$. The orthogonality conditions can be written in the form:

$$\begin{aligned} &\langle \Psi_{\mathrm{r}}(\cdot,z,\zeta), \Psi_{\mathrm{r}}(\cdot,z,\eta)\rangle = \delta(\zeta-\eta),\\ &\langle \Psi_{\mathrm{r}}(\cdot,z,\zeta), \Psi_{\mathrm{b},k}(\cdot,z)\rangle = 0,\\ &\langle \Psi_{\mathrm{b},k}(\cdot,z), \Psi_{\mathrm{b},j}(\cdot,z,)\rangle = \delta_{kj}. \end{aligned} \tag{11}$$

Since the L^2 inner product is invariant under the unitary flow of (10), the coefficients expressing $\Psi_{\mathrm{b},k}(x,z)$ in terms of $a(x,z,\zeta_k^*)\exp(-2\mathrm{i}\zeta_k^{*2}z)$ can be taken to be independent of z.

As indicated by the subscripts, the exact solutions $\Psi_{\mathrm{b},k}(x,z)$ represent bound state solutions (subscript 'b') that have finite L^2 norm, while the exact solutions $\Psi_{\mathrm{r}}(x,z,\zeta)$ represent for each real ζ unnormalizable solutions that superpose to form radiation fields (subscript 'r').

Remark. Any orthogonal basis of the bound state subspace spanned by $a(x,z,\zeta_k^*)\exp(-2\mathrm{i}\zeta_k^{*2}z)$ for $k=1,\ldots,M$ will do. This subspace is also spanned by $a(x,z,\zeta_k)\exp(-2\mathrm{i}\zeta_k^2 z)$ for $k=1,\ldots,M$. It follows from the above orthogonality relations that the entire bound state subspace is orthogonal to all radiations modes $\Psi_{\mathrm{r}}(x,z,\zeta)$ for real ζ. Within the bound state subspace, the following bi-orthogonality relations hold:

$$\begin{aligned} &\langle a(\cdot,z,\zeta_j)\exp(-2\mathrm{i}\zeta_j^2 z), a(\cdot,z,\zeta_k^*)\exp(-2\mathrm{i}\zeta_k^{*2}z)\rangle\\ &\quad = -\delta_{jk}\Im(\zeta_k)\prod_{m\neq k}(\zeta_k^*-\zeta_m)(\zeta_k^*-\zeta_m^*). \end{aligned} \tag{12}$$

These relations are often useful in finding a convenient orthogonal basis of bound states. □

To solve the linear Schrödinger equation (10) in general, one uses completeness (Miller and Akhmediev 1998) to expand the solution in terms of the basis:

$$f(x,z) = \sum_{k=1}^{M} B_{\mathrm{b},k}(z)\Psi_{\mathrm{b},k}(x,z) + \int_{-\infty}^{\infty} B_{\mathrm{r}}(z,\zeta)\Psi_{\mathrm{r}}(x,z,\zeta)\mathrm{d}\zeta. \tag{13}$$

Substituting this expansion into (10), using the fact that the modes are themselves exact solutions of (10) and then using the orthogonality relations to separate the projections, the Equation (10) becomes simply

$$\mathrm{i}\frac{\mathrm{d}B_{\mathrm{b},k}}{\mathrm{d}z}(z) = 0, \quad k = 1, \ldots, M, \quad \mathrm{i}\frac{\partial B_{\mathrm{r}}}{\partial z}(z, \zeta) = 0. \tag{14}$$

These equations are the starting point for developing coupled-mode equations for perturbations of the solitonic waveguide profile $V_0(x, z)$.

2.2. THE CASE OF $M = 2$ AND $N = 1$: TWO-SOLITON POTENTIALS

Here, we briefly consider the special cases of $M = 2$ and $N = 1$ in more detail. This special case already contains rich behavior that we will exploit further in later sections of this paper. We call these 'two-soliton' potentials because for all choices of the numbers ζ_1, ζ_2, and the scalars $g_0^{(1)}$ and $g_0^{(2)}$, the function $\psi(x, z) := 2\mathrm{i}b^{(1)}(x, z)$ is a two-soliton solution of the nonlinear Schrödinger equation

$$\mathrm{i}\frac{\partial \psi}{\partial z} + \frac{1}{2}\frac{\partial^2 \psi}{\partial x^2} + |\psi|^2\psi = 0. \tag{15}$$

The parameters ζ_k then play the role of the soliton eigenvalues for the Zakharov–Shabat eigenvalue problem (Zakharov and Shabat 1972). The parameters $g_0^{(k)}$ are the corresponding proportionality constants.

In the context of planar waveguide optics, the Equation (15) has meaning very similar to that of the linear Schrödinger equation. As is well-known, it describes the propagation of beams in homogeneous planar media subject to the Kerr effect in which the refractive index of the material is altered by the presence of an intense light beam. The field intensity $|\psi|^2 = -V_0$ is a nonlinear self-consistent potential generated by the light beam. This correspondence makes the study of solitonic potentials $V_0(x, z)$ especially attractive for optical applications, since the possibility exists of using a strong light beam to 'write' the waveguide structure in a nonlinear medium, and then subsequently using this structure to control the propagation of weaker signals that are unable to change the refractive index themselves, but certainly are influenced by the presence of the induced potential $V_0(x, z)$.

Taking $M = 2$ and $N = 1$, the algebraic procedure described above yields the following formula for the potential $V_0(x, z)$:

$$V_0(x, z) = -16\left|\frac{N}{D}\right|^2, \tag{16}$$

where with

$$J := \zeta_1 - \zeta_2, \quad K := \zeta_1^* - \zeta_2, \quad \text{and} \quad G_k := g_0^{(k)} \exp(-2\mathrm{i}(\zeta_k^* x + \zeta_k^{*2} z)), \tag{17}$$

the numerator and denominator are given by

$$N := JK^* \Im(\zeta_2) G_2 + JK \Im(\zeta_1) G_1 + J^* K \Im(\zeta_2) |G_1|^2 G_2 + J^* K^* \Im(\zeta_1) G_1 |G_2|^2, \tag{18}$$

$$D := |J|^2 (1 + |G_1|^2 |G_2|^2) + |K|^2 (|G_1|^2 + |G_2|^2) - 8 \Im(\zeta_1) \Im(\zeta_2) \Re(G_1 G_2^*). \tag{19}$$

The bi-orthogonality conditions above suggest that in the special case of $M = 2$, a convenient orthonormal basis of bound states is given by the two functions

$$\begin{aligned} \Psi_{\mathrm{b},1}(x,z) &:= \frac{a(x,z,\zeta_1^*) \exp(-2\mathrm{i}\zeta_1^{*2} z)}{\|a(\cdot,z,\zeta_1^*) \exp(-2\mathrm{i}\zeta_1^{*2} z)\|_2}, \\ \Psi_{\mathrm{b},2}(x,z) &:= \frac{a(x,z,\zeta_2) \exp(-2\mathrm{i}\zeta_2^2 z)}{\|a(\cdot,z,\zeta_2) \exp(-2\mathrm{i}\zeta_2^2 z)\|_2}. \end{aligned} \tag{20}$$

Alternatively, both ζ_1^* and ζ_2 may be exchanged for their respective complex conjugates. Explicit formulas for these modes are given in terms of the function a evaluated at $\zeta = \zeta_1^*$ and $\zeta = \zeta_2$:

$$a(x,z,\zeta_1^*) \exp(-2\mathrm{i}\zeta_1^{*2} z) = \frac{2\mathrm{i} J^* K \Im(\zeta_1)}{D g_0^{(1)}} (J^* |G_1|^2 G_2 - K^* G_1 |G_2|^2 - J G_1), \tag{21}$$

$$a(x,z,\zeta_2) \exp(-2\mathrm{i}\zeta_2^2 z) = \frac{2\mathrm{i} K \Im(\zeta_2) g_0^{(2)*}}{D} (2\mathrm{i} \Im(\zeta_1) G_1 - J^* |G_1|^2 G_2 - K^* G_2). \tag{22}$$

The radiation mode function $\Psi_{\mathrm{r}}(x,z,\zeta)$ can also be found explicitly, but we will not need it below.

2.2.1. *Periodic potentials and their bound states*

As a special case of the above construction, one may take the real parts of the numbers ζ_1 and ζ_2 to be the same. By a Galilean transformation, we may

without loss of generality assume the real parts are both zero, so that $\zeta_1 = \mathrm{i}m_1$ and $\zeta_2 = \mathrm{i}m_2$ for some positive real numbers $m_1 > m_2$. It is easily checked that in this case the potential function $V_0(x,z)$ is z-periodic with period $L = \pi/(m_1^2 - m_2^2)$. That is, the corresponding nonlinear Schrödinger field $\psi(x,z)$ is a second-order soliton.

The two functions $\Psi_{\mathrm{b},1}(x,z)$ and $\Psi_{\mathrm{b},2}(x,z)$ are independent exact bound state solutions of the linear Schrödinger equation with this periodic potential $V_0(x,z)$. Both solutions have Bloch form:

$$\Psi_{\mathrm{b},1}(x,z+L) = \mu\Psi_{\mathrm{b},1}(x,z), \quad \Psi_{\mathrm{b},2}(x,z+L) = \mu\Psi_{\mathrm{b},2}(x,z), \tag{23}$$

for all x and z, with the same Floquet multiplier $\mu = \exp(2\mathrm{i}m_1^2 L) = \exp(2\mathrm{i}m_2^2 L)$. It is indeed very unusual for z-periodic potentials to have exact bound state solutions; typically the Schrödinger equation with a z-periodic potential may be considered to be a PDE analogue of a parametrically forced oscillator, and as in the classical example of Mathieu's equation, periodic solutions usually do no exist. Therefore, one expects that these bound state solutions are structurally unstable, and that they would disappear with the introduction of a typical perturbation to the periodic potential $V_0(x,z)$. The mechanism for the destruction of the bound states in this case is a resonant coupling to the radiation modes expressed by the function $\Psi_{\mathrm{r}}(x,z,\zeta)$. For small perturbations, it is a second-order effect and thus the modes may be considered to be metastable. This structural instability and the associated dynamics of the decay process have recently been investigated in detail (Miller *et al.* 2000). Typical perturbations also introduce first-order effects that do not lead to the decay of the bound states but instead modify the location of the Floquet multipliers on the unit circle. A bifurcation occurs under which the multipliers for the two modes become distinct; this implies the appearance of a distinguished basis of Floquet eigenfunctions, which need not be proportional to $\Psi_{\mathrm{b},1}(x,z)$ and $\Psi_{\mathrm{b},2}(x,z)$ given above. These phenomena were also described recently (Besley *et al.* 1997, 1998a) in a context where the parameters $g_0^{(1)}$ and $g_0^{(2)}$ were selected to that $V_0(x,z)$ was an even function of x. In this case, the Floquet eigenfunctions that persist to first order under perturbation are of odd and even parity in x, and do not correspond to the function $a(x,z,\zeta)\exp(-2\mathrm{i}\zeta^2 z)$ evaluated at any of the particular ζ values $\zeta_1, \zeta_2, \zeta_1^*$, or ζ_2^*; they are orthonormal superpositions of these.

2.2.2. *Interacting well potentials and their bound states*

In the generic case when the real parts of ζ_1 and ζ_2 do not agree, the potential function $V_0(x,z)$ is qualitatively different from the special case described

above. As $z \to \pm\infty$, the potential is asymptotically a superposition of two potential wells that are moving with respect to each other:

$$V_0(x,z) \sim V_0^{(1)\pm}(x,z) + V_0^{(2)\pm}(x,z), \tag{24}$$

with the individual wells having the form

$$V_0^{(k)\pm}(x,z) = -4m_k^2 \operatorname{sech}^2(2m_k(x - v_k z) - \delta_k^{\pm}), \tag{25}$$

where $\zeta_k = -v_k/2 + \mathrm{i}m_k$ for v_k and m_k real, and where $\delta_k^{\pm}$ are numbers that can be computed from $g_0^{(1)}$ and $g_0^{(2)}$. That is, the nonlinear field $\psi(x,z)$ is a lossless collision of two soliton beams. Without loss of generality, we assume that $v_2 > v_1$. With this assumption, the two bound modes $\Psi_{\mathrm{b},1}(x,z)$ and $\Psi_{\mathrm{b},2}(x,z)$ are completely confined to one or the other potential well in the limit $z \to +\infty$. See Fig. 1 for pictures corresponding to the choices $\zeta_1 = 0.019 + \mathrm{i}$, $\zeta_2 = -0.047 + 0.4\mathrm{i}$, and $g_0^{(1)} = g_0^{(2)} = 1$. The two potential wells interact with each other for finite z, and a mode that is isolated in a single well for, say, large positive z will not necessarily be so confined for large negative z. In fact, the energy for large negative z is always shared between the wells (there is no scattering loss to radiation in the interaction), and the amount in each well can be computed exactly (Miller and Akhmediev 1996). Linear combinations of the two modes $\Psi_{\mathrm{b},1}(x,z)$ and $\Psi_{\mathrm{b},2}(x,z)$ can be arranged to be confined in one or the other well for large negative z. In this case ($M = 2$) these alternate solutions can be rather simply obtained by evaluating $a(x,z,\zeta)\exp(-2\mathrm{i}\zeta^2 z)$ for $\zeta = \zeta_1$ and $\zeta = \zeta_2^*$ and normalizing.

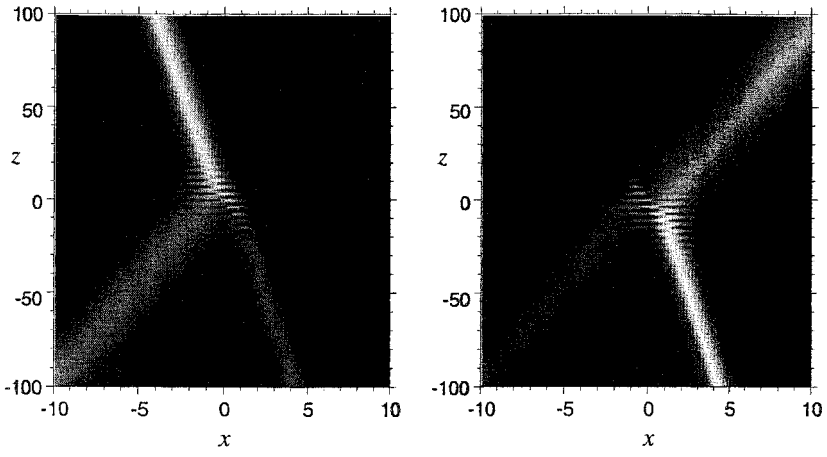

Fig. 1. Density plots of the square modulus of the orthogonal bound solutions $\Psi_{\mathrm{b},1}(x,z)$ (*left*) and $\Psi_{\mathrm{b},2}(x,z)$ (*right*) corresponding to $\zeta_1 = 0.019 + \mathrm{i}, \zeta_2 = -0.047 + 0.4\mathrm{i}$ and $g_0^{(1)} = g_0^{(2)} = 1$.

3. Y-junctions and longitudinal deformation of solitonic potentials

The solitonic z-dependent potentials described in detail above have been studied in several contexts, perhaps the most interesting of which is the application to the design of all-optically induced special-purpose waveguides for use in integrated planar optics. In this picture, the periodic potentials described first are models for periodic planar waveguides that have potential applications as frequency detectors and spectral filters. At a prescribed operating wavelength, such a device supports (within the applicability of the paraxial approximation, that is, neglecting back-reflection) exactly two bound modes that propagate permanently along the axis of the device. On the other hand, the interacting well potentials are models for waveguide junction devices. At a prescribed wavelength these devices support modes that represent the efficient channeling of optical energy through the X-shaped waveguide junction with no loss to radiation. There is crosstalk,[4] between the linear channels, but there is no loss to radiation, and the crosstalk is precisely quantifiable and can even be prescribed in the design parameters ζ_1 and ζ_2 (Miller and Akhmediev 1996).

One can try to generalize these results, by considering how the two output channels of the X-junction device might be viewed as branches of a Y-junction, and thus be effectively spliced onto a single trunk corresponding to the periodic waveguide device. Both pieces have two bound modes and one might like to consider how, if the two devices were somehow smoothly merged together, the light propagating along the trunk of the Y might be divided among the two branches, and how the result depends upon which modes of the trunk are excited.

The main idea is that one can obtain a Y-shaped potential by considering the parameters ζ_1 and ζ_2 to depend on z in such a way that along the length of the device they go from being purely imaginary (corresponding to the periodic waveguide trunk) to having distinct real parts (corresponding to the X-junction shape). This leads us to consider how the theory changes when the parameters depend on z.

Let the parameters $\zeta_1, \ldots, \zeta_M$ and the vectors $\vec{g}^{(k)}$ be arbitrary functions of z (i.e. instead of taking them to be respectively constants and pure exponential functions as before). Note that the linear algebraic procedure presented in Section 2 goes through as before, and specifies the various coefficients $a^{(p)}(x,z)$ and $\vec{b}^{(p)}(x,z)$. From the formula (8) one then obtains a potential function $V_0(x,z)$. Similarly, one obtains from (9) the function $\Psi_{\mathrm{r}}(x,z,\zeta)$ and then finds functions $\Psi_{\mathrm{b},k}(x,z)$ by applying the Gram–Schmidt orthonormalization procedure to the functions

[4]Zero-crosstalk X-junctions can also be designed using solitonic potential theory. See Miller (1996).

$$a(x,z,\zeta_k^*(z))\exp\left(-2\mathrm{i}\int_0^z \zeta_k(s)^{*2}\mathrm{d}s\right) \tag{26}$$

for $k=1,\ldots,M$. With the parameters taken to be arbitrary functions of z in this way,

1. A greater variety of waveguide shapes, encoded in the function $V_0(x,z)$, can be obtained.
2. The functions $\Psi_\mathrm{r}(x,z,\zeta)$ for $\zeta\in\mathbb{R}$ and $\Psi_{\mathrm{b},1}(x,z),\ldots,\Psi_{\mathrm{b},M}(x,z)$ still form a complete orthonormal basis of $L^2(\mathbb{R})$ for each fixed z, and can therefore be used to expand any function of x and z in terms of z-dependent coefficients. The orthogonality relations are exactly the same as before.

The theory diverges from the exact solitonic potential theory given in Section 2 when we consider the evolution of these 'modes' in z. They no longer satisfy the linear Schrödinger equation (10) corresponding to the deforming potential function $V_0(x,z)$. In Appendix A, the exact equations satisfied by these functions are derived and used to write the corresponding exact coupled-mode equations for deformed two-soliton potentials. These equations are obtained by projecting the solution of the linear Schrödinger equation (10) with deforming potential $V_0(x,z)$ onto the orthonormal basis, and they take the form

$$\begin{aligned}
\mathrm{i}\frac{\mathrm{d}B_{\mathrm{b},1}}{\mathrm{d}z}(z) &= M_{11}^{(\mathrm{NA})}(z)B_{\mathrm{b},1}(z)+M_{12}^{(\mathrm{NA})}(z)B_{\mathrm{b},2}(z)+\int_{-\infty}^{\infty}N_1^{(\mathrm{NA})}(z,\zeta)B_\mathrm{r}(z,\zeta)\mathrm{d}\zeta\\
\mathrm{i}\frac{\mathrm{d}B_{\mathrm{b},2}}{\mathrm{d}z}(z) &= M_{21}^{(\mathrm{NA})}(z)B_{\mathrm{b},1}(z)+M_{22}^{(\mathrm{NA})}(z)B_{\mathrm{b},2}(z)+\int_{-\infty}^{\infty}N_2^{(\mathrm{NA})}(z,\zeta)B_\mathrm{r}(z,\zeta)\mathrm{d}\zeta\\
\mathrm{i}\frac{\partial B_\mathrm{r}}{\partial z}(z,\zeta) &= O_1^{(\mathrm{NA})}(z,\zeta)B_{\mathrm{b},1}(z)+O_2^{(\mathrm{NA})}(z,\zeta)B_{\mathrm{b},2}(z)\\
&\quad+\int_{-\infty}^{\infty}K^{(\mathrm{NA})}(z,\zeta,\eta)B_\mathrm{r}(z,\eta)\mathrm{d}\eta,
\end{aligned} \tag{27}$$

where the coefficients are explicitly given in Appendix A.

The system (27) captures the exact dynamics of the Schrödinger Equation (10) with potential $V_0(x,z)$ constructed from the arbitrary given functions $\zeta_1(z)$, $\zeta_2(z)$, $g^{(1)}(z)$ and $g^{(2)}(z)$. The superscript (NA) on the matrix elements indicates that their contribution to the dynamics reflects the effect of 'non-adiabatic' deformations in the potential. However, if the dependence of the parameters ζ_k on z is adiabatic, so that $\mathrm{d}\zeta_k/\mathrm{d}z$ and $\mathrm{i}\mathrm{d}g^{(k)}/\mathrm{d}z-2\zeta_k^{*2}g^{(k)}$ are uniformly small in z, then the coefficients on the right-hand side of (27) will be small, with a size that is precisely quantified in Appendix A. This means that solutions of (27) will be close to solutions of the trivial system (14) over very long distances in z. To be more precise, it follows from the explicit formulas in Appendix A that if ϵ^2 is a small uniform bound for the derivatives $\mathrm{d}\zeta_k/\mathrm{d}z$ and the differences $\mathrm{i}\mathrm{d}g^{(k)}/\mathrm{d}z-2\zeta_k^{*2}g^{(k)}$, then the difference between the function

$B_{\mathrm{b},k}(z)$ computed on the one hand from the system (14) and on the other from the system (27) will be $O(\epsilon^2)$ over fixed z-intervals as ϵ goes to zero. On expanding z-intervals of size $O(\epsilon^{-1})$, the error will still be small, $O(\epsilon)$.

In this paper, we will always consider the functions $\zeta_k(z)$ to be of the form

$$\zeta_k(z) = -\frac{\epsilon}{2} v_k(\epsilon z) + \mathrm{i} m_k, \tag{28}$$

where the m_k are real constants, and where the real functions $v_k(Z = \epsilon z)$ are bounded and have bounded derivatives with respect to their arguments. Given these, we will further assume that

$$g^{(k)}(z) = g^{(k)}(0) \exp\left(-2\mathrm{i} \int_0^z \zeta_k(s)^{*2} \mathrm{d}s \right). \tag{29}$$

In this case, it is easy to verify the adiabaticity with the estimate

$$\sum_{k=1}^{2} \sup_{z\in\mathbb{R}} \left| \frac{\mathrm{d}\zeta_k}{\mathrm{d}z}(z) \right| = \frac{\epsilon^2}{2} \sum_{k=1}^{2} \sup_{Z\in\mathbb{R}} \left| \frac{\mathrm{d}v_k}{\mathrm{d}Z}(Z) \right| = O(\epsilon^2). \tag{30}$$

This means that on length scales of size ϵ^{-1} we may simply drop the terms on the right-hand side of (27) at the cost of uniformly small $O(\epsilon)$ error. However, one of the main messages we want to deliver in this paper is that with the above adiabaticity assumptions on $\zeta_k(z)$, the deformation in the transverse direction x is significant over length scales $z = O(\epsilon^{-1})$ where the coupling of the modes is negligible. That is, the 'centers of mass' of the individual potential wells can be changed by an order $O(1)$ amount over z distances that scale like ϵ^{-1}, and the effect of mode coupling can still be neglected. This opens the door the several new possibilities in all-optical device design for planar optical waveguides, including the design of effective Y-junctions.

4. Further corrections for wavelength detuning

In our analysis of beam propagation in solitonic Y-junction waveguides we will need to take into account small shifts in the operating wavelength. This leads to additional terms in the coupled-mode equations as we will now see. Consider the effect of slightly changing the frequency parameter β in the neighborhood of $\beta = 1$, starting with an equation of the form

$$\mathrm{i}(1+\xi)\frac{\partial f}{\partial z} + \frac{1}{2}\frac{\partial^2 f}{\partial \tilde{x}^2} - (1+\xi)^2 V_0(\tilde{x}, z) f = 0, \quad \beta = 1 + \xi, \tag{31}$$

for some fixed function $V_0(\tilde{x}, z)$, say specifying a Y-junction as a deformed solitonic potential. In transcribing this equation from (1) in the beginning Section 2, we have written $\tilde{x}$ instead of x because we anticipate the utility of a change of variables: $x = \tilde{x}\sqrt{1+\xi}$. Making this change of variables, we obtain for $f = f(x, z)$,

$$\mathrm{i}\frac{\partial f}{\partial z} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2} - (V_0(x,z) + W(x,z))f = 0, \tag{32}$$

where $W(x,z) = (1+\xi)V_0(x/\sqrt{1+\xi}, z) - V_0(x,z)$.

This perturbation $W(x,z)$ is uniformly small for all x and z as ξ goes to zero. This makes it easy to incorporate directly into the coupled mode equations. Expanding the solution of (32) in the complete orthonormal basis corresponding to the deforming background potential $V_0(x,z)$ as described in Section 3, and projecting onto the basis elements $\Psi_{\mathrm{b},k}(x,z)$ for $k = 1, 2$ and $\Psi_{\mathrm{r}}(x,z,\zeta)$, we find the equations

$$\begin{aligned}
\mathrm{i}\frac{\mathrm{d}B_{\mathrm{b},1}}{\mathrm{d}z}(z) &= M_{11}(z)B_{\mathrm{b},1}(z) + M_{12}(z)B_{\mathrm{b},2}(z) + \int_{-\infty}^{\infty} N_1(z,\zeta)B_{\mathrm{r}}(z,\zeta)\mathrm{d}\zeta \\
\mathrm{i}\frac{\mathrm{d}B_{\mathrm{b},2}}{\mathrm{d}z}(z) &= M_{21}B_{\mathrm{b},1}(z) + M_{22}B_{\mathrm{b},2}(z) + \int_{-\infty}^{\infty} N_2(z,\zeta)B_{\mathrm{r}}(z,\zeta)\mathrm{d}\zeta \\
\mathrm{i}\frac{\partial B_{\mathrm{r}}}{\partial z}(z,\zeta) &= O_1(z,\zeta)B_{\mathrm{b},1}(z) + O_2(z,\zeta)B_{\mathrm{b},2}(z) + \int_{-\infty}^{\infty} K(z,\zeta,\eta)B_{\mathrm{r}}(z,\eta)\mathrm{d}\eta,
\end{aligned} \tag{33}$$

where

$$\begin{aligned}
M_{kj}(z) &:= M_{kj}^{(\mathrm{NA})}(z) + \langle \Psi_{\mathrm{b},k}(\cdot,z), W(\cdot,z)\Psi_{\mathrm{b},j}(\cdot,\mathrm{z})\rangle, \\
N_k(z,\zeta) &:= N_k^{(\mathrm{NA})}(z,\zeta) + \langle \Psi_{\mathrm{b},k}(\cdot,z), W(\cdot,z)\Psi_{\mathrm{r}}(\cdot,z,\zeta)\rangle, \\
O_k(z,\zeta) &:= O^{(\mathrm{NA})}(z,\zeta) + \langle \Psi_{\mathrm{r}}(\cdot,z), W(\cdot,z)\Psi_{\mathrm{b},k}(\cdot,z)\rangle, \\
K(z,\zeta,\eta) &:= K^{(\mathrm{NA})}(z,\zeta,\eta) + \langle \Psi_{\mathrm{r}}(\cdot,z,\zeta), W(\cdot,z)\Psi_{\mathrm{r}}(\cdot,z,\eta)\rangle.
\end{aligned} \tag{34}$$

Note that as a consequence of the reality of the correction $W(x,z)$, the corresponding contributions to the matrix elements M_{jk} form a Hermitian symmetric matrix, the contributions to $O_k(z,\zeta)$ are the complex conjugates of those to $N_k(z,\zeta)$, and the contribution to $K(z,\zeta,\eta)$ is a Hermitian kernel.

In the application we will consider very soon, the correction W will scale like ϵ, where ϵ is the small parameter introduced at the end of Section 3. From the discussion at the end of that section, it should be clear that with this scaling, the corrections coming from the wavelength perturbation $W(x,z)$ will dominate any contributions from the nonadiabaticity of the potential

deformation over length scales in z of length proportional to ϵ^{-1}. This leads to very simple equations of motion in which nonadiabatic effects can be neglected.

5. Beam propagation in solitonic Y-junction waveguides

We now use the coupled-mode equations, which describe the exact dynamics of beam propagation in a deformed and possibly wavelength-detuned solitonic waveguide to describe the switching and frequency response properties of a Y-junction waveguide. As stated in Section 3, we will always take $\zeta_k(z)$ to be of the form (28), with m_k being constants and where the velocity functions $v_k(\cdot)$ are uniformly bounded. Usually we will further consider the velocity functions to satisfy $v_k(0) = 0$ and $v_k(+\infty) = v_k^\infty$, so the velocities ultimately saturate and become constant. Also, $v_2(Z) \geq v_1(Z)$ for all $Z \geq 0$. In this case, we know that for small ϵ, the correction terms on the right-hand side of (27) will be negligible over distances z of size ϵ^{-1}. At the same time, over these distances the deformation of the potential $V_0(x,z)$ in the transverse x direction is an order $O(1)$ effect.

5.1. A SOLITONIC Y-JUNCTION WRITTEN WITH LIGHT

As a specific example of such functions $v_k(Z)$, we now describe one way they may be obtained from a self-consistent nonlinear problem. Suppose one tried replacing the integrable cubic nonlinear Schrödinger equation (15) where it occurs in the solitonic potential theory with the cubic-quintic Schrödinger equation:

$$\mathrm{i}\frac{\partial\psi}{\partial z} + \frac{1}{2}\frac{\partial^2\psi}{\partial x^2} + |\psi|^2\psi = \epsilon^2|\psi|^4\psi, \tag{35}$$

where $\epsilon > 0$ is a small parameter. That is, one considers potential functions $V_0(x,z) = -|\psi(x,z)|^2$ where now $\psi(x,z)$ satisfies (35) rather than (15). With this sign on the quintic term, the Equation (35) describes the propagation of intense light beams in a homogeneous medium like PTS where the Kerr effect is corrected with a saturation of the dependence of the refractive index on the intensity. For extreme intensities the Kerr effect gives way to the quintic term which has the opposite sign. However, for more physical moderate intensities the quintic term introduces a slight weakening of the Kerr effect that is appropriate for the modeling of saturable media, like polymer and photorefractive materials. As we pointed out in Section 1, the advantage of this thought experiment is that adiabatic solitonic Y-junctions

associated with self-consistent potentials of the cubic–quintic model (35) can be written in the medium simply by launching a pump beam with an appropriate profile. Similar Y-junction waveguides have been created in the laboratory using exactly this idea (Luther-Davies and Xiaoping 1992a; Chen *et al.* 1996).

It has been observed numerically (Artigas *et al.* 1997) that the influence of the quintic term in (35) for small ϵ on the quasiperiodic two-soliton state of the cubic equation (that corresponds to the periodic waveguides described either) causes its destabilization. The two solition components gradually separate from one another, gaining independence. Ultimately, the two solitons split apart completely and the perturbation does not further affect them qualitatively.

These effects can be described analytically (Besley *et al.* 2000) using multiscale asymptotics. For small ϵ, the inverse-scattering transform that is the nonlinear mode decomposition diagonalizing the dynamics of the unperturbed Equation (15) remains useful in the study of (35) as a starting point for perturbation theory. The solition eigenvalues ζ_k of the Zakharov–Shabat eigenvalue problem (Zakharov and Shabat 1972) are no longer constant when ψ evolves according to (35) rather than (15), but become slowly varying. The method of multiple scales applied in the scattering transform domain shows that the soliton eigenvalues ζ_k may be taken to be of the form (28). This analysis results in a simple model for the interaction of two particles in one dimension with the coordinates $x_1(Z)$ and $x_2(Z)$, where $Z = \epsilon z$. The model takes the simple form of Newton's equations:

$$m_1 \frac{\mathrm{d}^2 x_1}{\mathrm{d}Z^2} = F_1(x_1, x_2; m_1, m_2), \quad m_2 \frac{\mathrm{d}^2 x_2}{\mathrm{d}Z^2} = F_2(x_1, x_2; m_1, m_2), \tag{36}$$

and the link to (28) is that the functions $v_k(Z)$ are just the velocities $\mathrm{d}x_k/\mathrm{d}Z$ of the respective particles. Most of the work involved is in the computation of the force functions F_k. However, these force functions conserve the total momentum, which allows the problem to be reduced to a similar equation for the relative displacement $y(Z) = x_2(Z) - x_1(Z)$, and then a scaling symmetry can be used to reduce the number of independent parameters from two (m_1 and m_2) to one, a normalized effective mass, $\tilde{M}$. Thus, with $q := y\sqrt{m_1 m_2}$ and $S := (m_1 m_1)^{3/2} Z$, the dynamical equations can be put in the form

$$\tilde{M} \frac{\mathrm{d}^2 q}{\mathrm{d}S^2} = F(q; \tilde{M}), \tag{37}$$

for a one-parameter family of force functions F.

As described in Besley *et al.* (2000), the (q, $\mathrm{d}q/\mathrm{d}S$) phase plane for this system exhibits a single unstable fixed point at the origin. Five orbits have the

same energy as the fixed point: the fixed point itself considered as an orbit, and two branches each of the stable and unstable manifolds making up the separatrix in the usual fashion. All of the separatrix branches become parallel to the q axis for large q, which corresponds to the force F being of essentially finite range and dying out for large q. So, typical orbits starting near the origin in the phase plane ultimately wind up along the unstable manifold, where the velocity saturates to a constant value.

From solutions of the dynamical system (37), one unravels the changes of variables and obtains corresponding solutions of (36). In turn, these give the soliton eigenvalues $\zeta_k(z)$ in the form (28) and the corresponding proportionality constants $g^{(k)}(z)$. These are then used to construct the approximate solution $\psi(x,z) = 2ib^{(1)}(x,z)$ of (35), according to the algebraic procedure described in Sections 2 and 3. The qualitative agreement between this perturbative calculation and numerical simulations of (35) is demonstrated in Fig. 2. The (essentially identical) pictures in Fig. 2 should be viewed by the reader as plots of the potential $V_0(x,z)$ of a Y-junction planar waveguide device (the potential well is deepest where the plot lightest, and the potential goes to zero on the black background) of exactly the type described earlier.

5.2. EFFICIENT TRUNK MODE SPLITTING

The main observation is that, regardless of whether the $v_k(Z)$ are supplied by a nonlinear theory or are simply cooked up subject to the assumptions

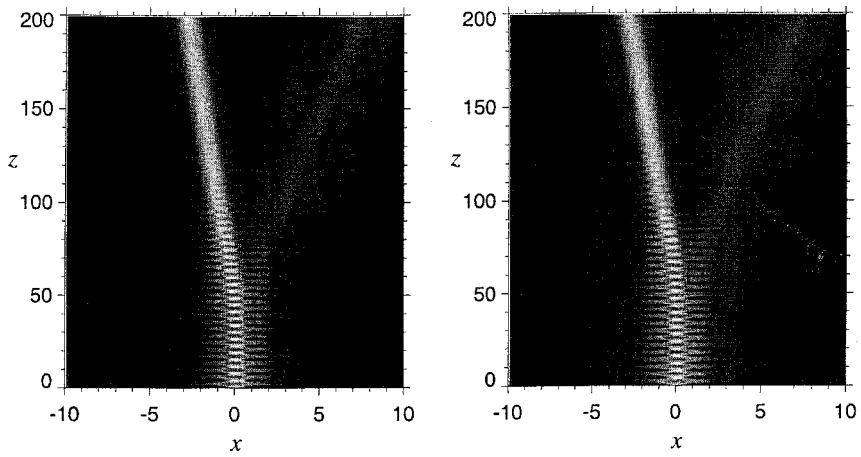

Fig. 2. *Left*: the field $|\psi(x,z)|^2$ as obtained by perturbation theory. *Right*: the field $|\psi(x,z)|^2$ as obtained by numerical simulation of the perturbed Equation (35). In both pictures $m_1 = 1, m_2 = 0.4$, and $\epsilon^2 = 0.00125$. The initial velocities were both taken to be zero, and $x_1(0)$ was a small negative number while $x_2(0)$ was a small positive number. This choice perturbs the initial phase point slightly away from the fixed point at the origin.

outlined above in a brute-force exercise of engineering (see Section 6 for an example of the latter), the splitting and separation of the coordinates x_1 and x_2 by an order one amount always takes place over scales of length ϵ^{-1}, and on this scale the dynamics is well approximated up to terms of size $O(\epsilon)$ by (14).

This means that the functions $\Psi_{b,1}(x,z)$ and $\Psi_{b,2}(x,z)$ defined in Section 3 when the parameters ζ_1 and ζ_2 satisfy (28) are independent approximate modes of the adiabatic solitonic Y-junction. The approximation improves as ϵ decreases in magnitude. Moreover, since each of these modes smoothly connects a Floquet mode of the z-periodic trunk of the junction onto exactly one of the two modes of the branches pictured in Fig. 1, these modes describe an adiabatic modal splitting process. These facts are illustrated in Figs. 3 and 4. Using the functions $v_k(Z)$ determined from perturbation theory applied to (35), the adiabatic potential theory of Section 3 is applied to compute the potential $V_0(x,z)$ of the Y-junction (this is basically what is illustrated in Fig. 2). Then, Fig. 3 compares the corresponding adiabatic bound mode $\Psi_{b,1}(x,z)$ with the result of a numerical integration of the linear Schrödinger equation with the Y-junction potential $V_0(x,z)$ for $\epsilon^2 = 0.00125$. Fig. 4 does the same with the adiabatic mode $\Psi_{b,2}(x,z)$. There is indeed a small error associated with the adiabatic approximation, but it is order $O(\epsilon)$ and thus vanishes in the adiabatic limit. We emphasize again the auxiliary role played by the particular choice of the functions $v_k(Z)$; virtually any bounded smooth functions can be used to the same effect. This fact underscores the enormous range of possibilities for device design available using solitonic potentials.

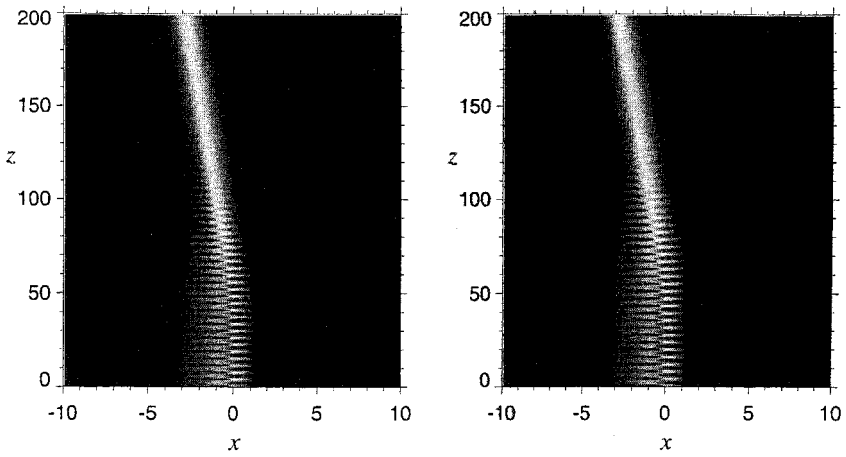

Fig. 3. *Left*: a plot of the square modulus of the approximate mode $\Psi_{b,1}(x,z)$ of the adiabatic solitonic Y-junction. *Right*: corresponding numerical simulation of the linear Schrödinger equation with adiabatically deforming two-soliton potential. The nonadiabaticity parameter is $\epsilon^2 = 0.00125$, and the channel depth parameters are $m_1 = 1$ and $m_2 = 0.4$.

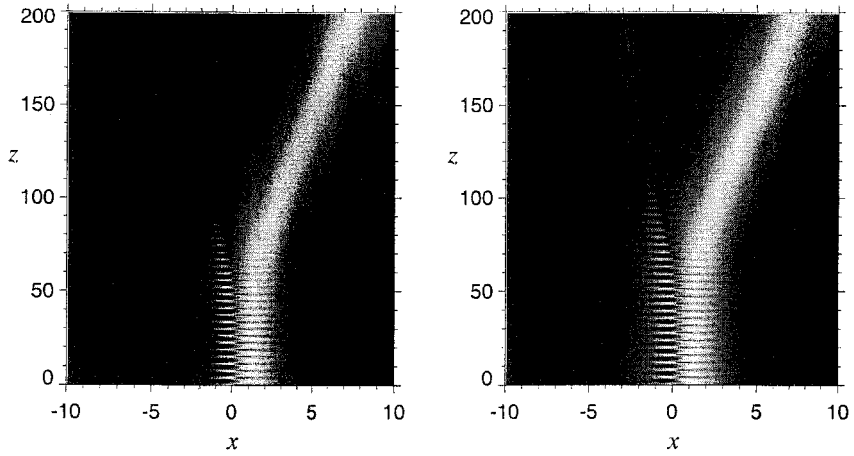

Fig. 4. *Left*: a plot of the square modulus of the approximate mode $\Psi_{b,2}(x,z)$ of the adiabatic solitonic Y-junction. *Right*: corresponding numerical simulation of the linear Schrödinger equation with adiabatically deforming two-soliton potential. The parameters are the same as in Fig. 3.

The pictures in Figs. 3 and 4 demonstrate the fact that solitonic Y-junctions of the type described here will perform very well as efficient switching devices. The fact that, once the two output ports of device separate, the approximate bound modes $\Psi_{b,k}(x,z)$ are very close to those of the X-junction (see Fig. 1) implies both the confinement of each mode to one or the other of the output ports (again up to a small error of size $O(\epsilon)$) and the explicit form of the excitation required at the trunk of the Y-junction in order to channel the light in a particular direction.

5.3. WAVELENGTH DETUNING AND FREQUENCY RESPONSE

The adiabatic Y-junction device described above is designed to operate at a particular wavelength λ_0 where by choice of units we have $\beta = 1$ in the linear Schrödinger equation. The effect of detuning the wavelength slightly is modeled by the same equation with the same potential function $V_0(x,z)$, but with $\beta \neq 1$. It is easily illustrated using numerical integration of the linear Schrödinger equation, as in Fig. 5. The effect of detuning the wavelength appears to be a dynamical exchange of energy between the two modes that resembles a beating process. Roughly speaking, how much light is transmitted into each of the two output ports depends on what superposition of the two modes is excited when the trunk of the Y opens up.

We want to consistently extend our model to capture this effect when $\beta = 1 + \xi$ and ξ is small. The main difficulty in a formal treatment by the method of multiple scales would appear to be the presence of two independent small parameters, ϵ and ξ. We remedy this situation immediately by

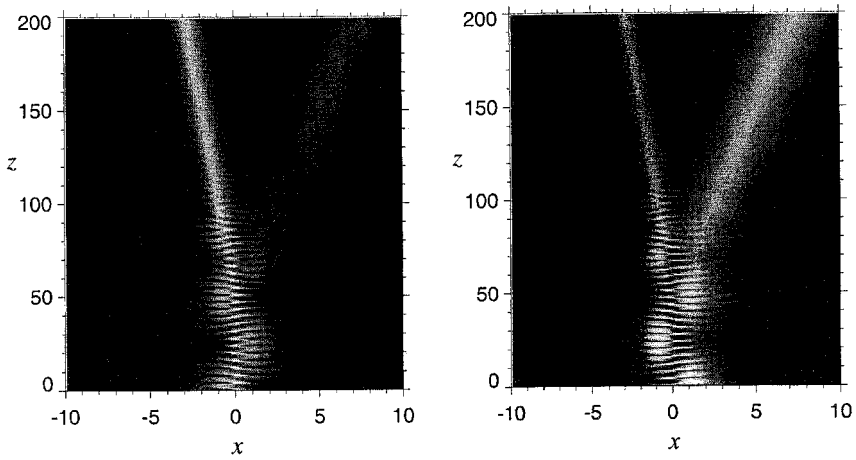

Fig. 5. The effect of wavelength detuning on the Y-junction. *Left*: numerical integration of the linear Schrödinger equation with $\beta = 1.1$ and initial condition $\Psi_{b,1}(x, 0)$. *Right*: the same for the initial condition $\Psi_{b,2}(x, 0)$. The nonadiabaticity parameter is $\epsilon^2 = 0.00125$. These pictures should be compared with Figs. 3 and 4.

making the assumption that $\xi = \rho\epsilon$, where ρ is considered to be a fixed parameter as ϵ tends to zero.

As explained in Section 4, the appropriate modified equation (in a slightly deformed transverse variable) is

$$i\frac{\partial f}{\partial z} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2} - (V_0(x, z) + W(x, z))f = 0, \tag{38}$$

where $V_0(x, z)$ is the potential for an adiabatic Y-junction planar waveguide (with nonadiabaticity parameter ϵ) the perturbation $W(x, z)$ has the form

$$\begin{aligned} W(x, z) &= (1 + \rho\epsilon)V_0(x/\sqrt{1 + \rho\epsilon}, z) - V_0(x, z) \\ &= \rho\epsilon\left(V_0(x, z) - \frac{x}{2}\frac{\partial V_0}{\partial x}(x, z)\right)(1 + O(\epsilon)). \end{aligned} \tag{39}$$

With ρ fixed as ϵ tends to zero, this perturbation to the potential is formally of size ϵ. Therefore, unlike the nonadiabatic correction terms on the right-hand side of (33) that are of size ϵ^2 and contribute negligibly to the dynamics for $z = O(\epsilon^{-1})$, the contribution from the change in operating wavelength is larger and will be important on these scales. Note that the $O(\epsilon)$ error in (39) is uniform in x because $V_0(x, z)$ and all of its x derivatives vanish exponentially for large x.

So, all of the terms on the right-hand side of (33) are uniformly $O(\epsilon)$, and the dominant terms are all contributed by the function $W(x, z)$ given by (39).

With these observations, we may begin the analysis of (33). Our initial conditions will be

$$B_{\mathrm{b},k}(0) = B_{\mathrm{b},k0}, \quad B_{\mathrm{r}}(0,\zeta) \equiv 0, \tag{40}$$

for some constants $B_{\mathrm{b},k0}$, so we assume only that there is initially no energy in the radiation spectrum. Assume power series expansions

$$B_{\mathrm{b},k} = B_{\mathrm{b},k}^{(0)} + \epsilon B_{b,k}^{(1)} + O(\epsilon^2), \quad B_{\mathrm{r}}(\zeta) = B_{\mathrm{r}}^{(0)}(\zeta) + O(\epsilon), \tag{41}$$

and introduce the slow scale $Z = \epsilon z$. For functions of both scales z and Z, we use the chain rule to replace

$$\frac{\partial}{\partial z} \to \frac{\partial}{\partial z} + \epsilon \frac{\partial}{\partial Z}. \tag{42}$$

Inserting these expansions into (33) and collecting powers of ϵ we find at leading order simply:

$$B_{\mathrm{b},k}^{(0)} = B_{\mathrm{b},k}^{(0)}(Z), \quad B_{\mathrm{r}}^{(0)}(\zeta) \equiv 0. \tag{43}$$

The vanishing of the radiation mode amplitude at this order is the only solution consistent with the initial conditions. At order $O(\epsilon)$, we obtain the equations:

$$\begin{aligned} \mathrm{i}\frac{\partial B_{\mathrm{b},1}^{(1)}}{\partial z} &= M_{11}^{(1)} B_{\mathrm{b},1}^{(0)}(Z) + M_{12}^{(1)} B_{\mathrm{b},2}^{(0)}(Z) - \mathrm{i}\frac{\partial B_{\mathrm{b},1}^{(0)}}{\partial Z}(Z), \\ \mathrm{i}\frac{\partial B_{\mathrm{b},2}^{(1)}}{\partial z} &= M_{21}^{(1)} B_{\mathrm{b},1}^{(0)}(Z) + M_{22}^{(1)} B_{\mathrm{b},2}^{(0)}(Z) - \mathrm{i}\frac{\partial B_{\mathrm{b},2}^{(0)}}{\partial Z}(Z), \end{aligned} \tag{44}$$

where $M_{jk}^{(1)}$ are the leading contributions of the matrix elements M_{jk} (see below). At this order, we also find an equation for the correction to the radiation mode amplitude that we will not need. Note that the result (43) from leading order implies that there are no contributions from the radiation mode amplitude B_{r} to these equations for the corrections to $B_{\mathrm{b},k}$ at this order.

To continue with the analysis, we need to have approximations – consistent with the assumptions about the functions $\zeta_1(z)$ and $\zeta_2(z)$ and convergent in the limit of ϵ going to zero – of the matrix elements M_{jk}. The approximations need to be uniformly valid as $\epsilon \to 0$ for long scales $z = O(\epsilon^{-1})$. Formally speaking, this means that we need to obtain approximations of the form

$$M_{jk}(z) = \epsilon M_{jk}^{(1)}(z, Z = \epsilon z)(1 + O(\epsilon)), \tag{45}$$

where the error term is uniformly small for all z and for Z held fixed. We already have from the estimates of the nonadiabatic contributions and from (39) that

$$\begin{aligned} M_{jk}(z) &= \langle \Psi_{\mathrm{b},j}(\cdot,z), W(\cdot,z)\Psi_{\mathrm{b},k}(\cdot,z)\rangle(1+O(\epsilon)) \\ &= \rho\epsilon\left\langle \Psi_{\mathrm{b},j}(\cdot,z), \left(V_0(\cdot,z) - \frac{(\cdot)}{2}\frac{\partial V_0}{\partial x}(\cdot,z)\right)\Psi_{\mathrm{b},k}(\cdot,z)\right\rangle(1+O(\epsilon)). \end{aligned} \tag{46}$$

At this point, we need corresponding approximations of the potential $V_0(x,z)$ and the bound modes $\Psi_{\mathrm{b},1}(x,z)$ and $\Psi_{\mathrm{b},2}(x,z)$ of the type described above that additionally are uniformly valid for all x since they appear inside the inner product integral. These approximations are of the form

$$\begin{aligned} V_0(x,z) &= \tilde{V}_0(x,z,Z)(1+O(\epsilon)), \\ \Psi_{\mathrm{b},k}(x,z) &= \tilde{\Psi}_{\mathrm{b},k}(x,z,Z)(1+O(\epsilon)), \end{aligned} \tag{47}$$

where the error terms are uniform in the sense described above. The required approximation $\tilde{V}_0$ is given exactly by the formula (16), wherein we simply substitute

$$G_k \to \exp(-2m_k(x-x_k(Z)))\exp(2\mathrm{i}m_k z + \mathrm{i}\theta_k), \quad \zeta_k \to \mathrm{i}m_k, \tag{48}$$

with the position $x_k(Z)$ being defined in terms of the velocity $v_k(Z)$ by

$$x_k(Z) := x_k(0) + \int_0^Z v_k(s)\mathrm{d}s, \tag{49}$$

and with the initial position and phase

$$x_k(0) := \frac{1}{2m_k}\log|g_0^{(k)}|, \quad \theta_k := \Im(\log(g_0^{(k)})). \tag{50}$$

Similar approximations $\tilde{\Psi}_{\mathrm{b},1}$ and $\tilde{\Psi}_{\mathrm{b},2}$ for the basis of bound modes are given by (20) along with the formulas (21) and (22) subject to the same substitutions (48). Note here that the normalization 'constants' in the denominators of (20) are indeed constant up to an $O(\epsilon)$ error term that is uniformly small for all $z = O(\epsilon^{-1})$. This procedure ultimately yields the required expressions:

$$M_{jk}^{(1)}(z,Z) = \rho\left\langle \tilde{\Psi}_{\mathrm{b},j}(\cdot,z,Z), \left(\tilde{V}_0(\cdot,z,Z) - \frac{(\cdot)}{2}\frac{\partial \tilde{V}_0}{\partial x}(\cdot,z,Z)\right)\tilde{\Psi}_{\mathrm{b},k}(\cdot,z,Z)\right\rangle. \tag{51}$$

For fixed Z, these functions are periodic in z with period $L = \pi/(m_1^2 - m_2^2)$.

Continuing with the solution of (44), we see that the corrections $B^{(1)}_{\mathrm{b},k}$ will only be bounded functions of z if the right-hand side has zero mean value. We therefore choose the dependence of $B^{(0)}_{\mathrm{b},k}$ on Z exactly to remove the mean secular terms. That is, we set

$$\begin{aligned} \mathrm{i}\frac{\partial B^{(0)}_{\mathrm{b},1}}{\partial Z}(Z) &= \overline{M^{(1)}_{11}}(Z)B^{(0)}_{\mathrm{b},1}(Z) + \overline{M^{(1)}_{12}}(Z)B^{(0)}_{\mathrm{b},2}(Z), \\ \mathrm{i}\frac{\partial B^{(0)}_{\mathrm{b},2}}{\partial Z}(Z) &= \overline{M^{(1)}_{21}}(Z)B^{(0)}_{\mathrm{b},1}(Z) + \overline{M^{(1)}_{22}}(Z)B^{(0)}_{\mathrm{b},2}(Z), \end{aligned} \tag{52}$$

where the mode coupling coefficients are defined by

$$\overline{M^{(1)}_{jk}}(Z) := \frac{1}{L}\int_0^L M^{(1)}_{jk}(z,Z)\mathrm{d}z. \tag{53}$$

It follows from the specification of the mode coupling coefficients that all of the dependence on Z enters via the two functions $x_1(Z)$ and $x_2(Z)$. Moreover, it is not hard to see that they only depend on the difference $y(Z) = x_2(Z) - x_1(Z)$ between these two functions. We may thus rewrite them as

$$\overline{M^{(1)}_{jk}}(Z) = \rho K_{jk}(y(Z); m_1, m_2), \tag{54}$$

where we have factored out the ratio parameter $\rho = \xi/\epsilon$. The advantage of this observation is that the same four universal functions $K_{jk}(y; m_1, m_2)$ describe the exchange of energy between the two bound modes regardless of the origin of the functions $x_1(Z)$ and $x_2(Z)$.

Computing these universal functions requires, according to (53) and (51), evaluating two integrals for each value of y and the parameters m_1 and m_2. In fact, by exchanging the order of integration, the averaging step (53) can be done first and moreover explicitly since the periodic dependence on z is quite simple; thus the integral (53) can be computed by residues. Rather than giving details here, we direct the interested reader to Besley *et al.* (2000) where a very similar (but more involved) calculation is presented. This analytical averaging step reduces the computation of the universal functions $K_{jk}(y; m_1, m_2)$ to a single integral (51) over $x \in \mathbb{R}$. For given values of the parameters m_1 and m_2, this integral may then be calculated numerically for each fixed value of y.

Fig. 6 shows a plot of the universal coupling coefficients for $m_1 = 1$ and $m_2 = 0.4$. The cross-coupling coefficients K_{12} and K_{21} vanish as $y \to \infty$ and the output waveguide ports separate. This is an expected result, since in this limit the approximate bound modes $\Psi_{\mathrm{b},k}(x,z)$ are localized in one or the other waveguide channel. These output channels are individually

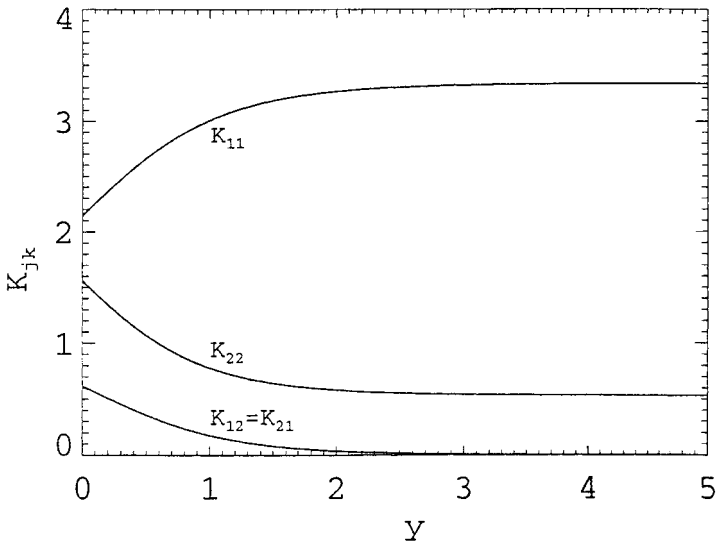

Fig. 6. The universal coupling functions $K_{jk}(y; 1, 0.4)$.

single-moded waveguides with distinct propagation constants, and it is well-known (Snyder and Love 1983) that if they are separated by a large transverse distance the coupling will be negligible.

It is also easy to see directly from the definition of the coupling coefficients that they are all real-valued, and that $K_{21} = K_{12}$. These features guarantee that the total bound power

$$P_{\mathrm{b}}^{(0)} := |B_{\mathrm{b},1}^{(0)}|^2 + |B_{\mathrm{b},2}^{(0)}|^2 \tag{55}$$

is a constant of motion under the asymptotic model (52). This statement means that the conservation of bound power only holds approximately in the true system (38) on length scales $Z = O(1)$ or $z = O(\epsilon^{-1})$. On longer scales, there will be significant losses to forward-propagating radiation modes.

We determine the behavior of the modes by integrating (52), which takes the form

$$\begin{aligned} \mathrm{i}\frac{\mathrm{d}B_{\mathrm{b},1}^{(0)}}{\mathrm{d}z}(z) &= \xi K_{11}(y(\epsilon z); m_1, m_2)B_{\mathrm{b},1}^{(0)}(z) + \xi K_{12}(y(\epsilon z); m_1, m_2)B_{\mathrm{b},2}^{(0)}(z), \\ \mathrm{i}\frac{\mathrm{d}B_{\mathrm{b},2}^{(0)}}{\mathrm{d}z}(z) &= \xi K_{21}(y(\epsilon z); m_1, m_2)B_{\mathrm{b},1}^{(0)}(z) + \xi K_{22}(y(\epsilon z); m_1, m_2)B_{\mathrm{b},2}^{(0)}(z), \end{aligned} \tag{56}$$

when we think of $B_{\mathrm{b},k}^{(0)}$ as functions of z rather than $Z = \epsilon z$. Although we assumed that ϵ and ξ were proportional to obtain (56), we may now simply take them to be independent. In fact, the physically important situation is to

consider ϵ fixed but small, which determines the geometry of the potential $V_0(x,z)$ and thus the waveguide dimensions, and then vary ξ. Since the ratio parameter $\rho = \xi/\epsilon$ was presumed fixed while $\epsilon \downarrow 0$ in out analysis, we expect that the integration of (56) should give valid results as long as ϵ is small and ξ is not too small or too large compared to ϵ. In Fig. 7, we compare the results of numerical integration of the model system (56) with direct numerical integration of (38) over a range of values of ξ for $\epsilon^2 = 0.00125$, $m_1 = 1$, and $m_2 = 0.4$. The initial condition for the Equation (38) is $f(x,0) = \Psi_{b,1}(x,0)$ and the corresponding initial condition for (56) is $B^{(0)}_{b,1}(0) = 1$ and

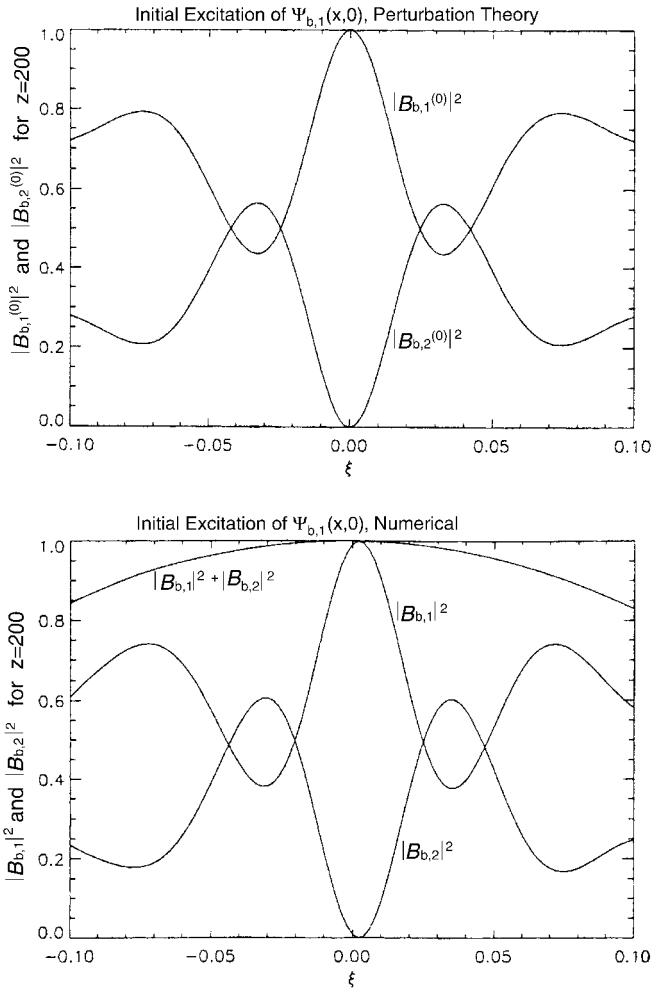

Fig. 7. Mode differentiation spectra for the solitonic Y-junction. *Top*: the result of integrating the system (52) for initial conditions $B^{(0)}_{b,1}(0) = 1$ and $B^{(0)}_{b,2}(0) = 0$. *Bottom*: the result of numerical integration of the partial differential equation for the initial conditions $f(x,0) = \Psi_{b,1}(x,0)$. In both cases, $\epsilon^2 = 0.00125$, $m_1 = 1.0$, $m_2 = 0.4$, and the integration is carried out to $z = 200$.

$B_{\mathrm{b},2}^{(0)}(0) = 0$. The plots show the values of $|B_{\mathrm{b},k}(z)|^2$ for $z = 200$, a position well beyond the opening of the branches of the Y-junction; this means that the relative velocity $y'(\epsilon z)$ is nearly saturated to its asymptotic constant value. It follows from the behavior of the universal coupling coefficients K_{jk} as functions of y, that for large z the two bound mode amplitudes become constant in modulus, and thus these mode differentiation spectra as computed for $z = 200$ are not expected to change further for larger z.

The fact that $P_{\mathrm{b}}^{(0)}$ is conserved by (56) is reflected in the fact that the two curves in the top plot add up to one. On the other hand, the higher-order scattering losses that are not captured by the system (56) are indeed evident in the bottom plot, where the bound power is plotted as well. Fig. 8 is similar to Fig. 7 but for an initial condition for (38) of $f(x,0) = \Psi_{\mathrm{b},2}(x,0)$ and correspondingly for (56) of $B_{\mathrm{b},1}^{(0)}(0) = 0$ and $B_{\mathrm{b},2}^{(0)}(0) = 1$.

The qualitative agreement between the predictions of the model system (56) and the numerical integration of the full Schrödinger Equation (38) is obvious from these figures. It is also clear that, as anticipated, the asymptotic results obtained from (56) are most accurate for $\xi \sim \epsilon$. As ξ increases for ϵ fixed, higher-order effects like radiative losses in the periodic trunk of the junction (these are quantified in Miller *et al.* (2000)) become more important. More subtle but also clear is the failure of the asymptotic theory of $\xi \ll \epsilon$. In this regime, the numerical simulations of (38) indicate a slight spectral shift of the peak of optimal performance. This effect is connected with the nonadiabatic terms formally of order $O(\epsilon^2)$ that were neglected, but that dominate in this regime.

Not only do these plots indicate the accuracy of the simple model (56) in comparison with a much more computationally intensive beam propagation, but they also show that in those experiments for which the signal must use a different wavelength from the pump beam that induces the Y-junction, some signal wavelengths are more desirable than others. It is clear that the most sharply defined mode splitting occurs in the neighborhood of the central peak in the transmission spectra, and therefore in such experiments, the best performance would be expected for signal wavelengths arbitrarily close to, but not equal to, the pump wavelength. The performance is therefore limited only by the spectral width of auxiliary filtering components used in the experiments. But reasonably good mode differentiation also occurs in certain sidebands of the transmission spectrum, the first of which occurs in these plots (that is, for the particular values of ϵ, m_1, and m_2 used here) near $\xi \approx \pm 0.075$. Of course, in those experiments in which four-wave mixing effects are made negligible, there is no difficulty with launching the signal beam in either trunk mode at exactly the optimal wavelength at the center of the transmission spectrum.

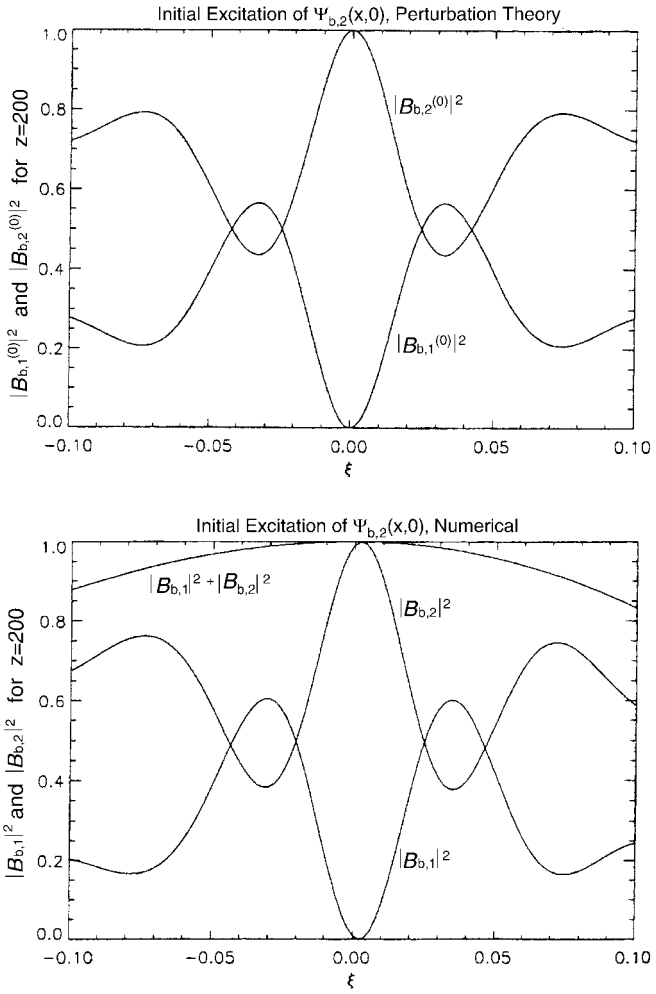

Fig. 8. Mode differentiation spectra for the solitonic Y-junction. *Top*: the result of integrating the system (52) for initial conditions $B_{b,1}^{(0)}(0) = 0$ and $B_{b,2}^{(0)}(0) = 1$. *Bottom*: the result of numerical integration of the partial differential equation for the initial conditions $f(x,0) = \Psi_{b,2}(x,0)$. In both cases, $\epsilon^2 = 0.00125$, $m_1 = 1.0$, $m_2 = 0.4$, and the integration is carried out to $z = 200$.

6. Conclusion

The solitonic waveguides are a family of special z-dependent potentials for the linear Schrödinger equation in which multi-moded beam propagation is completely lossless. The family of solitonic waveguides is indexed by parameters including the numbers ζ_k. For fixed values of the parameters, solitonic waveguides already exhibit complex spatial structure, allowing the design of multiport power dividing junctions, directional couplers, z-periodic channels, and so on. In this paper, we have shown that by allowing the

parameters to be adiabatic functions of z, in a way that is precise, the guidance properties are only slightly degraded over propagation distances in which the waveguide deformation is significant. We have applied these results to the analysis of a Y-junction waveguide that can be induced in a homogeneous nearly-Kerr nonlinear medium like PTS by the decay of a double soliton beam into two single solitons, showing that the device functions as an efficient mode splitter. In order to accommodate experimental situations where the signal beam must be of a different wavelength from the pump, we used perturbation theory to compute the effect of such detuning on the mode splitting characteristics, in the regime where the detuning is small, but still a larger effect than the nonadiabaticity of the junction. Efficient switching remains possible in a small neighborhood of the design wavelength as well as in the neighborhood of several side-peaks in the transmission spectrum.

As a concrete example, we have studied the particular adiabatic variation of the two-soliton waveguide that arises naturally in the nonlinear process of beam splitting in saturable materials, in which nonlinear propagation is modeled by the cubic–quintic nonlinear Schrödinger equation (35). This is because we believe that devices like this can already be implemented with current experimental techniques, and it will not be long before the leap is made from isolated laboratory experiments to practical, repeatable implementation.

On the other hand, we want to emphasize that our results all remain valid in the presence of arbitrary adiabatic variations of the solitonic waveguide parameters. The center of mass functions $x_1(Z)$ and $x_2(Z)$ do not have to come from a nonlinear theory for the resulting adiabatic solitonic waveguide to have useful mode splitting properties, as long as the nonadiabaticity parameter ϵ is sufficiently small. To demonstrate this point, we now give a design of an adiabatic solitonic Y-junction for which the two branch ports and the trunk port are all parallel outside the splitting region. We again choose $m_1 = 1$ and $m_2 = 0.4$ and $\epsilon^2 = 0.00125$. The two designs we propose are shown in Fig. 9. In the left-hand plot, we introduced adiabatic deformation of the parameters ζ_1 and ζ_2 with the use of the functions $x_1(Z) = -x_2(Z)$ where

$$x_2(Z) = \begin{cases} 0, & Z \le 2 \\ T_{\text{linear}}(Z), & 2 \le Z \le 5 \\ 5, & Z \ge 5, \end{cases} \tag{57}$$

with the transition function

$$T_{\text{linear}}(Z) := \frac{5}{3}(Z - 2). \tag{58}$$

These displacements provide a straight-line transition between two exact z-periodic solitonic waveguides, one in which the centers of mass are identical and thus the waveguide is obviously periodic, and one in which the centers of

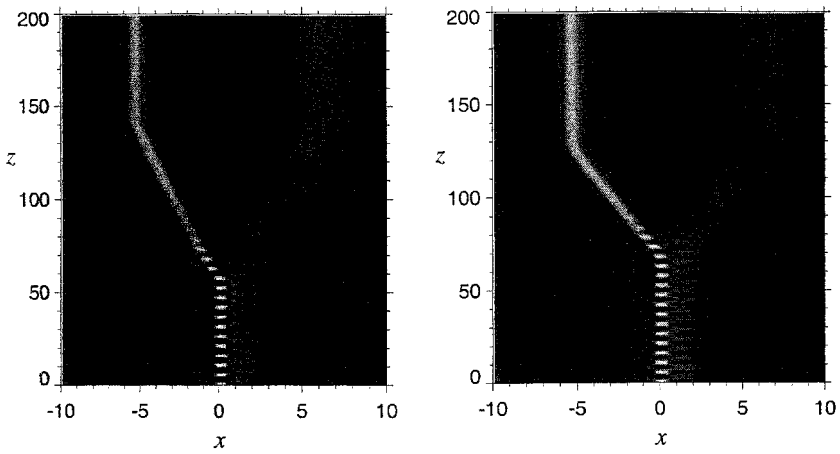

Fig. 9. Two adiabatic mode-splitting solitonic Y-junctions with parallel trunks and branches. *Left*: angular branching. *Right*: infinitely smooth branching. In both cases, $\epsilon^2 = 0.00125$.

mass are far apart, and thus the periodicity is less obvious. A smoother version of the same transition is given in the right-hand plot. Here, we used infinitely smooth functions $x_1(Z) = -x_2(Z)$ obtained by simply replacing the straight-line transition function by

$$T_{\text{smooth}}(Z) := \frac{5}{2}\left(1 + \tanh\left(\tan\left(\frac{\pi}{2}\left(\frac{2}{3}(Z-2) - 1\right)\right)\right)\right). \tag{59}$$

The pictures are qualitatively very similar, suggesting that not much is gained by the smoothness of the transition, as long as the derivatives of the functions $x_k(Z)$ (these are just the velocities $v_k(Z)$) are bounded. Then one may take ϵ sufficiently small for fixed coordinate functions and achieve arbitrarily accurate mode splitting. For both of these Y-junction waveguides, the mode splitting efficiency at this value of ϵ is expected to be qualitatively similar to that illustrated in Figs. 3 and 4 for the same excitations of the trunk channel.

Appendix A: Coupled-mode equations for deformed solitonic potentials

In this appendix, we show how the coupled-mode equations (14) are altered by allowing the parameters $\zeta_1, \ldots, \zeta_M$ to differ from constants as functions of z, and the vectors $\vec{g}^{(1)}, \ldots, \vec{g}^{(M)}$ to differ from pure exponentials. We begin by observing that in the fixed-parameter case described in Section 2, the function $a(x, z, \zeta)$ satisfies for each complex ζ the equation

$$\mathrm{i}\frac{\partial a}{\partial z} + \frac{1}{2}\frac{\partial^2 a}{\partial x^2} - V_0 a = -2\zeta^2 a. \tag{60}$$

Next we observe that in the fixed-parameter case, $a(x, z, \zeta)$ only depends on z through the quantities $\vec{g}^{(k)}$ which satisfy

$$\frac{\mathrm{d}\vec{g}^{(k)}}{\mathrm{d}z}(z) = -2\mathrm{i}\zeta_k^{*2}\vec{g}^{(k)}(z). \tag{61}$$

This means that (60) can be rewritten as

$$2\sum_{k=1}^{M}\left(\zeta_k^{*2}\sum_{n=1}^{N} g_n^{(k)}\frac{\partial a}{\partial g_n^{(k)}} - \zeta_k^2\sum_{n=1}^{N} g_n^{(k)*}\frac{\partial a}{\partial g_n^{(k)*}}\right) + \frac{1}{2}\frac{\partial^2 a}{\partial x^2} - V_0 a = -2\zeta^2 a. \tag{62}$$

There is no direct reference to any z-dependence in this equation. By its construction, a may be considered to be a complex analytic function of x, ζ, and quantities $\vec{g}^{(k)}$ and ζ_k and their complex conjugates for $k = 1, \ldots, M$:

$$a = a(x, \zeta, \{\zeta_k\}, \{\zeta_k^*\}, \{\vec{g}^{(k)}\}, \{\vec{g}^{(k)*}\}). \tag{63}$$

The potential V_0 depends on all of these quantities except for ζ:

$$V_0 = V_0(x, \{\zeta_k\}, \{\zeta_k^*\}, \{\vec{g}^{(k)}\}, \{\vec{g}^{(k)*}\}). \tag{64}$$

Equation (62) is therefore an identity that is a consequence of the algebraic construction alone.

This means that (62) continues to hold when the quantities $\{\vec{g}^{(j)}\}$, $\{\zeta_j\}$, as well as ζ are taken to be arbitrary functions of the longitudinal parameter z. In this more general case, the chain rule applied to (63) gives

$$\frac{\partial a}{\partial z} = \frac{\mathrm{d}\zeta}{\mathrm{d}z}\frac{\partial a}{\partial \zeta} + \sum_{k=1}^{M}\left(\frac{\mathrm{d}\zeta_k}{\mathrm{d}z}\frac{\partial a}{\partial \zeta_k} + \frac{\mathrm{d}\zeta_k^*}{\mathrm{d}z}\frac{\partial a}{\partial \zeta_k^*} + \sum_{n=1}^{N}\left(\frac{\mathrm{d}g_n^{(k)}}{\mathrm{d}z}\frac{\partial a}{\partial g_n^{(k)}} + \frac{\mathrm{d}g_n^{(k)*}}{\mathrm{d}z}\frac{\partial a}{\partial g_n^{(k)*}}\right)\right). \tag{65}$$

Comparing this expression with (62), we can write a Schrödinger-like equation that holds in the deformed case:

$$\begin{aligned}\mathrm{i}\frac{\partial a}{\partial z} + 2\zeta^2 a + \frac{1}{2}\frac{\partial^2 a}{\partial x^2} - V_0 a = {} & \mathrm{i}\frac{\mathrm{d}\zeta}{\mathrm{d}z}\frac{\partial a}{\partial \zeta} + \mathrm{i}\sum_{k=1}^{M}\left(\frac{\mathrm{d}\zeta_k}{\mathrm{d}z}\frac{\partial a}{\partial \zeta_k} + \frac{\mathrm{d}\zeta_k^*}{\mathrm{d}z}\frac{\partial a}{\partial \zeta_k^*}\right) \\ & + \sum_{k=1}^{M}\sum_{n=1}^{N}\left(\left(\mathrm{i}\frac{\mathrm{d}g_n^{(k)}}{\mathrm{d}z} - 2\zeta_k^{*2} g_n^{(k)}\right)\frac{\partial a}{\partial g_n^{(k)}}\right. \\ & \left. + \left(\mathrm{i}\frac{\mathrm{d}g_n^{(k)*}}{\mathrm{d}z} + 2\zeta_k^2 g_n^{(k)*}\right)\frac{\partial a}{\partial g_n^{(k)*}}\right).\end{aligned} \tag{66}$$

In this equation, the terms on the left-hand side may all be considered to be functions of the independent variables x and z. By contrast, the terms on the right-hand side involve derivatives with respect to quantities that are related to the algebraic structure of a, that is, derivatives with respect to the variables indicated in (63).

As explained in Section 3, the radiation mode function $\Psi_r(x, z, \zeta)$ for ζ real (independent of z) is obtained, just as in the undeformed case, from the formula (9). This leads to the partial differential equation satisfied by the radiation mode function:

$$\begin{aligned} &\mathrm{i}\frac{\partial \Psi_r}{\partial z} + \frac{1}{2}\frac{\partial^2 \Psi_r}{\partial x^2} - V_0 \Psi_r \\ &\quad = \mathrm{i}\frac{\Psi_r}{2}\sum_{k=1}^{M}\left((\zeta - \zeta_k)^{-1}\frac{\mathrm{d}\zeta_k}{\mathrm{d}z} + (\zeta - \zeta_k^*)^{-1}\frac{\mathrm{d}\zeta_k^*}{\mathrm{d}z}\right) \\ &\qquad + \frac{\mathrm{i}\exp(-2\mathrm{i}\zeta^2 z)}{U(\zeta, \{\zeta_k\}, \{\zeta_k^*\})}\sum_{k=1}^{M}\left(\frac{\mathrm{d}\zeta_k}{\mathrm{d}z}\frac{\partial a}{\partial \zeta_k} + \frac{\mathrm{d}\zeta_k^*}{\mathrm{d}z}\frac{\partial a}{\partial \zeta_k^*}\right) + \frac{\exp(-2\mathrm{i}\zeta^2 z)}{U(\zeta, \{\zeta_k\}, \{\zeta_k^*\})} \\ &\qquad \times \sum_{k=1}^{M}\sum_{n=1}^{N}\left(\left(\mathrm{i}\frac{\mathrm{d}g_n^{(k)}}{\mathrm{d}z} - 2\zeta_k^{*2} g_n^{(k)}\right)\frac{\partial a}{\partial g_n^{(k)}} + \left(\mathrm{i}\frac{\mathrm{d}g_n^{(k)*}}{\mathrm{d}z} + 2\zeta_k^2 g_n^{(k)*}\right)\frac{\partial a}{\partial g_n^{(k)*}}\right), \end{aligned} \tag{67}$$

where

$$U(\zeta, \{\zeta_k\}, \{\zeta_k^*\}) = \left(\pi \prod_{k=1}^{M} |\zeta - \zeta_k|^2\right)^{1/2}. \tag{68}$$

The terms containing a on the right-hand side could be re-written in terms of Ψ_r, but this is not essential for the observations to follow. For the bound modes of the two-soliton potential ($M = 2$ and $N = 1$), we use the bi-orthogonality relation to choose

$$\Psi_{b,1} = c_1 a(x, \zeta_1^*, \{\zeta_k\}, \{\zeta_k^*\}, \{g^{(k)}\}, \{g^{(k)*}\}) \exp\left(-2\mathrm{i}\int_0^z \zeta_1(s)^{*2}\mathrm{d}s\right), \tag{69}$$

and

$$\Psi_{b,2} = c_2 a(x, \zeta_2, \{\zeta_k\}, \{\zeta_k^*\}, \{g^{(k)}\}, \{g^{(k)*}\}) \exp\left(-2\mathrm{i}\int_0^z \zeta_2(s)^2\mathrm{d}s\right), \tag{70}$$

where c_1 and c_2 are chosen so that $\|\Psi_{b,k}\|_2 = 1$ for $k = 1, 2$. They depend only on ζ_1 and ζ_2, and therefore possibly on z. This leads to the equations satisfied by the bound modes:

$$\mathrm{i}\frac{\partial \Psi_{\mathrm{b},1}}{\partial z}+\frac{1}{2}\frac{\partial^2 \Psi_{\mathrm{b},1}}{\partial x^2}-V_0\Psi_{\mathrm{b},1}=\mathrm{i}\Psi_{\mathrm{b},1}\frac{\mathrm{d}}{\mathrm{d}z}\log c_1+c_1\exp\left(-2\mathrm{i}\int_0^z \zeta_1(s)^{*2}\mathrm{d}s\right)$$
$$\times\left[\mathrm{i}\frac{\mathrm{d}\zeta_1^*}{\mathrm{d}z}\frac{\partial a}{\partial \zeta}+\mathrm{i}\sum_{k=1}^{M}\left(\frac{\mathrm{d}\zeta_k}{\mathrm{d}z}\frac{\partial a}{\partial \zeta_k}+\frac{\mathrm{d}\zeta_k^*}{\mathrm{d}z}\frac{\partial a}{\partial \zeta_k^*}\right)\right.$$
$$+\sum_{k=1}^{M}\sum_{n=1}^{N}\left(\left(\mathrm{i}\frac{\mathrm{d}g_n^{(k)}}{\mathrm{d}z}-2\zeta_k^{*2}g_n^{(k)}\right)\frac{\partial a}{\partial g_n^{(k)}}\right.$$
$$\left.\left.+\left(\mathrm{i}\frac{\mathrm{d}g_n^{(k)*}}{\mathrm{d}z}+2\zeta_k^2 g_n^{(k)*}\right)\frac{\partial a}{\partial g_n^{(k)*}}\right)\right], \tag{71}$$

where on the right-hand side a is evaluated for $\zeta=\zeta_1^*$, and

$$\mathrm{i}\frac{\partial \Psi_{\mathrm{b},2}}{\partial z}+\frac{1}{2}\frac{\partial^2 \Psi_{\mathrm{b},2}}{\partial x^2}-V_0\Psi_{\mathrm{b},2}=\mathrm{i}\Psi_{\mathrm{b},2}\frac{\mathrm{d}}{\mathrm{d}z}\log c_2+c_2\exp\left(-2\mathrm{i}\int_0^z \zeta_2(s)^{2}\mathrm{d}s\right)$$
$$\times\left[\mathrm{i}\frac{\mathrm{d}\zeta_2}{\mathrm{d}z}\frac{\partial a}{\partial \zeta}+\mathrm{i}\sum_{k=1}^{M}\left(\frac{\mathrm{d}\zeta_k}{\mathrm{d}z}\frac{\partial a}{\partial \zeta_k}+\frac{\mathrm{d}\zeta_k^*}{\mathrm{d}z}\frac{\partial a}{\partial \zeta_k^*}\right)\right.$$
$$+\sum_{k=1}^{M}\sum_{n=1}^{N}\left(\left(\mathrm{i}\frac{\mathrm{d}g_n^{(k)}}{\mathrm{d}z}-2\zeta_k^{*2}g_n^{(k)}\right)\frac{\partial a}{\partial g_n^{(k)}}\right.$$
$$\left.\left.+\left(\mathrm{i}\frac{\mathrm{d}g_n^{(k)*}}{\mathrm{d}z}+2\zeta_k^2 g_n^{(k)*}\right)\frac{\partial a}{\partial g_n^{(k)*}}\right)\right], \tag{72}$$

where on the right-hand side a is evaluated for $\zeta=\zeta_2$.

The aim of writing down the exact equations satisfied by the basic functions is to use them in expressing the linear Schrödinger equation (10) as a coupled-mode system, by projecting onto the basis functions. Restricting to the case of $M=2$ and $N=1$, we first expand the solution of (10) in terms of the (deforming) orthonormal basis:

$$f(x,z)=B_{\mathrm{b},1}(z)\Psi_{\mathrm{b},1}(x,z)+B_{\mathrm{b},2}(z)\Psi_{\mathrm{b},2}(x,z)+\int_{-\infty}^{\infty}B_{\mathrm{r}}(z,\zeta)\Psi_{\mathrm{r}}(x,z,\zeta)\mathrm{d}\zeta. \tag{73}$$

The coefficients are simply

$$B_{\mathrm{b},k}(z)=\langle\Psi_{\mathrm{b},k}(\cdot,z),f(\cdot,z)\rangle,\quad B_{\mathrm{r}}(z,\zeta)=\langle\Psi_{\mathrm{r}}(\cdot,z,\zeta),f(\cdot,z)\rangle. \tag{74}$$

Next, insert this expansion into (10) and use the equations satisfied by the approximate modes to eliminate the z-derivatives of Ψ_{r} and $\Psi_{\mathrm{b},k}$ wherever

they appear. Finally, use the exact orthogonality conditions to separate the modes and therefore deduce (coupled) equations for the mode amplitudes $B_{\mathrm{b},1}(z)$, $B_{\mathrm{b},2}(z)$, and $B_{\mathrm{r}}(z,\zeta)$. We introduce some simplifying notation: let $R_1(x,z)$ denote the right-hand side of the evolution equation (71) for $\Psi_{\mathrm{b},1}$, let $R_2(x,z)$ denote the right-hand side of the evolution equation (72) for $\Psi_{\mathrm{b},2}$, and let $R(x,z,\zeta)$ denote the right-hand side of the evolution equation (67) for Ψ_{r}. Then, the *exact* equations of motion for the mode amplitude take the form of the system

$$\begin{aligned}
\mathrm{i}\frac{\mathrm{d}B_{\mathrm{b},1}}{\mathrm{d}z}(z) &= M_{11}^{(\mathrm{NA})}(z)B_{\mathrm{b},1}(z) + M_{12}^{(\mathrm{NA})}(z)B_{\mathrm{b},2}(z) + \int_{-\infty}^{\infty} N_1^{(\mathrm{NA})}(z,\zeta)B_{\mathrm{r}}(z,\zeta)\mathrm{d}\zeta \\
\mathrm{i}\frac{\mathrm{d}B_{\mathrm{b},2}}{\mathrm{d}z}(z) &= M_{21}^{(\mathrm{NA})}(z)B_{\mathrm{b},1}(z) + M_{22}^{(\mathrm{NA})}(z)B_{\mathrm{b},2}(z) + \int_{-\infty}^{\infty} N_2^{(\mathrm{NA})}(z,\zeta)B_{\mathrm{r}}(z,\zeta)\mathrm{d}\zeta \\
\mathrm{i}\frac{\partial B_{\mathrm{r}}}{\partial z}(z,\zeta) &= O_1^{(\mathrm{NA})}(z,\zeta)B_{\mathrm{b},1}(z) + O_2^{(\mathrm{NA})}(z,\zeta)B_{\mathrm{b},2}(z) \\
&\quad + \int_{-\infty}^{\infty} K^{(\mathrm{NA})}(z,\zeta,\eta)B_{\mathrm{r}}(z,\eta)\mathrm{d}\eta,
\end{aligned} \tag{75}$$

where for $k = 1, 2$ and $j = 1, 2$,

$$M_{jk}^{(\mathrm{NA})}(z) := -\langle \Psi_{\mathrm{b},j}(\cdot,z), R_k(\cdot,z)\rangle, \tag{76}$$

for $k = 1$ and $k = 2$,

$$\begin{aligned}
N_k^{(\mathrm{NA})}(z,\zeta) &:= -\langle \Psi_{\mathrm{b},k}(\cdot,z), R(\cdot,z,\zeta)\rangle, \\
O_k^{(\mathrm{NA})}(z,\zeta) &:= -\langle \Psi_{\mathrm{r}}(\cdot,z,\zeta), R_k(\cdot,z)\rangle,
\end{aligned} \tag{77}$$

and finally,

$$K^{(\mathrm{NA})}(z,\zeta,\eta) := -\langle \Psi_{\mathrm{r}}(\cdot,z,\zeta), R(\cdot,z,\eta)\rangle. \tag{78}$$

Note that according to the orthogonality relation, the kernel $K^{(\mathrm{NA})}(z,\zeta,\eta)$ contains a singular term proportional to $\delta(\zeta-\eta)$ coming from the first term on the right-hand side of (67).

For the adiabatic applications in the text, the crucial observation is that all of these coefficients are uniformly bounded in z when the deviations from the constant-parameter case are controlled. In particular, for each (operator-valued) coefficient in the coupled-mode system, say C, there is a corresponding constant K independent of all parameters, such that

$$\|C\| \leq K \sup_{z\in\mathbb{R}} \left(\left|\frac{\mathrm{d}\zeta_1}{\mathrm{d}z}\right| + \left|\frac{\mathrm{d}\zeta_2}{\mathrm{d}z}\right| + \left|\frac{\mathrm{d}g^{(1)}}{\mathrm{d}z} + 2\mathrm{i}\zeta_1^{*2} g^{(1)}\right| + \left|\frac{\mathrm{d}g^{(2)}}{\mathrm{d}z} + 2\mathrm{i}\zeta_2^{*2} g^{(2)}\right| \right), \tag{79}$$

where $\|\cdot\|$ indicates an appropriate operator norm. This bound is the basis for the analysis in the main text.

References

Artigas, D., L. Torner, J.P. Torres and N.N Akhmediev. *Opt. Commun.* **143** 322, 1997.
Besley, J.A., N.N. Akhmediev and P.D. Miller. *Opt. Lett.* **22** 1162, 1997.
Besley, J.A., N.N. Akhmediev and P.D. Miller. *Stud. Appl. Math.* **101** 343, 1998a.
Besley, J.A., J.D. Love and W. Langer. *J. Lightwave Technol.* **16** 678, 1998b.
Besley, J.A., P.D. Miller and N.N. Akhmediev. *Phys. Rev. E* **61** 7121, 2000.
Chen, Z., M. Mitchell and M. Segev. *Opt. Lett.* **21** 716, 1996.
De La Fuente, R., A. Barthelemy and C. Froehly. *Opt. Lett.* **16** 793, 1991.
Henry, W.M. and J.D. Love. *Opt. Quantum Electron.* **29** 379, 1997.
Hsu, J.-M. and C.-T. Lee. *IEEE J. Quantum Electron.* **34** 673, 1998.
Kang, J.U., G.I. Stegeman, J.S. Aitchison and N. Akhmediev. *Phys. Rev. Lett.* **76** 3699, 1996.
Kapon, E. and R.N. Thurston. *Appl. Phys. Lett.* **50** 1710, 1987.
Lawrence, B.L., W.E. Torruellas, M. Cha, M.L. Sundheimer, G.I. Stegeman, J. Meth, S. Etemad and G. Baker. *Phys. Rev. Lett.* **73** 597, 1994.
Love, J.D., R.W.C. Vance and A. Joblin. *Opt. Quantum Electron.* **28** 353, 1996.
Luther-Davies, B. and Y. Xiaoping. *Opt. Lett.* **17** 496, 1992a.
Luther-Davies, B. and Y. Xiaoping. *Opt. Lett.* **17** 1755, 1992b.
Marcatili, E.A.J. *IEEE J. Quantum Electron.* **21** 307, 1985.
Miller, P.D. *Phys. Rev. E* **53** 4137, 1996.
Miller, P.D. and N.N. Akhmediev. *Phys. Rev. E* **53** 4096, 1996.
Miller, P.D. and N.N. Akhmediev. *Physica D* **123** 513, 1998.
Miller, P.D., A. Soffer and M.I. Weinstein. *Nonlinearity* **13** 507, 2000.
Negami, T., H. Haga and S. Yamamoto. *Appl. Phys. Lett.* **54** 1080, 1989.
Ni, T.-D., D. Sturzebecher, M. Cummings and B. Perlman. *Appl. Phys. Lett.* **67** 1651, 1995.
Shirafuji, K. and S. Kurazono. *J. Lightwave Technol.* **9** 426, 1991.
Snyder, A.W. and J.D. Love. *Optical Waveguide Theory*, Chapman & Hall, London, 1983.
Weissman, Z., E. Marom and A. Hardy. *Opt. Lett.* **14** 293, 1989.
Weissman, A., D. Nir, S. Ruschin and A. Hardy. *Appl. Phys. Lett.* **67** 302, 1995.
Yuan, B., X. Song and H. Hu. *Opt. Commun.* **107** 205, 1994.
Zekharov, V.E. and A.V. Shabat. *Sov. Phys. JETP* **34** 62, 1972.