Available online at www.sciencedirect.com
ScienceDirect
Journal of Computational Physics 226 (2007) 1693–1709
JOURNAL OF COMPUTATIONAL PHYSICS
www.elsevier.com/locate/jcp

A mode elimination technique to improve convergence of iteration methods for finding solitary waves

T.I. Lakoba *, J. Yang [1]

Department of Mathematics and Statistics, 16 Colchester Ave., University of Vermont, Burlington, VT 05401, USA

Received 5 November 2006; received in revised form 24 May 2007; accepted 6 June 2007
Available online 27 June 2007

Abstract

We extend the key idea behind the generalized Petviashvili method of [T.I. Lakoba, J. Yang, A generalized Petviashvili iteration method for scalar and vector Hamiltonian equations with arbitrary form of nonlinearity, J. Comput. Phys., this issue, doi:10.1016/j.jcp.2007.06.009] by proposing a novel technique based on a similar idea. This technique systematically eliminates from the iteratively obtained solution a mode that is "responsible" either for the divergence or the slow convergence of the iterations. We demonstrate, theoretically and with examples, that this mode elimination technique can be used both to obtain some nonfundamental solitary waves and to considerably accelerate convergence of various iteration methods. As a collateral result, we compare the linearized iteration operators for the generalized Petviashvili method and the well-known imaginary-time evolution method and explain how their different structures account for the differences in the convergence rates of these two methods.

MSC: 35Qxx; 65B99; 65N99; 78A40; 78A99

Keywords: Nonlinear evolution equations; Solitary waves; Iteration methods; Convergence acceleration

1. Introduction

In the companion paper [1], we proposed a generalization of the Petviashvili iteration method for finding stationary solitary waves $u(\mathbf{x})$ of scalar and vector Hamiltonian equations with an arbitrary form of nonlinearity:

$$-Mu + F(\mathbf{x}, u) = 0, \quad u(|\mathbf{x}| \to \infty) \to 0, \tag{1.1}$$

DOI of original article: 10.1016/j.jcp.2007.06.009.
* Corresponding author. Tel.: +1 802 656 2610.
E-mail addresses: lakobati@cems.uvm.edu (T.I. Lakoba), jyang@cems.uvm.edu (J. Yang).
[1] Tel.: +1 802 656 4314.

0021-9991/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jcp.2007.06.010

where M is a self-adjoint differential operator and, in the vector case, the nonlinear term must satisfy a condition $\partial F_i/\partial u_j = \partial F_j/\partial u_i$. (Recall that the original Petviashvili method [2] was proposed for scalar equations with power-law nonlinearity $F(\mathbf{x}, u) = u^p$.) A common form of operator M (in the scalar case) is

$$M = \mu - \nabla^2, \tag{1.2}$$

where μ is the propagation constant of the solitary wave. Thus, the generalized Petviashvili method, that obtains solutions with a specified propagation constant, can be applied to the same class of equations as the well-known imaginary-time evolution method (ITEM) (see e.g. [3–6]) that is used to find solitary waves with a specified power.

In the present work, we extend the results of [1] as follows. In Section 2, we establish a mathematical relation between the generalized Petviashvili method and the ITEM. This discussion will also set the stage for the main result of this work, presented in Section 3. There, we develop the ideas behind the original and generalized Petviashvili methods [7,1] and propose a new technique that we refer to as the mode elimination. This technique can be used to obtain nonfundamental (see below) solitary waves, which the methods of [1–6] cannot obtain (the iterations would diverge). However, since alternative methods of obtaining nonfundamental solitary waves exist [3,8], we see the *main use* of the mode elimination in that it can considerably accelerate convergence of various iteration methods. The corresponding examples are presented in Section 4, and the summary of our results is given in Section 5.

2. Convergence rates of the Petviashvili and the imaginary-time evolution methods

In this section, we will compare the convergence properties of the generalized Petviashvili method [1] with those of the accelerated ITEM proposed in Ref. [6]. This discussion will highlight a feature of the generalized Petviashvili iteration scheme that will be important when we present our main result – the mode elimination technique – in the next section. Of the two versions of the ITEM (with power and amplitude normalizations) considered in [6], we will focus on the one with power normalization, because its linearized operator can be readily compared with that of the generalized Petviashvili method. In order not to obscure the main ideas by technical details, we restrict our presentation to the case of a single real-valued equation (1.1) with M given by Eq. (1.2), i.e., to:

$$\nabla^2 u + F(\mathbf{x}, u) = \mu u. \tag{2.1}$$

It is well known that the convergence of an iteration method is determined by the properties of the *linearized* iteration equation. Namely, let u_n be the solution obtained at the nth iteration, and let the "error" $\tilde{u}_n$ be defined as

$$\tilde{u}_n = u_n - u, \quad |\tilde{u}_n| \ll |u|. \tag{2.2}$$

As will be shown below, it satisfies a linearized iteration equation of the form

$$\tilde{u}_{n+1} = (1 + \Delta\tau\mathcal{L})\tilde{u}_n, \quad \Delta\tau > 0, \tag{2.3}$$

where $\mathcal{L}$ is the linear operator that results when the iteration method is linearized on the background of the solitary wave u, and $\Delta\tau$ is an auxiliary scaling parameter. From a conceptual point of view, the presence of $\Delta\tau$ emphasizes the analogy of iteration methods with numerical methods of solving time-dependent differential equations (see e.g. [9]); from a practical point of view, it can be used to ensure (in certain cases) or optimize the convergence of the method, as we will discuss later on.

Let us begin with general remarks regarding the convergence rate of the linearized iteration equation (2.3). Suppose that the eigenfunctions of $\mathcal{L}$ form a complete set in an appropriate functional space, so that $\tilde{u}_n$ can be expanded over them. Let the minimum and maximum eigenvalues of $\mathcal{L}$ be $\Lambda_{\min}$ and $\Lambda_{\max}$. Then the convergence rate of the iteration method can be defined as $\log(1/R)$, where the convergence factor R is the maximum (in magnitude) eigenvalue of the operator on the r.h.s. of (2.3):

$$R = \max\{|1 + \Lambda_{\max}\Delta\tau|,\ |1 + \Lambda_{\min}\Delta\tau|\}. \tag{2.4}$$

Clearly, $R < 1$ needs to hold in order for the iterations to converge, which implies

$$\Lambda_{\max} \leqslant 0 \quad \text{and} \quad 1 + \Lambda_{\min}\Delta\tau > -1. \tag{2.5}$$

Moreover, if $\Lambda_{\max} = 0$, then the corresponding eigenfunction of $\mathcal{L}$ needs to be a translational eigenmode (if it exists) of the linearized Eq. (2.1), which only shifts the solution in space and hence does not affect the convergence of the method. The smaller the convergence factor R, the faster the convergence. It can be readily shown [6] that the minimum value of R occurs at

$$\Delta\tau_* = \frac{-2}{\Lambda_{\min} + \Lambda_{\max}} \tag{2.6}$$

(recall that $\Lambda_{\min} < \Lambda_{\max} < 0$) and equals

$$R_* = \frac{1 - (\Lambda_{\max}/\Lambda_{\min})}{1 + (\Lambda_{\max}/\Lambda_{\min})}. \tag{2.7}$$

Therefore, the closer the ratio $(\Lambda_{\max}/\Lambda_{\min})$ to 1, the faster the convergence of the iteration method. Below we will compare the possible values of this ratio for the generalized Petviashvili method and the accelerated ITEM. To that end, we first need to cast the linearizations of these methods into the form of Eq. (2.3).

Let L_0 denote the nonlinear operator of Eq. (2.1) (or, more generally, of the stationary wave equation whose solution we are looking for), so that that equation is rewritten as

$$L_0 u = 0. \tag{2.8}$$

Let L be the corresponding linearized operator, so that

$$L_0(u + \tilde{u}) \equiv L_0 u + L\tilde{u} = L\tilde{u} \quad \text{for any } |\tilde{u}| \ll |u|. \tag{2.9}$$

Note that for Hamiltonian wave equations, L is always self-adjoint. With these notations, the generalized Petviashvili method is [1]:

$$u_{n+1} - u_n = \left(N^{-1}(L_0 u)_n - \gamma \frac{\langle u_n, (L_0 u)_n \rangle}{\langle u_n, N u_n \rangle} u_n \right) \Delta\tau, \tag{2.10}$$

where

$$\gamma = 1 + \frac{1}{\alpha\Delta\tau}. \tag{2.11}$$

Here and below, the inner product between two real-valued functions is defined in a standard way:

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(\mathbf{x}) g(\mathbf{x})\, d\mathbf{x}.$$

For the positive definite and self-adjoint operator N in (2.10), we take the simplest form used in [1]:

$$N = c - \nabla^2, \tag{2.12}$$

where the constant c is given by Eq. (3.11) of [1]. The constant α in Eq. (2.11) above is such that

$$Lu \approx \alpha N u \tag{2.13}$$

in a certain least-square sense; a formula for computing this constant at each iteration can be found in either of Eqs. (3.12) or (3.15) of [1], but will not be needed here. Following the steps of a calculation found at the beginning of Section 2 of [1], it is straightforward to show that the linearized form of the generalized Petviashvili method is:

$$\tilde{u}_{n+1} - \tilde{u}_n = \left(N^{-1} L\tilde{u}_n - \gamma \frac{\langle u, L\tilde{u}_n \rangle}{\langle u, N u \rangle} u \right) \Delta\tau. \tag{2.14}$$

Next, the accelerated ITEM of Ref. [6] is:

$$u_{n+1} = \left[\frac{P}{\langle \hat{u}_{n+1}, \hat{u}_{n+1}\rangle}\right]^{\frac{1}{2}} \hat{u}_{n+1}, \tag{2.15}$$

$$\hat{u}_{n+1} - u_n = K^{-1}(\nabla^2 u_n + F(\mathbf{x}, u_n) - \mu_n u_n)\Delta\tau, \tag{2.16}$$

$$\mu_n = \frac{\langle \nabla^2 u_n + F(\mathbf{x}, u_n), K^{-1}u_n\rangle}{\langle u_n, K^{-1}u_n\rangle}, \tag{2.17}$$

where $P = \int_{-\infty}^{\infty} u^2\, d\mathbf{x}$ is the specified power of the solitary wave. The positive definite and self-adjoint operator K is referred to as the acceleration operator for the ITEM [3,6]. For simplicity, we take K to have the same form (2.12) as the operator N in the generalized Petviashvili method, with the c being now an arbitrary positive constant. The linearized form of ITEM (2.15)–(2.17) is [6]:

$$\tilde{u}_{n+1} - \tilde{u}_n = \left(K^{-1}L\tilde{u}_n - \frac{\langle u, K^{-1}L\tilde{u}_n\rangle}{\langle u, K^{-1}u\rangle}K^{-1}u\right)\Delta\tau. \tag{2.18}$$

Thus, the "primordial" operator in the linearized equations of both the generalized Petviashvili method and the accelerated ITEM has the form:

$$\hat{L} = (c - \nabla^2)^{-1}L. \tag{2.19}$$

With L being the linearized operator of (2.1), the continuous spectrum of $\hat{L}$ is an interval (or, when $F(\mathbf{x}, u)$ is a periodic function of $\mathbf{x}$, a union of intervals), one of the end points of which is $\lambda = -1$ (see e.g. [6] and references therein). This eigenvalue of $\hat{L}$ corresponds to the eigenvalue $\lambda = -\infty$ of L. Then a possible spectrum of $\hat{L}$ is shown in Fig. 1a.

Even though the first terms on the r.h.s.'s of (2.14) and (2.18) have the same form (2.19), the eigenvalues of the corresponding operators $\mathcal{L}$ are different for two reasons. First, the values of c in operators N and K are, in general, different, which makes different the eigenvalues of the corresponding $\hat{L}$'s. Second, the nonlocal terms (involving inner products) in (2.14) and (2.18) modify the eigenvalues of $\hat{L}$ in different ways. We now consider this latter issue in more detail.

In regards to the operator of the linearized Petviashvili method (2.14), we recall a fact [1] that is *important for our discussion* both here and in the next section. Namely, the role of the nonlocal term in that operator is to (nearly) eliminate from $\tilde{u}_{n+1}$ the eigenfunction of $\hat{L} = N^{-1}L$ whose profile is close to that of the solitary wave u, while leaving the other eigenfunctions and their eigenvalues (nearly) unchanged. This is ensured by taking the

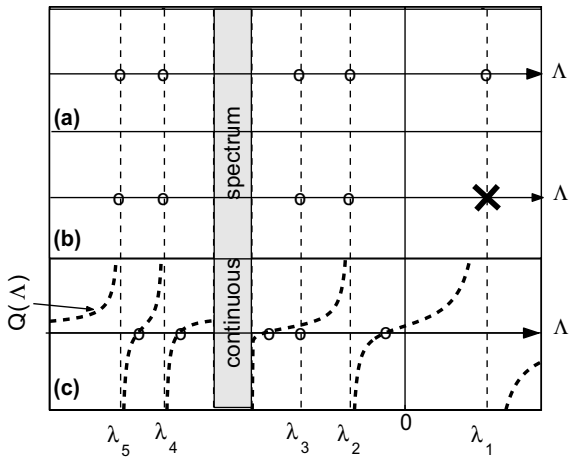

Fig. 1. Schematics of the spectra of the operators found on the r.h.s.'s of (2.19) (a), (2.14) (b), and (2.18) (c). The circles show the location of discrete eigenvalues. The cross on the right of (b) indicates the disappearance of the eigenvalue compared to (a). The thick dashed line in (c) shows a sample function of Eq. (2.21). The left edge of the continuous spectrum is located at $\Lambda = -1$. It is assumed that $\langle u,\ \psi_3\rangle = 0$; see text after (2.21).

constant γ and the operator N to satisfy (2.11) and (2.12), respectively. The adverb "nearly" is used above to account for the fact that relation (2.13) for Eq. (2.1) with a general nonlinear function $F(\mathbf{x}, u)$ holds only approximately. It is exact only for wave equations with power-law nonlinearity [7], for which the original Petviashvili method was proposed [2]. However, the special choice of the constant c in (2.12), as noted after that equation, makes the approximation in (2.13) sufficiently accurate at least near the "core" of the solitary wave.

Continuing with the discussion about the effect of the nonlocal term in (2.14) on the eigenvalues of the corresponding operator $\mathcal{L}$, let us suppose that u is a *fundamental* solution of the nonlinear wave equation. (E.g., in the case of Eq. (2.1), the fundamental solution, unlike nonfundamental ones, has no nodes.[2] For a more general Eq. (1.1) where the operator M is different from ∇^2, fundamental solutions may have nodes (as, e.g., the lump solution of the Kadomtsev–Petviashvili equation [10]); in that case, their distinguishing feature is that they have one "main" hump, while the nonfundamental solutions usually have several "main" humps.) Then the "u-like" eigenfunction of operator $N^{-1}L$ mentioned in the previous paragraph (see also (2.13)) corresponds to the largest eigenvalue, λ_1, of that operator; see Fig. 1a. Since this eigenfunction is eliminated by the nonlocal term at each iteration, then the resulting spectrum of the operator on the r.h.s. of (2.14) is as shown in Fig. 1b. Thus, for this operator, $\Lambda_{\max} \approx \lambda_2$ and $\Lambda_{\min} \approx \lambda_{\min}$; the reason for using "$\approx$" instead of "$=$" is that relation (2.13) holds approximately, as we noted above. Now, if $\lambda_2 < 0$ and the step size $\Delta\tau$ satisfies a condition

$$1 + \lambda_{\min}\Delta\tau > -1, \tag{2.20}$$

then according to (2.5), the generalized Petviashvili method converges to u. As a sidenote, we mention that for equations with power-law nonlinearity, L is known [11] to have only one positive eigenvalue, and hence the Sylvester inertia law (see, e.g., Theorem 7.6.3 in [12]) guarantees that λ_1 is the only positive eigenvalue of $\hat{L} = N^{-1}L$.

Now let us consider the linearized operator $\mathcal{L}$ in (2.18) for the ITEM (2.15)–(2.17). In [6], we showed that the set of discrete eigenvalues of this $\mathcal{L}$ is the union of two sets: (i) the roots of a function

$$Q(\Lambda) = \sum_j \frac{|\langle u, \psi_j\rangle|^2}{\lambda_j - \Lambda} + \int_{\text{continuum}} \frac{|\langle u, \psi(\lambda)\rangle|^2 \, \mathrm{d}\lambda}{\lambda - \Lambda}, \tag{2.21}$$

where ψ_j is the eigenfunction of $\hat{L}$ corresponding to the eigenvalue λ_j, and also (ii) the set of those λ_j for which $\langle u, \psi_j\rangle = 0$. This is shown schematically in Fig. 1c, with ψ_3 there satisfying $\langle u, \psi_3\rangle = 0$. (Note that $Q(\Lambda)$ does not need to be defined for the continuum eigenvalues Λ.) Thus, for the operator $\mathcal{L}$ in (2.18), $\Lambda_{\min} \geqslant \lambda_{\min}$ and $\Lambda_{\max} > \lambda_2$.

The consideration of the two preceding paragraphs shows that even when the acceleration operators N and K in (2.14) and (2.18) are the same (i.e., have the same c), one cannot, in general, make a definite statement on whether the ratio $(\Lambda_{\max}/\Lambda_{\min})$, and hence the convergence rate, is greater for the generalized Petviashvili method or for the accelerated ITEM. Moreover, the fact that the values of c in N and K are generally different, and hence so are the eigenvalues λ_j of the corresponding two $\hat{L}$'s, further obstructs the comparison of the convergence rates of the two methods. The only two statements that can be made here are the following. (i) For Eqs. (2.1) with an arbitrary nonlinearity, if the ITEM converges to a fundamental solution, then we expect that in most cases (see below), so does the generalized Petviashvili method. (ii) For Eqs. (2.1) with power-law nonlinearity $F(\mathbf{x}, u) = u^p$, the Petviashvili method with the optimal choice of $\Delta\tau$ converges to the fundamental solution faster than does the optimally accelerated ITEM (2.15)–(2.17).

To justify statement (i), first recall that for fundamental solitary waves,

$$(\Lambda_{\max})_{\text{Petviashvili}} \approx \lambda_2 < (\Lambda_{\max})_{\text{ITEM}}, \tag{2.22}$$

as long as the value c in the operator (2.19) is taken to be the same for both methods. Next, if the ITEM converges, then according to (2.5), $(\Lambda_{\max})_{\text{ITEM}} < 0$, thereby implying that $\lambda_2 < 0$. However, by the Sylvester inertia law, the sign of λ_2 does not depend on the actual value of c (as long as $c > 0$). Therefore, with a possible exception of those cases where λ_2 is close to zero, the left part of (2.22) yields $(\Lambda_{\max})_{\text{Petviashvili}} < 0$, which means that the generalized Petviashvili method converges. To prove statement (ii), first note that operator L in this case

[2] By nodes in $D > 1$ spatial dimensions, we mean sets of points of dimension less than D where $u(\mathbf{x}) = 0$.

satisfies the conditions of Theorem 4 of Ref. [6], so that $c = \mu$ is the optimal value for K and $(\Lambda_{\min})_{\mathrm{ITEM}} = \lambda_{\min}(= -1)$. Next, in the Petviashvili method for the equation with $F(\mathbf{x}, u) = u^p$, $N = M$ [2,1] and hence $c = \mu$ as well, whence $(\Lambda_{\min})_{\mathrm{Petviashvili}} = \lambda_{\min}$. Thus, in this case,

$$(\Lambda_{\min})_{\mathrm{Petviashvili}} = (\Lambda_{\min})_{\mathrm{ITEM}}. \tag{2.23}$$

Combining Eq. (2.23) and inequality (2.22), where now the sign "≈" must be replaced with "=", one concludes that $(\Lambda_{\max}/\Lambda_{\min})$ should be greater for the Petviashvili method; hence statement (ii) follows.

A simple example illustrating statement (ii) is the stationary nonlinear Schrödinger equation in one dimension:

$$u_{xx} + u^3 = u, \quad |u| \to 0 \quad \text{as } |x| \to \infty, \tag{2.24}$$

for which the ITEM with the parameters $c = \mu(= 1)$ and $\Delta\tau = 1.5$, corresponding to the optimal acceleration, converges to the accuracy of 10^{-10} in 33 iterations. The Petviashvili method (2.10) with $\Delta\tau = 1.5$, $\alpha = 2$ (as in the original Petviashvili method; see [1]), and γ given by (2.11), converges to the same accuracy in 19 iterations. Here both methods start with the initial condition $u_0 = \mathrm{e}^{-x^2}$. In our numerical experiments of finding the fundamental solutions of non-power-law equations (not covered by the above statement (ii)), we also observed that the generalized Petviashvili method is faster than the optimally accelerated ITEM (2.15)–(2.17); see, e.g., Example 3.1 in [1]. (The ITEM with amplitude normalization [6] can still be faster than the generalized Petviashvili method.)

However, in a situation where both methods converge to a *nonfundamental* solitary wave, the optimally accelerated ITEM can be faster than the generalized Petviashvili method. As an example, let us revisit the equation with a double-well potential:

$$u_{xx} + V(x)u - u^3 = \mu u, \quad V(x) = 6(\mathrm{sech}^2(x-1) + \mathrm{sech}^2(x+1)), \tag{2.25}$$

considered in Example 3.2 of [1]. We will focus on its anti-symmetric solution (see Fig. 2a) with the propagation constant $\mu = 1.43$ and the corresponding power $P \equiv \int_{-\infty}^{\infty} u^2\,\mathrm{d}x = 10$. This solution is nonfundamental since it has a node; the fundamental solution in this case is a two-humped pulse with its maxima located near the maxima of the potential. The solid and dashed lines in Fig. 2b show the evolutions of the error norm, defined as

$$E_n = \left(\frac{\langle u_n - u_{n-1}, u_n - u_{n-1}\rangle}{\langle u_n, u_n\rangle}\right)^{1/2}, \tag{2.26}$$

for the generalized Petviashvili method and the optimally accelerated ITEM, respectively. In both cases, the parameter $\Delta\tau$ was emprically optimized (see (2.6)) to yield the maximum convergence rates; the respective values are $\Delta\tau_{*,\mathrm{Petviashvili}} = 1.6$ and $\Delta\tau_{*,\mathrm{ITEM}} = 0.7$. Also, in the case of the generalized Petviashvili method, the value

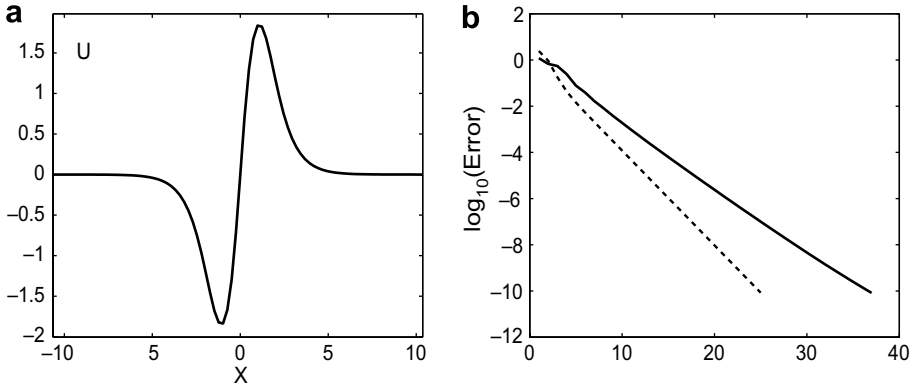

Fig. 2. (a) The anti-symmetric solution of Eq. (2.25) with $\mu = 1.43$ ($P = 10$). (b) Error evolutions, starting with an anti-symmetric initial condition, for the generalized Petviashvili method (solid line) the optimally accelerated ITEM (dashed line).

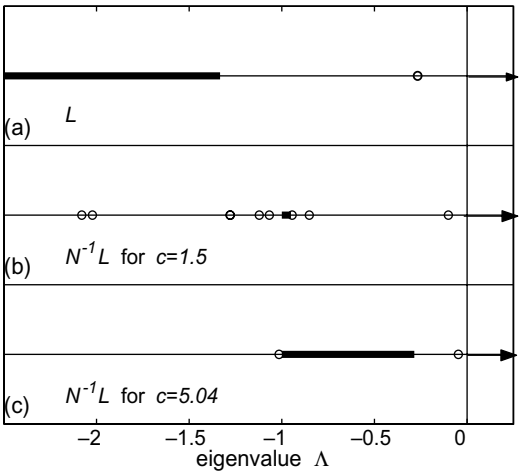

Fig. 3. The actual spectra of the linearized operator L of Eq. (2.25) (a) and of the operators $N^{-1}L$, where N is given by (2.12), with $c = 1.5$ (b) and $c = 5.04$ (c). The operator in (b) has a very short interval of continuous spectrum between -1 and -0.95.

$c = 5.04$ was algorithmically computed [1], while for the ITEM, $c = 1.5$ was empirically found to yield the optimal convergence rate. As the initial condition for both these methods, we took $u_0 = 2xe^{-x^2}$. As seen from Fig. 2b, the optimally accelerated ITEM is about one and a half times faster than the generalized Petviashvili method. The reason behind this can be understood by looking at the spectra of the corresponding operators $\hat{L}$ in (2.19) with the above values $c = 1.5$ for the accelerated ITEM (Fig. 3b) and $c = 5.04$ for the generalized Petviashvili method (Fig. 3c). Namely, when one starts with an anti-symmetric initial condition (as we did above), the symmetric eigenmodes corresponding to $\lambda_{2k+1},\ k = 0, 1, \ldots$ do not contribute to the error $\tilde{u}_n$. Then from (2.7) and Figs. 3b and c,

$$R_{\mathrm{ITEM}} < \frac{1 - (\lambda_{2,c=1.5}/\lambda_{\min,c=1.5})}{1 - (\lambda_{2,c=1.5}/\lambda_{\min,c=1.5})} = \frac{1 - 0.41}{1 + 0.41} = 0.42,$$

$$R_{\mathrm{Petviashvili}} \approx \frac{1 - (\lambda_{\max\ \mathrm{continuum},c=5}/\lambda_{\min,c=5})}{1 + (\lambda_{\max\ \mathrm{continuum},c=5}/\lambda_{\min,c=5})} = \frac{1 - 0.28}{1 + 0.28} = 0.56,$$

and hence the corresponding numbers of iterations to reach the accuracy of 10^{-10} can be estimated as:

$$n_{\max,\mathrm{ITEM}} \approx \frac{-10 \ln 10}{\ln R_{\mathrm{ITEM}}} = 26, \quad n_{\max,\mathrm{Petviashvili}} \approx \frac{-10 \ln 10}{\ln R_{\mathrm{Petviashvili}}} = 40.$$

These estimates are in very good agreement with the numbers of iterations (25 and 37, respectively) reported in Fig. 2b. Note also that the empirically found optimal values of $\Delta\tau_*$ reported above agree with Eq. (2.6) and the spectra shown in Figs. 3b and c.

3. Mode elimination technique for improving convergence of iteration methods

Here we develop the ideas of Ref. [1] and extend the generalized Petviashvili method so that it could be employed for two additional purposes: (i) obtaining certain nonfundamental solutions of stationary nonlinear wave equations; and (ii) accelerating convergence of iterations methods. We emphasize that the technique we propose can be applied to *any* iteration method and to single and coupled equations as well. For simplicity of the presentation, below we illustrate it for single equations of the form (2.1).

We begin with the observation that in most cases (with Eq. (2.25) being a notable exception), the generalized Petviashvili method would not converge to a nonfundamental solution of a given wave equation. The reason for that can be understood from the following simple example. Consider an equation

$$u_{xx} + (6\mathrm{sech}^2 x + u^2)u = \mu u. \tag{3.1}$$

When the amplitude of u is small, (3.1) has two solutions: the fundamental, $\{u^{(1)} \approx \epsilon \mathrm{sech}^2 x,\ \mu^{(1)} \approx 4\}$, and the nonfundamental, $\{u^{(2)} \approx \epsilon \mathrm{sech} x \tanh x,\ \mu^{(2)} \approx 1\}$, where $\epsilon \ll 1$. Then the operator obtained by the linearization of Eq. (3.1) on the background of the nonfundamental solution,

$$L \approx \partial_x^2 + 6\mathrm{sech}^2 x - \mu^{(2)}, \tag{3.2}$$

has two largest eigenvalues: $\lambda_1 \approx \mu^{(1)} - \mu^{(2)} \approx 3 > 0$ and $\lambda_2 \approx \mu^{(2)} - \mu^{(2)} = 0$, with the corresponding eigenfunctions being approximately $u^{(1)}$ and $u^{(2)}$. As we noted in Section 2, the nonlocal term in the linearized iteration equation (2.14) nearly eliminates the eigenfunction of operator $\hat{L} = N^{-1}L$ which is "similar" to the background solution $u^{(2)}$. However, the eigenfunction of $\hat{L}$ corresponding to the eigenvalue $\lambda_1 > 0$ of $\hat{L}$ is not eliminated, and hence, according to the discussion found before Eq. (2.20), the generalized Petviashvili method will not converge to solution $u^{(2)}$.

The above example suggests a simple way in which the generalized Petviashvili method (2.10) can be modified so that it would converge to a nonfundamental solution u (given, of course, an initial condition close to u). In the general form, this modified method is

$$u_{n+1} - u_n = \left[N^{-1}(L_0 u)_n - \gamma \frac{\langle u_n, (L_0 u)_n \rangle}{\langle u_n, N u_n \rangle} u_n - \sum_{j=1}^{J_{\mathrm{unst}}} \gamma_{\mathrm{unst}}^{(j)} \frac{\langle \phi_{\mathrm{unst}}^{(j)}, (L_0 u)_n \rangle}{\langle \phi_{\mathrm{unst}}^{(j)}, N \phi_{\mathrm{unst}}^{(j)} \rangle} \phi_{\mathrm{unst}}^{(j)} \right] \Delta\tau, \tag{3.3}$$

where γ and N are defined as in (2.11) and (2.12), $\phi_{\mathrm{unst}}^{(j)}$ are the functions that approximate the eigenmodes of operator $(N^{-1}L)$ with positive eigenvalues (excluding the background solution u), J_{unst} is the number of such eigenmodes, and

$$\gamma_{\mathrm{unst}}^{(j)} = 1 + \frac{1}{\alpha_{\mathrm{unst}}^{(j)} \Delta\tau}, \quad \alpha_{\mathrm{unst}}^{(j)} = \frac{\langle \phi_{\mathrm{unst}}^{(j)}, L\phi_{\mathrm{unst}}^{(j)} \rangle}{\langle \phi_{\mathrm{unst}}^{(j)}, N\phi_{\mathrm{unst}}^{(j)} \rangle}. \tag{3.4}$$

Here $\alpha_{\mathrm{unst}}^{(j)}$, defined analogously to (2.13):

$$L\phi_{\mathrm{unst}}^{(j)} \approx \alpha_{\mathrm{unst}}^{(j)} N \phi_{\mathrm{unst}}^{(j)}, \tag{3.5}$$

is computed according to Eq. (3.12) of [1]. In the context of the example in the previous paragraph, $J_{\mathrm{unst}} = 1$ and $\phi_{\mathrm{unst}}^{(1)} = u^{(1)}$.

Following the lines of the analysis of Section 2 in Ref. [1], it is straightforward to show that in method (3.3), (3.4), the components of the error $\tilde{u}_n$ "aligned along" the modes $\phi_{\mathrm{unst}}^{(j)}$, $j = 1, \ldots, J_{\mathrm{unst}}$, are nearly eliminated at every iteration; this is guaranteed by the form of the coefficients $\gamma_{\mathrm{unst}}^{(j)}$. Therefore, in what follows, we refer to method (3.3) as the mode elimination method. In Section 4, we will present the results of applying this method to a two-dimensional equation of the form (3.1) to obtain its nonfundamental solutions.

Remark. It is clear that the success of the mode elimination method hinges upon the knowledge of the "unstable" eigenmodes $\phi_{\mathrm{unst}}^{(j)}$. However, in many cases, an approximate knowledge of $\phi_{\mathrm{unst}}^{(j)}$ may suffice.

We now show how the mode elimination technique can be used to accelerate convergence of iteration methods. The reason that a given method converges slowly is, according to (2.7), that the ratio $\Lambda_{\max}/\Lambda_{\min}$ is small. Since for an appropriately chosen operator N, $|\Lambda_{\min}| = \mathrm{O}(1)$ (see, e.g., Figs. 3b and c), then for a slowly convergent method, the eigenvalue $|\Lambda_{\max}|$ must be small. Then if one can eliminate the corresponding eigenmode, similarly to how it is done in (3.3), one essentially replaces $(\Lambda_{\max})_{\mathrm{old}}$ with $(\Lambda_{\max})_{\mathrm{new}} < (\Lambda_{\max})_{\mathrm{old}} (< 0)$. Then the ratio $\Lambda_{\max}/\Lambda_{\min}$ increases and so does the convergence rate of the iteration method. The practical issue here is how to find the mode, ϕ_{slow}, which slows down the convergence. Fortunately, this is rather easy to do using the following observation. For $\Delta\tau < \Delta\tau_*$, where $\Delta\tau_*$ is defined in (2.6), the factor $(1 + \Lambda_{\mathrm{slow}}\Delta\tau) \equiv (1 + \Lambda_{\max}\Delta\tau)$, which governs the decay of ϕ_{slow}, is the largest among such factors for all the eigenmodes of $(N^{-1}L)$. Then after some iterations, the content of the error $\tilde{u}_n \equiv u_n - u$ becomes dominated by the eigenmode ϕ_{slow}, and hence

$$\phi_{\mathrm{slow}} \propto (u_n - u_{n-1}). \tag{3.6}$$

The elimination of the function $(u_n - u_{n-1})$ is carried out in exactly the same way as in (3.3), yielding the method:

$$u_{n+1} - u_n = \left[N^{-1}(L_0 u)_n - \gamma \frac{\langle u_n, (L_0 u)_n \rangle}{\langle u_n, N u_n \rangle} u_n - \gamma_{\text{slow},n} \frac{\langle \phi_{\text{slow},n}, (L_0 u)_n \rangle}{\langle \phi_{\text{slow},n}, N \phi_{\text{slow},n} \rangle} \phi_{\text{slow},n} \right] \Delta\tau, \tag{3.7}$$

where

$$\phi_{\text{slow},n} = u_n - u_{n-1}, \quad \gamma_{\text{slow},n} = 1 + \frac{s}{\alpha_{\text{slow},n} \Delta\tau}, \quad \alpha_{\text{slow},n} = \frac{\langle \phi_{\text{slow},n}, L \phi_{\text{slow},n} \rangle}{\langle \phi_{\text{slow},n}, N \phi_{\text{slow},n} \rangle}. \tag{3.8}$$

Note the coefficient s in (3.8), which we will comment on in the next paragraph. We will also provide examples that demonstrate the efficiency of the accelerated Petviashvili method (3.7), (3.8) and its extensions to other iteration methods, in the next section.

Similarly to the analysis of Ref. [1], one can show that the role of coefficient s in (3.8) is to control how much of the mode $\phi_{\text{slow},n}$ is subtracted at each iteration. We found empirically that in most cases, it is beneficial for the convergence rate to subtract not the entire $\phi_{\text{slow},n}$-component from u_n but only part of it, usually somewhere between 40% and 80% (i.e., use $s \sim 0.4$–0.8). (However, even using the value $s = 1$ leads to a significant increase in convergence rate compared to the corresponding non-accelerated method when the latter is slow.) The justification of using $0 < s < 1$ (or, alternatively, $1 < s < 2$) rather than $s = 1$ is based on the same considerations, found before Eq. (3.6), which led us to propose the accelerated method (3.7). Namely, to uphold those considerations, $\phi_{\text{slow},n}$ is to remain the most slowly decaying eigenmode of $(N^{-1}L)$ at every iteration. In the case where the entire amount of it is subtracted at the $(n+1)$st iteration, it is not obvious (and probably not true) that the error $\tilde{u}_{n+2}$ at the next iteration would consist mainly of the mode $\phi_{\text{slow},n+1} \equiv u_{n+1} - u_n$, which will be subtracted at the $(n+2)$nd iteration. However, if only $s \cdot 100\%$ of mode $\phi_{\text{slow},n}$ is subtracted, this mode can still remain the most slowly decaying as long as

$$|(1-s) \cdot (1 + \Lambda_{\text{slow}} \Delta\tau)| > |1 + \Lambda_{\text{next}} \Delta\tau|, \tag{3.9}$$

where Λ_{slow} is the eigenvalue corresponding to $\phi_{\text{slow},n}$, and Λ_{next} is the eigenvalue corresponding to the next most slowly decaying mode. Yet, for s not too small, the l.h.s. of (3.9) is still considerably less than $|1 + \Lambda_{\text{slow}} \Delta\tau|$, and hence the convergence rate of the original iteration method is increased.

To conclude this section, we compare our mode elimination technique for convergence acceleration with the Steffensen's method (see e.g. [13]), which is based on applying the Aitken's acceleration algorithm every given number of iterations. The idea of the Steffensen's method is the following. Suppose one has three consecutive iterative solutions u_n, u_{n+1}, u_{n+2} about which one knows that they satisfy

$$\frac{\tilde{u}_{n+2}(\mathbf{x})}{\tilde{u}_{n+1}(\mathbf{x})} \approx \frac{\tilde{u}_{n+1}(\mathbf{x})}{\tilde{u}_n(\mathbf{x})} \quad \text{for all } \mathbf{x}, \tag{3.10}$$

where $\tilde{u}_n$ is the error defined in (2.2). Using these solutions, one applies the Aitken's algorithm:

$$u_{n+3} \equiv u_n^A = u_n - \frac{(u_{n+1} - u_n)^2}{u_{n+2} - 2u_{n+1} + u_n}, \tag{3.11}$$

and then proceeds to computing the next few iterations $u_{n+4}, \ldots, u_{n+n_{\text{accel}}+2}$ with the original iteration method, where $n_{\text{accel}} \geqslant 3$. Then one uses $u_{n+n_{\text{accel}}}$, $u_{n+n_{\text{accel}}+1}$, $u_{n+n_{\text{accel}}+2}$ to compute $u^A_{n+n_{\text{accel}}}$ by (3.11) with $n \to n + n_{\text{accel}}$, and so on. In [14], this method was successfully used to accelerate the convergence of the original Petviashvili method for the nonlinear Schrödinger equation in three spatial dimensions.

Aitken's algorithm (3.11) systematically reduces the error $(u_n^A - u)$ only when (3.10) holds sufficiently well, which occurs under the same condition (3.6) that must hold in order for the mode elimination method to work. However, the *sense* in which (3.6) is to hold is drastically different for these two acceleration techniques. For the mode elimination, it suffices if (3.6) holds approximately *near the "core"* of the solitary wave, since ϕ_{slow} enters Eqs. (3.7) and (3.8) via the inner products with functions that are essentially nonzero only in that spatial region. On the contrary, for the Steffensen's method, (3.10) has to hold *pointwise* and, in particular, far away from the "core" of $u(\mathbf{x})$. In the latter spatial region, the denominator of (3.11) is nearly zero, and hence even a small ripple in u_n, u_{n+1}, or u_{n+2} can result in a large distortion of u_n^A. This was indeed observed in our numerical experiments, except in the cases where $|\Lambda_{\max}| \ll |\Lambda_{\text{next}}|$, where $|\Lambda_{\text{next}}|$ is defined after (3.9). Thus, we expect our mode elimination technique and the Steffensen's method to be competitive in those latter cases, but

expect the mode elimination technique to have superior performance over that of the Steffensen's method when there are more than one eigenmodes with $\Lambda \approx \Lambda_{\max}$. This expectation is borne out by Examples 4.2 and 4.3.

4. Examples of the mode elimination technique

Below we illustrate the application of the mode elimination technique to obtaining nonfundamental solitary waves and to accelerating convergence of iteration methods for Eq. (2.1). In Ref. [8], we already showed by extensive simulations that this technique can greatly accelerate convergence of a class of universally-convergent iteration methods for both single and coupled equations. (Method (4.4) presented below is a particular member of that class.) Therefore, here we will focus on clarifying the role of parameter s in Eq. (3.8) for optimizing the convergence rate and also on demonstrating the applicability of the mode elimination technique to *various* classes of iteration methods.

Example 4.1. Here we will demonstrate that method (3.3), (3.4) can be used to obtain nonfundamental solitary waves when approximate information about the unstable eigenmodes of $(N^{-1}L)$ is available. We will also compare the performance of this method with that of a universally-convergent method proposed in [8]. The following equation

$$\nabla^2 u + V_0(\mathrm{sech}x\,\mathrm{sech}y)^2 u + u^3 = \mu u, \quad V_0 = 20 \tag{4.1}$$

is a two-dimensional counterpart of Eq. (3.1). Since the potential well in (4.1) is sufficiently deep ($V_0 \gg 1$), this equation admits several nonfundamental solutions. Below we report the details of finding the first of them which corresponds to $\mu = 8$ and is shown in Fig. 4. For this solution, we expect the generalized Petviashvili method to have one unstable eigenmode (in addition to the mode approximated by u that may possibly also be unstable), and approximate this eigenmode by

$$\phi_{\mathrm{unst}} = \mathrm{e}^{-\frac{1}{2}(r/W)^2}, \quad r^2 = x^2 + y^2. \tag{4.2}$$

The width W in (4.2) is found iteratively from the formula

$$W_n^2 = \frac{2}{3}\frac{\langle u_n, x^2 u_n\rangle}{\langle u_n, u_n\rangle}, \tag{4.3}$$

in deriving which we assumed that $u \propto x\phi_{\mathrm{unst}}$. Starting with the initial condition $u_0 = 2x\mathrm{e}^{-(x^2+y^2)}$, method (3.3), (3.4) with a nearly optimal $\Delta\tau = 0.7$ took about 50 iterations to reach the accuracy of 10^{-10}. Thus, the general-

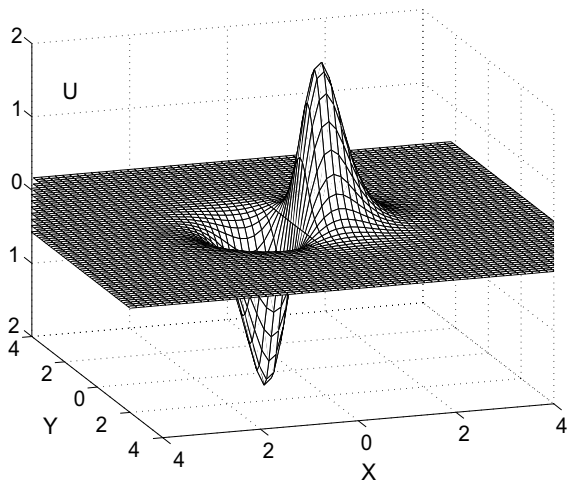

Fig. 4. The first nonfundamental solution of Eq. (4.1) with $\mu = 8$.

ized Petviashvili method with mode elimination (3.3), (3.4) converges to this nonfundamental solution, while the generalized Petviashvili method (2.10) without the mode elimination diverges.

We also obtained the same solution by a method based on the "squared" operator $(N^{-1}L)$:

$$u_{n+1} - u_n = -\left[(N^{-1}LN^{-1}L_0u)_n - \Gamma_n \frac{\langle u_n, (LN^{-1}L_0u)_n\rangle}{\langle u_n, Nu_n\rangle} u_n\right]\Delta\tau. \tag{4.4}$$

(The name "squared" comes from the fact that $(N^{-1}L)^2$ appears in the linearized version of (4.4).) In [8], we showed that this method belongs to a family of universally-convergent methods (i.e., methods which can converge to *any* nonfundamental solution of a given equation provided that the initial condition is sufficiently close to that solution) for either of the following choices of Γ_n: $\Gamma_n = 0$ or

$$\Gamma_n = 1 - \frac{1}{(\langle u_n, (LN^{-1}Lu)_n\rangle / \langle u_n, Nu_n\rangle)\Delta\tau}. \tag{4.5}$$

Note that this Γ_n is defined similarly to γ_n in the generalized Petviashvili method (see Eqs. (2.11) and (3.4)). Since we are looking for a nonfundamental solution of (4.1), then using the value for Γ_n given by (4.5) as opposed to $\Gamma_n = 0$ will not eliminate the mode with the maximum eigenvalue (see the discussion after Eq. (3.2)), and hence will not speed up the convergence of the iterations. Therefore, in the remainder of this example we report the results for method (4.4) with $\Gamma_n = 0$. Starting with the same initial condition as above, this method with the operator N computed as in [1] and with a nearly optimal $\Delta\tau = 0.5$ took about 190 iterations to converge to the accuracy of 10^{-10}. Thus, the mode elimination method (3.3), (3.4) is several times faster than the squared-operator method (4.4) for finding the first nonfundamental solution of (4.1). (We also observed that method (3.3), (3.4) is less sensitive to the choice of initial conditions than method (4.4).) However, when we additionally included the step of eliminating the slow mode, as in Eqs. (3.6)–(3.8), into both methods, the difference in their convergence rates was significantly reduced. Namely, the convergence of method (3.3), (3.4), which has already been quite rapid, was not improved by this additional step (and the number of iterations remained around 50), while the squared-operator method now took about 70 iterations to converge.

We also applied both methods to finding the second nonfundamental solution of (4.1), which has the shape similar to $A(1 - Br^2)e^{-(r/C)^2}$ with $r^2 = x^2 + y^2$ and $A, B, C = \text{const}$ (see Fig. 7). For this solution we found, through experimentation, that one needs to include five unstable modes into (3.3). For the respective optimal $\Delta\tau$'s, the generalized Petviashvili method with mode elimination (3.3), (3.4) was found to be about 50% faster than the squared-operator method (4.4). However, this advantage in the convergence rate is offset by the increased complexity arising from the need to guess the number and profiles of unstable modes and then to estimate their parameters (namely, the widths). Therefore, we conclude that the mode elimination method may be more efficient than the squared-operator method for finding the *lowest-order* nonfundamental solitary waves, as long as some reasonable guess about the unstable modes can be made. However, for finding second- and higher-order nonfundamental solutions, method (4.4) appears to be easier to implement and hence more practical.

Example 4.2. In this and the next two examples, we demonstrate the efficiency of the convergence acceleration technique based on the mode elimination, as in (3.7) and (3.8), for three different iteration methods. In this example, we apply this technique to the generalized Petviashvili method.

We look for the fundamental solitary wave of an equation arising in the theory of nonlinear photonic lattices:

$$\nabla^2 u + V_0(\cos^2 x + \cos^2 y)u + u^3 = \mu u \tag{4.6}$$

for three choices of the potential amplitude and the propagation constant:

$$(a): V_0 = 4, \mu = 4.95; \quad (b): V_0 = 4, \mu = 6.5; \quad (c): V_0 = 0, \mu = 1.$$

In case (a), the propagation constant is close to the edge of the continuous spectrum band, and the solitary wave occipies many "sites" of the potential, while in case (b), the propagation constant is sufficiently far away from the band edge, and the solitary wave is well localized. (The profiles of the corresponding solutions are similar to those of the top and bottom solutions shown in Fig. 3 of [1].) Case (c) is that of the nonlinear

Schrödinger equation in two spatial dimensions. In all cases, we apply three methods: the generalized Petviashvili method (2.10) without any acceleration, the same method with the Aitken's acceleration (3.11) performed after every third iteration ($n_{\text{accel}} = 3$), and the mode elimination method (3.7), (3.8) with various values of s (see the paragraph including Eq. (3.9)). The initial condition in all cases is $u_0 = 1.5e^{-(x^2+y^2)}$, and the step size $\Delta\tau = 1$.

In case (a), the generalized Petviashvili method (2.10) takes about 950 iterations to converge to the accuracy of 10^{-10}. When the mode elimination technique is applied, starting at the moment when the error becomes less or equal to some small value (we chose 10^{-2}), the convergence occurs in about 180 iterations, i.e. more than five times as fast. The evolution of the error is shown in Fig. 5 by the thick solid line for the choice $s = 1$; for smaller values of s up to 0.4 which we tried, the error evolution is similar (and the convergence is slightly faster). The characteristic feature of this error evolution is that it is nonmonotonic and rather irregular. This irregularity is somewhat abated for $s < 1$, in agreement with our discussion in Section 3. Now, when we attempted to apply the Aitken's acceleration to the generalized Petviashvili method, we observed quick divergence of the so "accelerated" method. We actually tried various values of n_{accel} and $\Delta\tau$ but were unable to make the iterations converge. The reason for this is explained at the end of Section 3. In fact, by monitoring the error $\tilde{u}_n$ at every iteration, we observed that it contains many nonlocalized modes, so that the condition (3.10) of applicability of the Aitken's acceleration is clearly violated in this case.

The corresponding results for case (b) are also shown in Fig. 5. There, the mode elimination technique accelerates the convergence of the generalized Petviashvili method by about a factor of four. The error evolution is much smoother than in case (a). This appears to be correlated with the fact, which follows from our monitoring of the error, that the latter is dominated by a single eigenmode. Consequently, condition (3.10) is now satisfied, and the Steffensen's method (i.e., the generalized Petviashvili method with Aitken's acceleration) also converges; see the dotted line in Fig. 5. Let us note that the irregular behavior of the error of the Steffensen's method at low values of the error leads to a rather high sensitivity of the total number of iterations to the initial condition. For example, we verified that if the acceleration is started when the error reaches 10^{-3} instead of 10^{-2}, then the Steffensen's method converges to the accuracy of 10^{-10} in about 30 iterations.

The error evolutions for case (c) are shown in Fig. 6. The convergence acceleration in this case (as, actually, also in case (b)) is not of practical importance because the convergence of the non-accelerated generalized Petviashvili method (2.10) is quite fast (see the thin solid line in Fig. 6). Therefore, below we discuss the results for this method accelerated by the mode elimination technique for the sole purpose of highlighting this

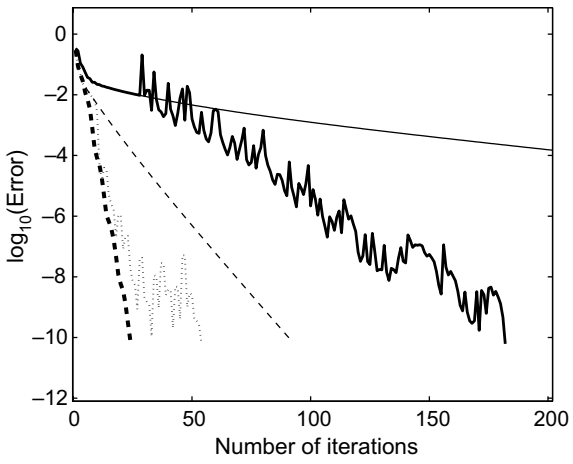

Fig. 5. The evolution of the error in cases (a) and (b) of Example 4.2. Thin solid: non-accelerated method (2.10) for case (a); thick solid: method (3.7), (3.8) with $s = 1$ for case (a); thin dashed: non-accelerated method (2.10) for case (b); thick dashed: method (3.7), (3.8) with $s = 1$ for case (b); dotted line: Steffensen's method for case (b).

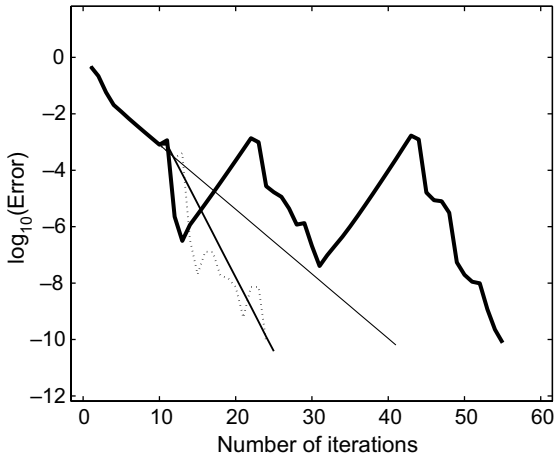

Fig. 6. The evolution of the error in case (c) of Example 4.2. Thin solid: non-accelerated method (2.10); thick solid: method (3.7), (3.8) with $s = 1$; medium solid: method (3.7), (3.8) with $s = 0.7$; dotted line: Steffensen's method.

technique's dependence on the parameter s. The error evolution of method (3.7), (3.8) with $s = 1$, where the acceleration is started when the error becomes less or equal to 10^{-3}, is very irregular (see the thick solid line in Fig. 6), and as a result, the accelerated method takes more iterations to converge than the non-accelerated one. Moreover, the evolution of the error also strongly depends on the initial condition and on when the acceleration is started. For example, when we began the acceleration at the moment of the error reaching 10^{-2} or 10^{-4}, rather than 10^{-3}, the convergence occurred in about 190 or 100 iterations, respectively. In both cases, the error evolution curves were irregular, with several "ups and downs". However, when we used values $0.4 < s < 0.8$ instead of $s = 1$, the behavior of the accelerated iterations greatly improved. The optimal case of $s = 0.7$ is shown in Fig. 6 by the medium solid line. Both the sensitivity to the "starting moment" of the acceleration and the irregularity of the error evolution are suppressed for $s < 1$, in agreement with the discussion in Section 3. We also applied the Steffensen's method to this case and found it to converge in about the same number of iterations as the mode elimination method with the optimal s; see the dotted line in Fig. 6.

Example 4.3. In this and the following examples, we show that the mode elimination technique can be used to accelerate convergence of other iterative methods. In this example, we apply this technique to the squared-operator method (4.4), which can converge [8] to any given nonfundamental solitary wave of the underlying stationary wave equation. It should be noted that in [8], the efficiency of the so accelerated squared-operator

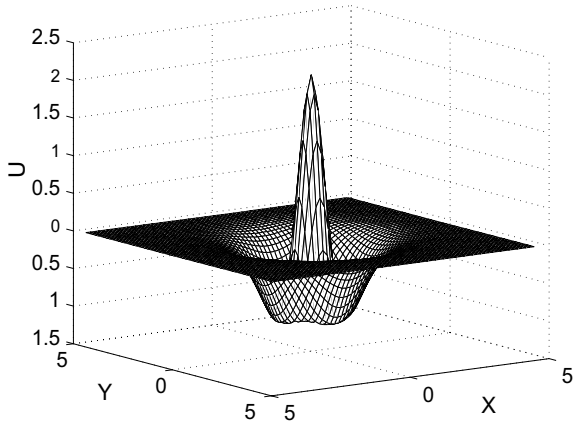

Fig. 7. The second nonfundamental solution of Eq. (4.1) with $\mu = 3$.

methods (referred to there as modified squared-operator methods) was amply demonstrated for a number of single and coupled stationary wave equations, both Hamiltonian and dissipative. In all simulations reported in [8], the value of the parameter s in (3.8) was taken to equal 1. Therefore, below we will focus on the dependence of the error evolution on the parameter s.

We apply the squared-operator methods with and without mode elimination to find the second nonfundamental solution of Eq. (4.1). This solution for $\mu = 3$ is shown in Fig. 7. In all cases considered below, we used the initial condition $u_0 = (1 - 2r^2)e^{-r^2}$, $r^2 = x^2 + y^2$ and the step size $\Delta\tau = 0.3$ (nearly optimal). As the method without mode elimination, we used (4.4). The method with mode elimination is then a straightforward modification of methods (3.7), (3.8) and (4.4):

$$u_{n+1} - u_n = -\left[\left(N^{-1}LN^{-1}L_0u\right)_n - \Gamma_n \frac{\langle u_n, (LN^{-1}L_0u)_n\rangle}{\langle u_n, Nu_n\rangle} u_n - \Gamma_{\text{slow},n} \frac{\langle \phi_{\text{slow},n}, (LN^{-1}L_0u)_n\rangle}{\langle \phi_{\text{slow},n}, N\phi_{\text{slow},n}\rangle} \phi_{\text{slow},n}\right]\Delta\tau, \tag{4.7}$$

where, similarly to (3.8):

$$\phi_{\text{slow},n} = u_n - u_{n-1}, \quad \Gamma_{\text{slow},n} = 1 - \frac{s}{\mathcal{A}_{\text{slow},n}\Delta\tau}, \quad \mathcal{A}_{\text{slow},n} = \frac{\langle \phi_{\text{slow},n}, LN^{-1}L\phi_{\text{slow},n}\rangle}{\langle \phi_{\text{slow},n}, N\phi_{\text{slow},n}\rangle}. \tag{4.8}$$

In both cases, with and without mode elimination, we found empirically that the methods with Γ_n given by (4.5) require the initial condition to be closer to the exact solution than do the corresponding methods with $\Gamma_n = 0$. On the other hand, the former methods were significantly faster than those with $\Gamma_n = 0$. Therefore, we initially used methods (4.4) or (4.7) with $\Gamma_n = 0$, and when the error reached a small value (we chose 5×10^{-3}), switched Γ_n to the expression (4.5). The corresponding error evolutions for the accelerated method (4.7) with $s = 1$ and $s = 0.7$ (optimal) are shown in Fig. 8 by the thick and medium lines, while for the non-accelerated method (4.4) without mode elimination, the error evolution is shown by the thin line. Note that the behavior of the accelerated method with $s < 1$ compared to that behavior with $s = 1$ follows the same trends as observed in Example 4.2. Namely, the error evolution for the schemes with $s < 1$ is smoother and much less sensitive to the moment when the acceleration starts. Overall, the mode elimination is found to accelerate the convergence by a factor between three and four, depending on the choice of the parameter s. Finally, we note that the Steffensen's method in this case does not converge.

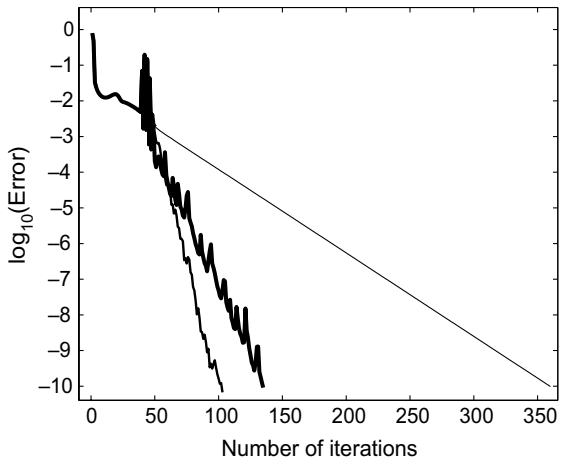

Fig. 8. The evolution of the error in Example 4.3. In all cases, the application of (4.5) and/or acceleration is begun when the error norm reaches 5×10^{-3}. Thin line: method (4.4) (no mode elimination); thick line: method (4.7) with $s = 1$; medium line: method (4.7) with $s = 0.7$.

Example 4.4. In this last example, we show that the convergence acceleration technique based on mode elimination can also be applied to the ITEM. Here we chose to present the results for the version of this method (2.15)–(2.17) with power normalization, but the technique can be used as well for the ITEM with amplitude normalization [6].

For the stationary wave equation (2.1) written in an equivalent form:

$$L_0 u \equiv L_{00} u - \mu u = 0, \tag{4.9}$$

the ITEM (2.15)–(2.17) with mode elimination can be written as follows:

$$u_{n+1} = \left[\frac{P}{\langle \hat{u}_{n+1}, \hat{u}_{n+1}\rangle}\right]^{\frac{1}{2}} \hat{u}_{n+1}, \tag{4.10}$$

$$\hat{u}_{n+1} - u_n = \left[K^{-1}(L_0 u)_n - \gamma_{\mathrm{slow},n} \frac{\langle \phi_{\mathrm{slow},n}, (L_0 u)_n\rangle}{\langle \phi_{\mathrm{slow},n}, K\phi_{\mathrm{slow},n}\rangle} \phi_{\mathrm{slow},n}\right]\Delta\tau. \tag{4.11}$$

Here K is a positive definite self-adjoint operator with constant coefficients (as, e.g., in (2.12)).

$$(L_0 u)_n = L_{00} u_n - \mu_n u_n, \quad \mu_n = \frac{\langle L_{00} u_n, K^{-1} u_n\rangle}{\langle u_n, K^{-1} u_n\rangle}, \tag{4.12}$$

and

$$\phi_{\mathrm{slow},n} = u_n - u_{n-1}, \quad \gamma_{\mathrm{slow},n} = 1 + \frac{s}{\alpha_{\mathrm{slow},n}\Delta\tau}, \quad \alpha_{\mathrm{slow},n} = \frac{\langle \phi_{\mathrm{slow},n}, L\phi_{\mathrm{slow},n}\rangle}{\langle \phi_{\mathrm{slow},n}, K\phi_{\mathrm{slow},n}\rangle}. \tag{4.13}$$

We apply the methods without and with mode elimination – (2.15)–(2.17), (4.10)–(4.13), respectively – to Eq. (4.6) with $V_0 = 4$ and $P = 1.94$, whose solution looks similar to the top solution in Fig. 3 of [1]. The corresponding propagation constant $\mu = 5.01$ is close to the bandgap edge, and the ITEM without mode elimination converges slowly; see the thin line in Fig. 9. In all simulations, we took $\Delta\tau = 1$ and the operator K of the form (2.12) with $c = 1$, which yielded the (nearly) optimal convergence rate of the ITEM (2.15)–(2.17). The error evolutions for the ITEM (4.10)–(4.13) with mode elimination are shown in Fig. 9 by the thick and medium lines. As in Examples 4.2 and 4.3, the scheme with mode elimination provides a severalfold improvement to the convergence rate of the ITEM. Also as in those examples, the error evolution with $s < 1$ is more regular than that with $s = 1$.

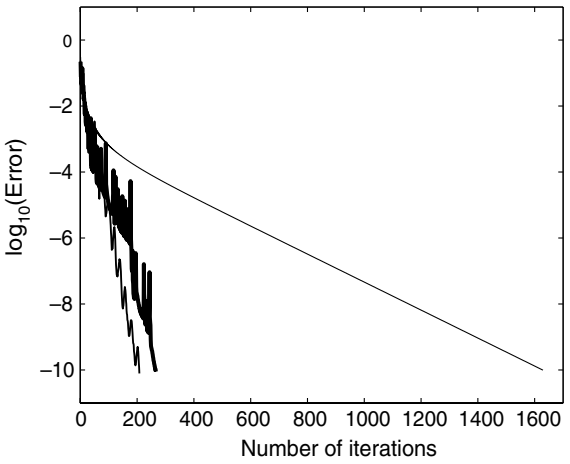

Fig. 9. The evolution of the error in Example 4.4. Thin line: optimally accelerated (with respect to parameter c in operator K, see text) ITEM (2.15)–(2.17) without mode elimination. Thick and medium lines: ITEM (4.10)–(4.13) with mode elimination with $s = 1$ (thick) and $s = 0.7$ (medium). The application of mode elimination begins at the first iteration.

Thus, from the last three examples, we conclude that in those cases when the iterations converge slowly and their acceleration is highly desirable, the mode elimination method provides a considerable improvement of the convergence rate (by a factor of several times). Taking $s < 1$, so that only part of the mode $(u_n - u_{n-1})$ would be eliminated, usually results in smoother convergence; however, the choice $s = 1$ still yields a considerable improvement of the convergence rate in comparison with that of the non-accelerated iteration method. For these slowly convergent cases, the Steffensen's method, based on the Aitken's acceleration, often diverges.

Remark. In those cases when the step size $\Delta\tau$ is nearly optimal, the error is expected to be dominated by two eigenmodes, corresponding to $\Lambda_{\max}$ and $\Lambda_{\min}$, since

$$(1 + \Lambda_{\max}\Delta\tau) \approx -(1 + \Lambda_{\min}\Delta\tau) \tag{4.14}$$

for this $\Delta\tau$ (see (2.4) and (2.6)). Then it seems logical that one would need to eliminate *both* of these eigenmodes, which are proportional to $(u_n - u_{n-2})$ and $(u_n - 2u_{n-1} + u_{n-2})$, respectively. We found, however, that although this does result in a smoother error evolution than the elimination of just the single mode $(u_n - u_{n-1})$, it does not yield any consistent improvement of the convergence rate compared to the latter case.

5. Summary

In this work, we obtained the following results.

In Section 2, we compared the linearized operators of the generalized Petviashvili method and the ITEM with power normalization. In particular, we showed that while the "primordial" part of those operators has the same form (2.19), their nonlocal parts involving the inner products are different, leading to the eigenvalues of the corresponding operators being different. In our simulations we observed that the generalized Petviashvili method converges to fundamental solitary waves faster than does the ITEM (although we could prove this rigorously only for equations with power-law nonlinearity). On the other hand, in those (rare) cases when both methods converge to a nonfundamental solitary wave, we produced an explicit example where the ITEM is faster.

In Section 3, we proposed a new technique, which we referred to as the mode elimination. One application of this technique is that it can obtain nonfundamental solitary waves, for which the generalized Petviashvili method would otherwise diverge. The corresponding iteration scheme is given by Eqs. (3.3) and (3.4). In Example 4.1 in Section 4, we demonstrated that this technique can be superior to an alternative, squared-operator, technique [8] when applied to finding *lowest-order* nonfundamental solutions. However, for finding higher-order solutions, the technique of Ref. [8] appears to be more practical.

As a *more important application* for the mode elimination technique, we showed that it can accelerate the convergence of various iteration methods. This acceleration is most significant (by a factor of several times) in those cases when it is most needed, i.e., when the convergence of the non-accelerated method is slow. The iteration schemes implementing this technique are: Eqs. (3.7), (3.8) for the generalized Petviashvili method; Eqs. (4.7), (4.8) for a squared-operator method (see also Ref. [8]); and Eqs. (4.10)–(4.13) for the ITEM with power normalization.

Acknowledgments

The work of T.I.L was supported in part by the National Science Foundation under Grant DMS-0507429, and the work of J.Y. was supported in part by The Air Force Office of Scientific Research under grant USAF 9550-05-1-0379.

References

[1] T.I. Lakoba, J. Yang, A generalized Petviashvili iteration method for scalar and vector Hamiltonian equations with arbitrary form of nonlinearity, J. Comput. Phys., this issue, doi:10.1016/j.jcp.2007.06.009.

[2] V.I. Petviashvili, Equation for an extraordinary soliton, Sov. J. Plasma Phys. 2 (1976) 257–258.

[3] J.J. Garcia-Ripoll, V.M. Perez-Garcia, Optimizing Schrödinger functionals using Sobolev gradients: applications to quantum mechanics and nonlinear optics, SIAM J. Sci. Comput. 23 (2001) 1316–1334.

[4] W. Bao, Q. Du, Computing the ground state solution of Bose–Einstein condensates by a normalized gradient flow, SIAM J. Sci. Comput. 25 (2004) 1674–1697.
[5] V.S. Shchesnovich, S.B. Cavalcanti, Rayleigh functional for nonlinear systems. Available from: <http://www.arXiv.org> (preprint nlin.PS/0411033).
[6] J. Yang, T.I. Lakoba, Convergence and acceleration of imaginary-time evolution methods for solitary waves in arbitrary spatial dimensions, submitted.
[7] D.E. Pelinovsky, Yu.A. Stepanyants, Convergence of Petviashvili's iteration method for numerical approximation of stationary solutions of nonlinear wave equations, SIAM J. Numer. Anal. 42 (2004) 1110–1127.
[8] J. Yang, T. Lakoba, Universally-convergent squared-operator iteration methods for solitary waves in general nonlinear wave equations, Stud. Appl. Math. 118 (2007) 153–197.
[9] J.-P. Chehab, J. Laminie, Differential equations and solution of linear systems, Numer. Algorithms 40 (2005) 103–124.
[10] S.V. Manakov, V.E. Zakharov, L.A. Bordag, A.R. Its, V.B. Matveev, Two-dimensional solitons of the Kadomtsev–Petviashvili equation and their interaction, Phys. Lett. A 63 (1977) 205.
[11] M. Weinstein, Lyapunov stability of ground states of nonlinear dispersive evolution equations, Comm. Pure Appl. Math. 39 (1986) 51–67.
[12] R. Horn, C. Johnson, Matrix Analysis, Cambridge University Press, New York, 1991.
[13] L.W. Johnson, R.D. Riess, Numerical Analysis, Addison-Wesley Publishing Co., Reading, 1977, Problem 8 for Section 4.3.2.
[14] L. Demanet, W. Schlag, Numerical verification of a gap condition for a linearized nonlinear Schrödinger equation, Nonlinearity 19 (2006) 829–852.