Numerical Solution to a Linearized KdV Equation on Unbounded Domain

Chunxiong Zheng, 1 Xin Wen, 2 Houde Han 1,3

1 Department of Mathematical Sciences, Tsinghua University, Beijing 100084, People’s Republic of China

2 Institute of Computational Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100080, China

3 Department of Mathematics, University of Science and Technology of China, Hefei 230026, People’s Republic of China

Received 3 October 2006; accepted 10 March 2007
Published online 4 June 2007 in Wiley InterScience (www.interscience.wiley.com).
DOI 10.1002/num.20267

Exact absorbing boundary conditions for a linearized KdV equation are derived in this paper. Applying these boundary conditions at artificial boundary points yields an initial-boundary value problem defined only on a finite interval. A dual-Petrov-Galerkin scheme is proposed for numerical approximation. Fast evaluation method is developed to deal with convolutions involved in the exact absorbing boundary conditions. In the end, some numerical tests are presented to demonstrate the effectiveness and efficiency of the proposed method. © 2007 Wiley Periodicals, Inc. Numer Methods Partial Differential Eq 24: 383–399, 2008

Keywords: absorbing boundary conditions; dual-Petrov-Galerkin method; linearized KdV equation; unbounded domains

I. INTRODUCTION

The Korteweg-de Vries (KdV) equation

\[ u_t - 6uu_x + u_{xxx} = 0 \]  

(1.1)

has attracted attention of both physical scientists and mathematicians, since it was found to admit soliton solutions and be able to model the propagation of solitary wave on the water surface, a phenomena first discovered by Scott Russell in 1834 [1]. As a matter of fact, later studies revealed
that the KdV equation can be widely derived when nonlinearity is of the same order as cubic dispersion. Such situations are very typical in hydrodynamics, physics, and acoustics [2, 3].

In the case when a PDE is employed to model waves on an unbounded domain and the numerical simulation is performed, it is a common practice to truncate the unbounded domain by introducing artificial boundaries, which necessitates additional boundary conditions to be designed. A proper choice of these boundary conditions should ensure both stability and accuracy of the truncated initial-boundary value problem. In the engineering language, these boundary conditions should mimic the absorption of waves traveling through the artificial boundaries to the exterior domain. Right in this context, they are popularly referred to as absorbing boundary conditions (ABC) in the literature.

A simple treatment of ABCs is to make the truncated computational domain large enough, such that the numerical simulation terminates before significant wave reflection comes into effect on the region which bears special physical interest, see [4, 5]. Obviously, this treatment greatly increases the computational cost in terms of time and storage, and more delicate ABCs are thus needed to reduce the computational burden. In [6], for the modified KdV equation
\[
\frac{d^2}{dx^2} u^2 + u_x + 6 u^2 u_x + u_{xxx} = 0,
\]
Zheng modified the ABCs proposed by Fokas [7] based on the inverse scattering theory, and designed a suitable numerical scheme. But unfortunately, the extension to the nonlinear KdV equation (1.1) seems not straightforward.

This paper is concerned with the ABCs for a linearized KdV equation of the form
\[
\frac{d}{dt} u_t + u_x = h,
\]
where \( h \) is the source term. Though much simplified compared with the nonlinear KdV equation (1.1), this equation itself has many applications. For example, it is usually used to study the propagation of fairly long waves in the shallow water [3]. Linearity property makes possible to use many effective analytical tools such as the Laplace transformation [8]. With this tool, we can derive the exact ABCs for equation (1.2) at fixed boundary points.

The organization of the rest is as follows. In Section II, we derive exact ABCs for equation (1.2) when a locally supported initial function is prescribed. In Section III, a dual-Petrov-Galerkin spectral method is proposed for the numerical approximation. In Section IV, by following Jiang and Greengard [9] we develop fast evaluation method for the exact ABCs. This method is indispensable for long time computation due to the existence of convolutions in the exact ABCs. Numerical tests are presented in Section V, and we conclude this paper in Section VI.

II. DERIVATION OF EXACT ABCs

We study the Cauchy problem
\[
\begin{align*}
\frac{d}{dt} u_t + u_x & = h(x,t), \quad x \in \mathcal{R}, \\
\frac{d}{dt} u(x,0) & = u_0(x), \quad x \in \mathcal{R}, \\
u & \to 0, \quad x \to \infty.
\end{align*}
\]

The initial function \( u_0 \) and the source function \( h \) are assumed to be compactly supported in a finite interval \([a, b]\), with \( a < b \). To compute the numerical solution of (2.1)–(2.3) restricted
to \([a, b]\), some boundary conditions are needed to be designed and applied at boundary points \([a, b]\). As formulated as follows, this goal can be achieved by analyzing the equation (2.1) on the complementary domain of \([a, b]\).

More precisely, let us consider the following two problems:

\[
\begin{align*}
 u_t + u_{xxx} &= 0, \quad x > b, \\
 u(x, 0) &= 0, \quad x > b, \\
 u &\to 0, \quad x \to +\infty,
\end{align*}
\]

and

\[
\begin{align*}
 u_t + u_{xxx} &= 0, \quad x < a, \\
 u(x, 0) &= 0, \quad x < a, \\
 u &\to 0, \quad x \to -\infty.
\end{align*}
\]

These problems are incomplete, and cannot be solved independently. If \(u(b, t)\) is prescribed, we can derive two relations satisfied by the Dirichlet function \(u(b, t)\) and the Neumann functions \(u_x(b, t)\) and \(u_{xx}(b, t)\) of the problem (2.4)–(2.6). Analogously, for the problem (2.7)–(2.9), if \(u(a, t)\) is prescribed, we can obtain one relation satisfied by the Dirichlet function \(u(a, t)\) and the Neumann functions \(u_x(a, t)\) and \(u_{xx}(a, t)\).

Doing the Laplace transformation on equations (2.4) and (2.7) we get

\[
s \hat{u} + \hat{u}_{xxx} = 0, \quad (2.10)
\]

where \(s \in C\) with \(\Re s > 0\) is the argument in the Laplace-transformed space, and \(\hat{u}\) is the transformed function of \(u\). The general solution to the above equation (2.10) is given explicitly as

\[
\hat{u}(x, s) = c_1(s)e^{\lambda_1(s)x} + c_2(s)e^{\lambda_2(s)x} + c_3(s)e^{\lambda_3(s)x}, \quad (2.11)
\]

with

\[
\lambda_1(s) = -\sqrt{s}, \quad \lambda_2(s) = -\sqrt{s}\omega, \quad \lambda_3(s) = -\sqrt{s}\omega^2, \quad \omega = e^{\frac{2\pi i}{3}}.
\]

It is straightforward to verify that

\[
\Re \lambda_1(s) < 0, \quad \Re \lambda_2(s) > 0, \quad \Re \lambda_3(s) > 0. \quad (2.12)
\]

For the problem (2.4)–(2.6), \(c_2(s)\) and \(c_3(s)\) in (2.11) must equal zero thanks to the infinity condition (2.6). Thus then, we have

\[
\hat{u}(b, s) - \frac{1}{\lambda_1^2} \hat{u}_{xx}(b, s) = 0, \quad \hat{u}_x(b, s) - \frac{1}{\lambda_1} \hat{u}_{xx}(b, s) = 0. \quad (2.13)
\]

The inverse Laplace transformation on equations (2.13) yields

\[
\begin{align*}
 u(b, t) - I_1^{\frac{1}{2}} u_{xx}(b, t) &= 0, \\
 u_x(b, t) + I_1^{\frac{1}{2}} u_{xx}(b, t) &= 0,
\end{align*}
\]

where \(I_\alpha^\alpha\) with \(\alpha > 0\) is the fractional integral operator defined by

\[
I_\alpha^\alpha f(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t - \tau)^{\alpha - 1} f(\tau)d\tau, \quad t > 0. \quad (2.15)
\]

The readers are referred to [10] for more detail.
For the problem (2.7)–(2.9), \( c_1(s) \) in the general solution (2.11) must equal zero owing to the infinity condition (2.9). In this case, we have
\[
\hat{u}(a, s) + \frac{1}{\lambda_1} \hat{u}_x(a, s) + \frac{1}{\lambda_1^2} \hat{u}_{xx}(a, s) = 0.
\]
The inverse Laplace transformation thus gives
\[
u(a, t) - I_1^\frac{1}{3} u_x(a, t) + I_2^\frac{2}{3} u_{xx}(a, t) = 0.
\] (2.16)
In fact, the equations (2.14) and (2.16) define three exact ABCs satisfied by the solution \( u \) of (2.1)–(2.3). Applying them at boundary points \( \{a, b\} \), we derive a truncated initial-boundary value problem:
\[
\begin{align*}
\frac{\partial u}{\partial t} + \frac{1}{3} u_{xxx} &= h(x, t), \quad x \in [a, b], \\
u(x, 0) &= u_0(x), \quad x \in [a, b], \\
u(a, t) - I_1^\frac{1}{3} u_x(a, t) + I_2^\frac{2}{3} u_{xx}(a, t) &= 0, \\
u(b, t) - I_1^\frac{2}{3} u_{xx}(b, t) &= 0,
\end{align*}
\] (2.17)–(2.21)
According to our analysis, if the truncated problem (2.17)–(2.20) has only one solution, this solution should be the same as that of (2.1)–(2.3) being restricted to \([a, b]\). The existence of solution for (2.17)–(2.20) is unclear to us, and we leave it open in this paper. As to the uniqueness of solution, it is implied by the stability estimate in the following theorem, which is derived by showing the kernel functions involved in the exact ABCs (2.19)–(2.21) are of positive type in the sense of memory type, a similar technique used in [13] for the linear Schrödinger equation.

**Theorem 1.** The problem (2.17)–(2.20) is \( L^2 \)-stable. More precisely, for any \( t > 0 \), there is a constant positive number \( c(t) \), such that
\[
\int_a^b u^2(x, t)dx \leq c(t) \left( \int_a^b u_0^2(x)dx + \int_0^t \int_a^b h^2(x, t)dxdt \right).
\] (2.22)

**Proof.** Multiplying both sides of (2.17) with \( 2u \) and integrating by parts gives
\[
\frac{d}{dt} \int_a^b u^2(x, t)dx = -2 \int_a^b uu_{xxx}dx + \int_a^b 2hudx
\]
\[
= -2 \int_a^b \left( uu_{xx} - \frac{1}{2} u^2_x \right)_x dx + \int_a^b 2hudx
\]
\[
= 2 \left( uu_{xx} - \frac{1}{2} u^2_x \right)_x (a, t) - 2 \left( uu_{xx} - \frac{1}{2} u^2_x \right)_x (b, t) + \int_a^b 2hudx
\]
\[
= \left( 2(I_1^\frac{1}{3} u_x - I_2^\frac{2}{3} u_{xx})_x (a, t) - 2(2u_x I_1^\frac{1}{3} u_{xx} - (I_2^\frac{2}{3} u_{xx})^2)_x (b, t) + \int_a^b 2hudx. \right)
\]

The last equality holds thanks to the ABCs (2.19)–(2.21). Integrating both sides of the above equation with respect to \( t \) we get

\[
\int_a^b u^2(x,t)dx - \int_a^b u_0^2(x)dx = I_t^1 \left( 2 \left( I_t^1 u_x - I_t^2 u_{xx} \right) u_x - u_x^2 \right) (a,t) \\
- I_t^1 \left( 2u_{xx} I_t^3 u_x - \left( I_t^3 u_x \right)^2 \right) (b,t) + \int_0^t \int_a^b 2hudxdt.
\]

In view of Lemma 1 in the following, we have

\[
\int_a^b u^2(x,t)dx - \int_a^b u_0^2(x)dx \leq \int_0^t \int_a^b u^2(x,t)dxdt + \int_0^t \int_a^b h^2(x,t)dxdt.
\]

The proof finishes by the standard Gronwall’s inequality. ■

**Lemma 1.** For any smooth functions \( f(t) \) and \( g(t) \) with \( f(0) = g(0) = 0 \), it holds

\[
I_t^1 \left( 2f I_t^3 f - \left( I_t^3 f \right)^2 \right) \geq 0, \tag{2.23}
\]

\[
I_t^1 \left( 2 \left( I_t^3 g - I_t^3 f \right) f - g^2 \right) \leq 0. \tag{2.24}
\]

**Proof.** For any \( T > 0 \), we set

\[
F(t) = f(t)I_{[0,T]}, \quad G(t) = g(t)I_{[0,T]}.
\]

Thus then,

\[
I_t^1 \left( 2f I_t^3 f - \left( I_t^3 f \right)^2 \right) \big|_{t=T} = I_t^1 \left( 2F I_t^3 F - \left( I_t^3 F \right)^2 \right) \big|_{t=T} \geq \int_0^{+\infty} \left( 2FI_t^3 F - \left( I_t^3 F \right)^2 \right) dt.
\]

Using the Plancherel equality for the Laplace transformation yields

\[
\int_0^{+\infty} \left( 2FI_t^3 F - \left( I_t^3 F \right)^2 \right) dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left( 2(i\xi)^{-\frac{2}{3}} - |\xi|^{-\frac{2}{3}} \right) |\hat{F}(i\xi)|^2 d\xi = 0.
\]

The last equality holds since \( |\hat{F}(i\xi)|^2 = |\hat{F}(-i\xi)|^2 \). In addition, we have

\[
I_t^1 \left( 2 \left( I_t^3 g - I_t^3 f \right) f - g^2 \right) \big|_{t=T} \\
= \int_0^{+\infty} \left( 2I_t^1 G - I_t^3 F \right) F - G^2 \right) dt \\
= \int_0^{+\infty} \left( 2FI_t^1 F - \left( I_t^3 F \right)^2 \right) dt \\
= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left( 2(i\xi)^{-\frac{1}{3}} \hat{G}(i\xi) \hat{F}(i\xi) - |\xi|^{-\frac{2}{3}} |\hat{F}(i\xi)|^2 - |\hat{G}(i\xi)|^2 \right) d\xi
\]

*Numerical Methods for Partial Differential Equations* DOI 10.1002/num
\[
\leq \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left( 2|\xi|^{-\frac{3}{2}} |\hat{G}(i\xi)||\hat{F}(i\xi)| - |\xi|^{-\frac{3}{2}} |\hat{F}(i\xi)|^2 - |\hat{G}(i\xi)|^2 \right) d\xi \\
\leq 0.
\]

The proof is complete.

### III. NUMERICAL IMPLEMENTATION

Let \( \Delta t \) denote the time step, and \( t_m = m \Delta t \) the \( m \)-th discrete time point. For any function \( f(t) \) let \( f^m \approx f(t_m) \). We approximate the fractional integral \( I^\alpha f \) by replacing \( f \) with a suitable interpolating function \( \tilde{f} \) at points \( \{t_0, t_1, \ldots\} \). If the piecewise linear interpolation is used, we obtain

\[
I^\alpha_t f(t_m) \approx I^\alpha_t \tilde{f}(t_m) = \sum_{j=0}^{m} \beta_{\alpha,m-j} f_j,
\]

with

\[
\beta_{\alpha,j} = \frac{(\Delta t)\alpha}{\alpha(\alpha + 1)\Gamma(\alpha)} \times \left\{ \begin{array}{cl} 1, & j = 0, \\ (j + 1)^{\alpha+1} + (j - 1)^{\alpha+1} - 2j^{\alpha+1}, & j > 0. \end{array} \right.
\]

At \( t = t_m \), the exact ABCs (2.19)–(2.21) are discretized as

\[
u^m(a) - \beta_{\frac{1}{3},0} u_x^m(a) + \beta_{\frac{2}{3},0} u_{xx}^m(a) = \sum_{j=1}^{m} \beta_{\frac{1}{3},j} u_x^{m-j}(a) - \sum_{j=1}^{m} \beta_{\frac{2}{3},j} u_{xx}^{m-j}(a) \equiv c_1^m, \tag{3.1}
\]
\[
u^m(b) - \beta_{\frac{2}{3},0} u_{xx}^m(b) = \sum_{j=1}^{m} \beta_{\frac{2}{3},j} u_{xx}^{m-j}(b) \equiv c_2^m, \tag{3.2}
\]
\[
u_x^m(b) + \beta_{\frac{1}{3},0} u_{xx}^m(b) = -\sum_{j=1}^{m} \beta_{\frac{1}{3},j} u_{xx}^{m-j}(b) \equiv c_3^m. \tag{3.3}
\]

To time variable of the governing equation (2.1), we use the standard Crank-Nicolson discretization, and thus obtain

\[
\frac{u^m - u^{m-1}}{\Delta t} + \frac{u_{xxx}^m + u_{xxx}^{m-1}}{2} = \frac{h^m + h^{m-1}}{2}. \tag{3.4}
\]

The spatial discretization can be accomplished in many different ways. In this paper, we propose a spectral scheme adapted from the dual-Petrov-Galerkin method in Shen [11].

We define

\[
V_N = \left\{ u \in P_N : u(a) - \beta_{\frac{1}{3},0} u_x(a) + \beta_{\frac{2}{3},0} u_{xx}(a) = 0, u(b) - \beta_{\frac{2}{3},0} u_{xx}(b) = 0, u_x(b) + \beta_{\frac{1}{3},0} u_{xx}(b) = 0 \right\}
\]
and its dual space

\[ V_N^* = \left\{ v \in P_N : v(b) + \beta_{\frac{3}{2},0} v_x(b) + \beta_{\frac{3}{2},0} v_{xx}(b) = 0, v(a) - \beta_{\frac{3}{2},0} v_x(a) = 0, v_x(a) - \beta_{\frac{3}{2},0} v_{xx}(a) = 0 \right\}, \]

where \( P_N \) stands for the \( N \)-degree polynomial space. The dual-Petrov-Galerkin approximation to (3.4) reads:

Find \( u^m \in P_N \), such that \( \forall v \in V_N^* \),

\[
\frac{1}{\Delta t}(u^m, v) + \frac{1}{2}(u_{xxx}^m, v) = \frac{1}{\Delta t}(u^{m-1}, v) - \frac{1}{2}(u_{xxx}^{m-1}, v) + \frac{1}{2}(h^m + h^{m-1}, v),
\]

and the equations (3.1)–(3.3) are satisfied.

Suppose \( L_k \) is the \( k \)-th Legendre polynomial function on \([a, b]\). By defining

\[
\begin{pmatrix} \phi_{N-2} \\ \phi_{N-1} \\ \phi_N \end{pmatrix} = A \begin{pmatrix} L_0 \\ L_1 \\ L_2 \end{pmatrix},
\]

with \( 3 \times 3 \) matrix \( A \) determined (it is possible for any \( \Delta t \)) by

\[
(1 - \beta_{\frac{3}{2},0} \partial_x + \beta_{\frac{3}{2},0} \partial_x^2)\phi_{N-2}(a) = 1, \quad (1 - \beta_{\frac{3}{2},0} \partial_x + \beta_{\frac{3}{2},0} \partial_x^2)\phi_{N-2}(b) = 0, \quad (\partial_x + \beta_{\frac{3}{2},0} \partial_x^2)\phi_{N-2}(b) = 0,
\]

\[
(1 - \beta_{\frac{3}{2},0} \partial_x + \beta_{\frac{3}{2},0} \partial_x^2)\phi_{N-1}(a) = 0, \quad (1 - \beta_{\frac{3}{2},0} \partial_x + \beta_{\frac{3}{2},0} \partial_x^2)\phi_{N-1}(b) = 1, \quad (\partial_x + \beta_{\frac{3}{2},0} \partial_x^2)\phi_{N-1}(b) = 0,
\]

\[
(1 - \beta_{\frac{3}{2},0} \partial_x + \beta_{\frac{3}{2},0} \partial_x^2)\phi_N(a) = 0, \quad (1 - \beta_{\frac{3}{2},0} \partial_x + \beta_{\frac{3}{2},0} \partial_x^2)\phi_N(b) = 0, \quad (\partial_x + \beta_{\frac{3}{2},0} \partial_x^2)\phi_N(b) = 1,
\]

the variational problem (3.5) is transformed into:

Find \( u^m = u^m + c_1^m \phi_{N-2} + c_2^m \phi_{N-1} + c_3^m \phi_N \) with \( w^m \in V_N \), such that \( \forall v \in V_N^* \),

\[
\frac{1}{\Delta t}(w^m, v) + \frac{1}{2}(w_{xxx}^m, v) = \frac{1}{\Delta t}(w^{m-1}, v) - \frac{1}{2}(w_{xxx}^{m-1}, v) - \frac{1}{2}(h^m + h^{m-1}, v),
\]

where constants \( c_1^m, c_2^m \) and \( c_3^m \) are given in the equations (3.1)–(3.3).

For any \( 0 \leq k \leq N - 3 \), set

\[
\phi_k = L_k + \alpha_k L_{k+1} + \beta_k L_{k+2} + \gamma_k L_{k+3},
\]

\[
\psi_k = L_k - \alpha_k L_{k+1} + \beta_k L_{k+2} - \gamma_k L_{k+3}
\]

such that \( \phi_k \in V_N \) and \( \psi_k \in V_N^* \). It can be verified directly that \( \alpha_k, \beta_k \) and \( \gamma_k (\neq 0) \) can be uniquely determined. Thus then, \( \{\phi_k\}^{N-3}_{k=0} \) and \( \{\psi_k\}^{N-3}_{k=0} \) constitute a basis of \( V_N \) and \( V_N^* \), respectively. The advantage of the dual-Petrov-Galerkin method lies in the fact that

\[
(\partial_x^2 \phi_i, \psi_j) = - (\phi_i, \partial_x^2 \psi_j) = 0, \quad i \neq j.
\]

Thanks to the orthogonality of Legendre polynomials, the coefficient matrix for (3.6) is seven-diagonal, and the linear algebraic system can be solved efficiently within \( O(N) \) operations. See [11] for more detailed discussion.

KdV EQUATION ON UNBOUNDED DOMAIN 389
IV. FAST EVALUATION METHOD

On the $m$-th time step, solving the variational problem (3.6) necessitates computing $c_1^m$, $c_2^m$ and $c_3^m$ defined in equations (3.1)–(3.3), which requires $O(m)$ operations. This becomes much costly if $m$ is large. Following the idea of Jiang and Greengard [9] on the exact ABC of the linear Schrödinger equation, we divide the fractional integral into two parts

$$I_1 = \int_{t_{m-1}}^{t_m} (t_m - \tau)^{\alpha - 1} f(\tau) d\tau$$

$I_1$ is called the local part, which is approximated by replacing $f$ with its piecewise linear interpolating function $\tilde{f}$. A direct computation gives

$$I_1 \approx \frac{1}{\Gamma(\alpha)} \int_{t_{m-1}}^{t_m} (t_m - \tau)^{\alpha - 1} \tilde{f}(\tau) d\tau = \frac{(\Delta t)^{\alpha}}{\Gamma(\alpha)} \left( \frac{1}{\alpha(\alpha + 1)} f^m + \frac{1}{\alpha + 1} f^{m-1} \right).$$

$I_2$ is called the history part, which is a normal integral without singularity. On a finite time interval $[\Delta t, T]$, the kernel function can be approximated uniformly with a sum of decaying exponentials for any prescribed relative error $\epsilon$, i.e.,

$$\left| t^{\alpha - 1} - \sum_{i=1}^{M_\alpha} \delta_{i,\alpha} e^{-\sigma_{i,\alpha} t} \right| \leq \epsilon t^{\alpha - 1}, \forall t \in [\Delta t, T]. \quad (4.1)$$

Thus then,

$$I_2 \approx \frac{1}{\Gamma(\alpha)} \sum_{i=1}^{M_\alpha} \delta_{i,\alpha} \int_{0}^{t_{m-1}} e^{-\sigma_{i,\alpha}(t_{m-1} - \tau)} \tilde{f}(\tau) d\tau = \frac{1}{\Gamma(\alpha)} \sum_{i=1}^{M_\alpha} \delta_{i,\alpha} A_{i,\alpha}^m(f)$$

with

$$A_{i,\alpha}^m(f) \equiv \int_{0}^{t_{m-1}} e^{-\sigma_{i,\alpha}(t_{m-1} - \tau)} \tilde{f}(\tau) d\tau.$$ 

Note that $A_{i,\alpha}^1(f) = 0$, and for $m > 1$,

$$A_{i,\alpha}^m(f) = \int_{0}^{t_{m-2}} e^{-\sigma_{i,\alpha}(t_{m-2} - \tau)} \tilde{f}(\tau) d\tau + \int_{t_{m-2}}^{t_{m-1}} e^{-\sigma_{i,\alpha}(t_{m-1} - \tau)} \tilde{f}(\tau) d\tau$$

$$= e^{-\sigma_{i,\alpha} \Delta t} \left( A_{i,\alpha}^{m-1}(f) + \frac{\sigma_{i,\alpha} \Delta t - 1}{\sigma_{i,\alpha}^2 \Delta t} f^{m-1} + \frac{1 - e^{-\sigma_{i,\alpha} \Delta t}}{\sigma_{i,\alpha}^2 \Delta t} f^{m-2} \right).$$

The fractional integral $I_1^\alpha f$ is then approximated by

$$I_1^\alpha f(t_m) \approx \frac{(\Delta t)^\alpha}{\Gamma(\alpha)} \left( \frac{1}{\alpha(\alpha + 1)} f^m + \frac{1}{\alpha + 1} f^{m-1} \right) + \frac{1}{\Gamma(\alpha)} \sum_{i=1}^{M_\alpha} \delta_{i,\alpha} A_{i,\alpha}^m(f).$$

The dual-Petrov-Galerkin approximation with fast evaluation to the problem (2.17)–(2.21) is: Find $u^m = u^m + d_1^m \phi_{N-2} + d_2^m \phi_{N-1} + d_3^m \phi_N$ with $u^m \in V_N$, such that $\forall v \in V_N^*$,

$$
\frac{1}{\Delta t}(u^m, v) + \frac{1}{2}(u^m_{xx}, v) = \frac{1}{\Delta t}(u^{m-1} - d_1^m \phi_{N-2} - d_2^m \phi_{N-1} - d_3^m \phi_N, v) - \frac{1}{2}(u^{m-1}_{xx}, v) + \frac{1}{2}(h^m + h^{m-1}, v),
$$

with

$$
d_1^m = \frac{2}{5} (\Delta t)^\frac{1}{3} (u^{m-1}_x(a) + \frac{1}{\Gamma(\frac{1}{3})} \sum_{i=1}^{M_1} \delta_i \frac{1}{3} A^m_{i \frac{2}{3}} (u_x(a, \cdot))
$$

$$
- \frac{2}{5} (\Delta t)^\frac{1}{3} u^{m-1}_x(a) - \frac{1}{\Gamma(\frac{2}{3})} \sum_{i=1}^{M_2} \delta_i \frac{2}{3} A^m_{i \frac{2}{3}} (u_x(a, \cdot)),
$$

$$
d_2^m = \frac{2}{5} (\Delta t)^\frac{1}{3} (u^{m-1}_x(b) + \frac{1}{\Gamma(\frac{1}{3})} \sum_{i=1}^{M_1} \delta_i \frac{1}{3} A^m_{i \frac{2}{3}} (u_x(b, \cdot))
$$

$$
- \frac{2}{5} (\Delta t)^\frac{1}{3} u^{m-1}_x(b) - \frac{1}{\Gamma(\frac{2}{3})} \sum_{i=1}^{M_2} \delta_i \frac{2}{3} A^m_{i \frac{2}{3}} (u_x(b, \cdot)),
$$

At the $m$-th time step, solving (4.2) requires $O(N + M_1^\frac{1}{3} + M_2^\frac{1}{3})$ operations, which is independent of $m$. Compared with $O(N + m)$ operations required for solving (3.6), this is a remarkable saving of computational cost if the total number of time steps is large.

Now we explain how to arrive at (4.1) for $\alpha = 1/3$ and $\alpha = 2/3$. In the following, for any real number $x$, $|x|$ denotes the smallest integer which is not less than $x$, while $\lfloor x \rfloor$ denotes the largest integer which is not larger than $x$.

We start with the integral equality

$$
\frac{1}{3} \Gamma \left( \frac{1}{3} \right) r^{-\frac{1}{3}} = \int_0^{+\infty} e^{-x^3} ds = \int_0^{2^M} e^{-x^3} ds + \int_{2^M}^{2^N} e^{-x^3} ds + \int_{2^N}^{+\infty} e^{-x^3} ds \equiv I_1 + I_2 + I_3 \equiv I,
$$

where $M$ and $N$ are two integers to be determined. Since

$$
I_3 = \int_2^{+\infty} e^{-x^3} ds = \int_0^{+\infty} e^{-(s+2^N)^3} ds \leq e^{-2^{3N} \Delta t}, \forall t \geq \Delta t,
$$

the following lemma is straightforward.

**Lemma 2.** For $\epsilon > 0$ and $t \geq \Delta t$, let $N = \lfloor \frac{1}{3} \log_{\frac{1}{2}} \frac{\ln \epsilon}{2 \Delta t} \rfloor$, then we have $I_3 \leq \epsilon I$.

Next we consider $I_1$. Let $\{s_{M,n,k}\}_{k=1}^n$ and $\{w_{M,n,k}\}_{k=1}^n$ be the nodes and weights for $n$-point Gauss-Legendre quadrature on the interval $[0, 2^M]$, then the standard error estimate gives

$$
\left| I_1 - \sum_{k=1}^n w_{M,n,k} e^{-s_{M,n,k}} \right| = \frac{2^{M(n+1)}}{2n + 1} \frac{(n!)^4}{(2n)!} \left| D_{\xi}^{2n} e^{-\xi^3} \right|, 0 < \xi < 2^M.
$$

Let $x = \xi t^{\frac{1}{3}}$. Then

$$|D_t^{2n} e^{-x^3}| = t^{\frac{2n}{3}} |D_x^{2n} e^{-x^3}| \leq t^{\frac{1}{3} - 1} T^{\frac{2n+1}{3}} |D_x^{2n} e^{-x^3}|, \quad \forall t \leq T.$$ 

Denoting

$$G_1(n) = \frac{3}{\Gamma\left(\frac{1}{3}\right)} \frac{1}{2n + 1} \frac{(n!)^4}{((2n)!)^3} \max_{x \in \mathbb{R}^n} |D_x^{2n} e^{-x^3}|,$$

we have

$$\left| I_1 - \sum_{k=1}^{n} w_{M,n,k}^* e^{-\left(s_{M,n,k}^*\right)^3} \right| \leq \frac{\left(2MT^{\frac{1}{3}}\right)^{2n+1}}{2n + 1} \frac{(n!)^4}{((2n)!)^3} t^{\frac{1}{3}} |D_x^{2n} e^{-x^3}| \leq \left(2MT^{\frac{1}{3}}\right)^{2n+1} G_1(n) I.$$ 

**Lemma 3.** Let $M = \left\lfloor -\frac{1}{3} \log_2 T \right\rfloor$, and let $n^* = n^*(\epsilon_1)$ be the smallest positive integer satisfying $G_1(n^*) \leq \epsilon_1$ for $\epsilon_1 > 0$, then we have

$$\left| I_1 - \sum_{k=1}^{n^*} w_{M,n,k}^* e^{-\left(s_{M,n,k}^*\right)^3} \right| \leq \epsilon_1 I.$$ 

This lemma holds since $2MT^{\frac{1}{3}} \leq 1$ and $G_1(n) \to 0$ as $n \to \infty$, see Figure 1.

Now we consider $I_2$. Since

$$I_2 = \sum_{j=M}^{N-1} \int_{2j}^{2j+1} e^{-x^3} ds,$$

FIG. 1. Plot of functions $G_1(\cdot)$ and $G_3(\cdot)$.
we compute the integral on the dyadic interval $[2^j, 2^{j+1}]$ first. Let \( \{s_{j,n,k}\}_{k=1}^{n} \) and \( \{w_{j,n,k}\}_{k=1}^{n} \) be the nodes and weights for n-point Gauss-Legendre quadrature on the interval, then we have

\[
\left| \int_{2^j}^{2^{j+1}} e^{-x^3} \, ds - \sum_{k=1}^{n} w_{j,n,k} e^{-s_{j,n,k}^3} \right| = \frac{2^{j(2n+1)}}{2n+1} \left( \frac{(n!)^4}{[(2n)!]^3} \right) |D_{\xi}^{2n} e^{-\xi^3}|, \quad 2^j < \xi < 2^{j+1}.
\]

Let \( x = \xi \frac{1}{t^3} \). Then

\[
|D_{\xi}^{2n} e^{-\xi^3}| = t^{\frac{2n}{3}} |D_{x}^{2n} e^{-x^3}| = t^{\frac{1}{3}} \left( \frac{x}{\xi} \right)^{2n+1} |D_{x}^{2n} e^{-x^3}| \leq t^{\frac{1}{3}} 2^{-j(2n+1)} x^{2n+1} |D_{x}^{2n} e^{-x^3}|.
\]

Denoting

\[
G_2(n) \equiv \frac{3}{\Gamma \left( \frac{1}{3} \right)} \frac{1}{2n+1} \frac{(n!)^4}{[(2n)!]^3} \max_{x \in \mathbb{R}^+} x^{2n+1} |D_{x}^{2n} e^{-x^3}|,
\]

we have

\[
\left| \int_{2^j}^{2^{j+1}} e^{-x^3} \, ds - \sum_{k=1}^{n} w_{j,n,k} e^{-s_{j,n,k}^3} \right| \leq \frac{1}{2n+1} \left( \frac{(n!)^4}{[(2n)!]^3} \right) t^{\frac{1}{3}} x^{2n+1} |D_{x}^{2n} e^{-x^3}| \leq G_2(n) I.
\]

**Lemma 4.** For \( \epsilon_2 > 0 \), let \( n = n(\epsilon_2) \) be the smallest positive integer satisfying \( G_2(n) \leq \epsilon_2 \), then we have

\[
\left| I_2 - \sum_{j=M}^{N-1} \sum_{k=1}^{n} w_{j,n,k} e^{-s_{j,n,k}^3} \right| \leq (N - M) \epsilon_2 I.
\]

This lemma holds since \( G_2(n) \to 0 \) as \( n \to \infty \), see Figure 2.

**FIG. 2.** Plot of functions \( G_2(\cdot) \) and \( G_4(\cdot) \).
Theorem 2. If $\frac{T}{\Delta t} \leq 2^{30}$ and $\epsilon > 10^{-222}$, let $M = \left\lfloor -\frac{1}{3} \log_2 T \right\rfloor$ and $N = \left\lceil \frac{1}{3} \log_2 \frac{-\ln \epsilon}{\Delta t} \right\rceil$, then we have

$$\left| I - \sum_{k=1}^{n^*} w^{*}_{M,n^*,k} e^{-\left(\tau^*_{M,n^*,k}\right)^3} - \sum_{j=M}^{N-1} \sum_{k=1}^{n} w_{j,n,k} e^{-\tau_{j,n,k}^3} \right| \leq 16\epsilon I.$$

Here, $n^* = n^*(\epsilon)$ and $n = n(\epsilon)$ are determined as in Lemma 2 and Lemma 3.

Proof. By Lemma 1, Lemma 2, and Lemma 3, we have

$$\left| I - \sum_{k=1}^{n^*} w^{*}_{M,n^*,k} e^{-\left(\tau^*_{M,n^*,k}\right)^3} - \sum_{j=M}^{N-1} \sum_{k=1}^{n} w_{j,n,k} e^{-\tau_{j,n,k}^3} \right| \leq (N - M + 2)\epsilon I.$$

The theorem follows since

$$N - M < \frac{1}{3} \log_2 T + \frac{1}{3} \log_2 \frac{-\ln \epsilon}{\Delta t} + 2 = \frac{1}{3} \log_2 \frac{T}{\Delta t} + \frac{1}{3} \log_2 (-\ln \epsilon) + 2 \leq 15.$$

In the case when $\alpha = \frac{1}{3}$ in (4.1), we use the following integral equality

$$\frac{1}{3} \Gamma\left(\frac{2}{3}\right) t^{-\frac{2}{3}} = \int_0^{\infty} s e^{-s^{\frac{3}{2}}} ds = \int_0^{2^M} s e^{-s^{\frac{3}{2}}} ds + \int_{2^M}^{\infty} s e^{-s^{\frac{3}{2}}} ds = I_1^* + I_2^* + I_3^* \equiv I^*.$$

The analysis for $\alpha = \frac{2}{3}$ can be applied here analogously.

Theorem 3. If $\frac{T}{\Delta t} \leq 2^{30}$ and $\epsilon > 10^{-222}$, let $M = \left\lfloor -\frac{1}{3} \log_2 T \right\rfloor$ and $N = \left\lceil \frac{1}{3} \log_2 \frac{-\ln \epsilon}{\Delta t} \right\rceil$, we have

$$\left| I^* - \sum_{k=1}^{n^*} w^{*}_{M,n^*,k} e^{-\left(\tau^*_{M,n^*,k}\right)^3} - \sum_{j=M}^{N-1} \sum_{k=1}^{n} w_{j,n,k} e^{-\tau_{j,n,k}^3} \right| \leq 16\epsilon I^*,$$

where $n^* = n^*(\epsilon)$ and $n = n(\epsilon)$ are the smallest integers which satisfy $G_3(n^*) \leq \epsilon$ and $G_4(n) \leq \epsilon$ with $G_3(\cdot)$ and $G_4(\cdot)$ defined by

$$G_3(n) = \frac{3}{\Gamma\left(\frac{2}{3}\right)} \frac{1}{2n + 1} \frac{1}{((2n)!)^3} \max_{x \in \mathbb{R}^+} \left| D_{x}^{2n} xe^{-x^3} \right|,$$

$$G_4(n) = \frac{3}{\Gamma\left(\frac{2}{3}\right)} \frac{1}{2n + 1} \frac{1}{((2n)!)^3} \max_{x \in \mathbb{R}^+} x^{2n+1} \left| D_{x}^{2n} xe^{-x^3} \right|.$$
The number of exponentials needed to approximate the kernel functions $t^{-\frac{1}{3}}$ and $t^{-\frac{2}{3}}$ with a relative error $16\epsilon$.

<table>
<thead>
<tr>
<th>$\epsilon$</th>
<th>$2^9$</th>
<th>$2^{12}$</th>
<th>$2^{15}$</th>
<th>$2^{18}$</th>
<th>$2^{21}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^{-4}$</td>
<td>36</td>
<td>44</td>
<td>52</td>
<td>60</td>
<td>68</td>
</tr>
<tr>
<td>$10^{-7}$</td>
<td>60</td>
<td>73</td>
<td>86</td>
<td>99</td>
<td>102</td>
</tr>
<tr>
<td>$10^{-10}$</td>
<td>81</td>
<td>99</td>
<td>117</td>
<td>135</td>
<td>153</td>
</tr>
<tr>
<td>$10^{-13}$</td>
<td>103</td>
<td>126</td>
<td>149</td>
<td>172</td>
<td>195</td>
</tr>
</tbody>
</table>

The number of exponentials can be further reduced if $\Delta t$ and $T$ are fixed.

V. NUMERICAL TESTS

We consider the following problem

$$u_t + u_{xxx} = 0, \quad x \in \mathbb{R},$$

$$u(x,0) = e^{-x^2}, \quad x \in \mathbb{R},$$

$$u \to 0, \quad x \to \infty.$$  \hspace{1cm} (5.1) (5.2) (5.3)

The fundamental solution of the linearized KdV equation (5.1) is

$$E(x,t) = \frac{1}{\sqrt{3t}} \text{Ai}\left(\frac{x}{\sqrt{3t}}\right),$$

where $\text{Ai}(\cdot)$ is the Airy function (see [8] for details). The exact solution of problem (5.1)–(5.3) can be written out explicitly in terms of $E(x,t)$ as

$$u_{exa}(x,t) = E(x,t) \ast e^{-x^2},$$

where $\ast$ denotes the convolution operator on the whole real axis. Function $u_{exa}(x,4)$ is depicted in Figure 3.

Thanks to the fast decaying of Gaussian functions, the initial function given by (5.2) can be taken to have a compact support. In fact, in the outside of $[-6, 6]$, the function $e^{-x^2}$ equals zero with a tolerance less than $10^{-15}$. Thus then, being restricted to $[-6, 6]$, the exact solution $u_{exa}(x,t)$ is a good approximation of the exact solution to the following problem

$$u_t + u_{xxx} = 0, \quad x \in [-6, 6],$$

$$u(x,0) = e^{-x^2}, \quad x \in [-6, 6],$$

$$u(-6,t) - L^\frac{1}{3}_x u_x(-6,t) + L^\frac{2}{3}_x u_{xx}(-6,t) = 0,$$

$$u(6,t) - L^\frac{1}{3}_x u_x(6,t) = 0,$$

$$u_x(6,t) + L^\frac{1}{3}_x u_{xx}(6,t) = 0.$$
Now we use our schemes to compute the numerical solutions. For ease of reference, we will refer to the dual-Petrov-Galerkin scheme without fast evaluation as the standard scheme, and the scheme with fast evaluation as the fast scheme. The ending time point is set to be $T_f = 4$. For the fast scheme, the tolerance is $\epsilon = 10^{-16}$, and altogether 265 decaying exponentials are used to approximate the kernel functions. All computations are performed on a personal computer with a 1.70GHz CPU and 1GB bytes of memory.

We plot the numerical solution in Figure 3. It can be observed that even with a small number of spatial modes ($N = 32$) and a relative large time step ($\Delta t = \frac{1}{8}$), the numerical solution matches the exact solution very well. Next we study the accuracy of our fast scheme. The logarithms of $L^2$-errors are plotted in Figure 4. We set $N = 256$ in the top plot, which is made large enough so that error from spatial discretization is neglectable. We see that the $L^2$-errors degenerate with a rate of 4 if the time step is halved. This means that our fast scheme is of second order in time. In the bottom plot of Figure 4, the time step is set to be sufficiently small, $\Delta t = \frac{1}{32}$, and the number of modes is limited in the interval $[16, 52]$. In this context, the error from spatial discretization dominates, and the error from time discretization is neglectable. An almost straight line indicates that the $L^2$-error converges like $\exp(-cN^2)$, which is the typical supergeometric convergence for analytic functions by spectral methods (see [12]). The errors presented by the standard scheme are not shown here, since they are almost the same as those presented by the fast scheme if the computational parameters are set equal.

To compare the performance of the standard scheme and the fast scheme, we depict computation time in Figure 5 for different number of time steps. We observe that the fast scheme greatly reduce the computational effort. The computation cost for $\Delta t = \frac{1}{700000}$ is even less than that of the standard scheme for $\Delta t = \frac{1}{30000}$. We would also remark that no stability had been detected during our numerical implementation.

VI. CONCLUSION

In this paper, we have considered numerical simulation of a linearized KdV equation on the whole real axis. We have derived the exact nonreflecting boundary conditions. These boundary
conditions are local in space and nonlocal in time. It turns out that the exact nonreflecting boundary conditions are quite different at the left and right boundary points: at the right boundary point we need two boundary conditions, but only one at the left boundary point. Applying these boundary conditions gives a reduced problem defined on a finite interval, and a dual-Petrov-Galerkin method has been proposed for numerical approximation.

Our numerical scheme can be easily adapted for other settings of problem. For example, when the definition domain is a semi-infinite interval, the procedure of designing fast numerical scheme can be followed. The minor difference lies in the construction of trial and test function spaces. A more interesting problem which deserves future consideration is how to simulate the solution to the fully nonlinear KdV equation.

References


