

A NEW DUAL-PETROV–GALERKIN METHOD FOR THIRD AND HIGHER ODD-ORDER DIFFERENTIAL EQUATIONS: APPLICATION TO THE KDV EQUATION*

JIE SHEN†

Abstract. A new dual-Petrov–Galerkin method is proposed, analyzed, and implemented for third and higher odd-order equations using a spectral discretization. The key idea is to use trial functions satisfying the underlying boundary conditions of the differential equations and test functions satisfying the “dual” boundary conditions. The method leads to linear systems which are sparse for problems with constant coefficients and well conditioned for problems with variable coefficients. Our theoretical analysis and numerical results indicate that the proposed method is extremely accurate and efficient and most suitable for the study of complex dynamics of higher odd-order equations.

Key words. dual-Petrov–Galerkin, KDV equation, Legendre polynomials, spectral approximation

AMS subject classifications. 65M70, 35Q53, 65M12

DOI. 10.1137/S0036142902410271

1. Introduction. Over the last thirty years, spectral methods have been playing an increasingly important role in scientific and engineering computations. Most work on spectral methods is concerned with elliptic and parabolic-type equations; there has also been active research on spectral methods for hyperbolic problems (see, for instance, [11, 7, 14] and the references therein). However, there is only a limited body of literature on spectral methods for dispersive, namely, third and higher odd-order, equations. In particular, relatively few studies are devoted to third and higher odd-order equations in finite intervals. This is partly due to the fact that direct collocation methods for higher odd-order boundary problems lead to very much higher condition numbers—more precisely, of order N^{2k} , where N is the number of modes and k is the order of the equation—and often exhibit unstable modes if the collocation points are not properly chosen (see, for instance, [17, 21]).

In a sequence of papers [22, 23, 25, 26], the author constructed efficient spectral-Galerkin algorithms for elliptic equations in various situations. In this paper, we extend the main idea for constructing efficient spectral-Galerkin algorithms—using *compact combinations* of orthogonal polynomials, which satisfy *essentially* all the underlying homogeneous boundary conditions, as basis functions—to third and higher odd-order equations. Since the main differential operators in these equations are not symmetric, it is quite natural to employ a Petrov–Galerkin method.

The key idea of the new spectral dual-Petrov–Galerkin method is the innovative choice of the test and trial functional spaces. More precisely, we choose the trial functions to satisfy the underlying boundary conditions of the differential equations, and we choose the test functions to satisfy the “dual” boundary conditions.

Recently, Ma and Sun [19, 20] studied an interesting Legendre–Petrov–Galerkin method for third-order equations. The main difference between this paper and [19, 20]

*Received by the editors June 26, 2002; accepted for publication (in revised form) February 18, 2003; published electronically October 2, 2003. This research was partially supported by NFS grant DMS-0074283.

<http://www.siam.org/journals/sinum/41-5/41027.html>

†Department of Mathematics, Purdue University, West Lafayette, IN 47907 (shen@math.purdue.edu).

lies in the choice of the test and trial function spaces and their basis functions. The critical feature of test and trial spaces used here is that they allow us to integrate by parts freely without introducing any additional boundary terms. With this property and our choice of using “compact combinations” (minimal interactions) of Legendre polynomials as basis functions for the test and trial spaces, we obtain linear systems which are *compactly sparse* for problems with constant coefficients and *well conditioned* (i.e., the condition number is independent of the number of unknowns) for problems with variable coefficients. This is rather remarkable considering the fact that the problems at hand are nonsymmetric and involve high-order derivatives.

Together with the so-called Chebyshev–Legendre approach [10, 24], i.e., using the Legendre formulation and Chebyshev–Gauss–Lobatto points, our method has a quasi-optimal computational complexity and is well conditioned to permit the use of very large numbers of modes without suffering from large round-off errors, which are necessary for simulations of very complex dynamics of challenging scientific and engineering problems.

The new spectral dual-Petrov–Galerkin method not only leads to quasi-optimal numerical algorithms; it is also equivalent to a *natural* weighted variational formulation for third and higher odd-order equations. By showing that the basis functions for the trial (and test) spaces form a sequence of orthogonal polynomials in a weighted Sobolev space, we are able to establish optimal error estimates in appropriate weighted Sobolev spaces.

The paper is organized as follows. In sections 2 and 3, we study the third-order and fifth-order equations, respectively. As an example of application, we consider the Korteweg–de Vries (KDV) equation on a finite interval in section 4. In section 5, we discuss miscellaneous issues/extensions of the spectral dual-Petrov–Galerkin methods. In section 6, we present various numerical results exhibiting the accuracy and efficiency of our numerical algorithms. We end the paper with a few concluding remarks.

We now introduce some notation. Let $\omega(x)$ be a positive weight function on $I = (-1, 1)$. One usually requires that $\omega \in L^1(I)$. However, in this paper, we shall be interested mainly in the case $\omega \notin L^1(I)$. We shall use the weighted Sobolev spaces $H_\omega^m(\Omega)$ ($m = 0, \pm 1, \dots$) whose norms are denoted by $\|\cdot\|_{m,\omega}$. In particular, the norm and inner product of $L_\omega^2(\Omega) = L_\omega^2(\Omega)$ are denoted by $\|\cdot\|_\omega$ and $(\cdot, \cdot)_\omega$, respectively. To account for homogeneous boundary conditions, we define

$$H_{0,\omega}^m(\Omega) = \{v \in H_\omega^m(\Omega) : v(\pm 1) = v'(\pm 1) = \dots = v^{(m-1)}(\pm 1) = 0\}, \quad m = 1, 2, \dots$$

The subscript ω will be omitted from the notation in the case where $\omega \equiv 1$.

We denote by c a generic constant that is independent of any parameters and functions. In most cases, we shall simply use the expression $A \lesssim B$ to mean that there exists a generic constant c such that $A \leq cB$.

Let L_k be the k th degree Legendre polynomial. We now recall some basic properties of Legendre polynomials (cf. [27]) which will be used in this paper.

$$(1.1) \quad \int_{-1}^1 L_k(x)L_j(x)dx = \frac{2}{2k+1}\delta_{kj};$$

$$(1.2) \quad L_n(x) = \frac{1}{2n+1}(L'_{n+1}(x) - L'_{n-1}(x)), \quad n \geq 1;$$

$$(1.3a) \quad L'_n(x) = \sum_{\substack{k=0 \\ k+n \text{ odd}}}^{n-1} (2k+1)L_k(x);$$

$$(1.3b) \quad L''_n(x) = \sum_{\substack{k=0 \\ k+n \text{ even}}}^{n-2} \left(k + \frac{1}{2}\right) (n(n+1) - k(k+1)) L_k(x);$$

$$(1.4a) \quad L_n(\pm 1) = (\pm 1)^n,$$

$$(1.4b) \quad L'_n(\pm 1) = \frac{1}{2}(\pm 1)^{n-1}n(n+1),$$

$$(1.4c) \quad L''_n(\pm 1) = (\pm 1)^n(n-1)n(n+1)(n+2)/8.$$

2. Third-order equations.

2.1. Dual-Petrov-Galerkin method. Consider the model third-order equation

$$(2.1) \quad \begin{aligned} \alpha u - \beta u_x - \gamma u_{xx} + u_{xxx} &= f, \quad x \in I = (-1, 1), \\ u(\pm 1) = u_x(1) &= 0, \end{aligned}$$

where α, β, γ are given constants. Without loss of generality, we consider only homogeneous boundary conditions, for nonhomogeneous boundary conditions $u(-1) = c_1, u(1) = c_2,$ and $u_x(1) = c_3$ can be handled easily by considering $v = u - \hat{u},$ where \hat{u} is the unique quadratic polynomial satisfying the nonhomogeneous boundary conditions.

Denoting by P_N the space of polynomials of degree $\leq N,$ we set

$$(2.2) \quad V_N = \{u \in P_N : u(\pm 1) = u_x(1) = 0\}, \quad V_N^* = \{u \in P_N : u(\pm 1) = u_x(-1) = 0\}.$$

For any constants a and $b,$ let $\omega^{a,b}(x) = (1-x)^a(1+x)^b.$ We also define

$$(2.3) \quad V = \{u : u \in H_0^1(I), u_x \in L_{\omega^{-2,0}}^2(I)\}, \quad V^* = \{u : u \in H_0^1(I), u_x \in L_{\omega^{0,-2}}^2(I)\}.$$

It is clear that $V_N \subset V$ and $V_N^* \subset V^*.$

We consider the following Legendre dual-Petrov-Galerkin approximation for (2.1): Find $u_N \in V_N$ such that

$$(2.4) \quad \alpha(u_N, v_N) - \beta(\partial_x u_N, v_N) + \gamma(\partial_x u_N, \partial_x v_N) + (\partial_x u_N, \partial_x^2 v_N) = (f, v_N) \quad \forall v_N \in V_N^*,$$

where $(u, v) = \int_I uv dx, \partial_x u,$ and $\partial_x^2 u$ denote $\frac{du}{dx}$ and $\frac{d^2u}{dx^2},$ respectively.

Notice that for any $u_N \in V_N$ we have $\omega^{-1,1}u_N \in V_N^*.$ Thus the above dual-Petrov-Galerkin formulation is equivalent to the following weighted spectral-Galerkin approximation: Find $u_N \in V_N$ such that

$$(2.5) \quad \begin{aligned} \alpha(u_N, v_N)_{\omega^{-1,1}} - \beta(\partial_x u_N, v_N)_{\omega^{-1,1}} + \gamma(\partial_x u_N, \omega^{1,-1}\partial_x(v_N \omega^{-1,1}))_{\omega^{-1,1}} \\ + (\partial_x u_N, \omega^{1,-1}\partial_x^2(v_N \omega^{-1,1}))_{\omega^{-1,1}} = (f, v_N)_{\omega^{-1,1}} \quad \forall v_N \in V_N, \end{aligned}$$

where $(u, v)_{\omega^{-1,1}} = \int_I uv \omega^{-1,1} dx.$

We shall see that the dual-Petrov-Galerkin formulation (2.4) is most suitable for implementation, while the weighted Galerkin formulation (2.5) is more convenient for error analysis.

2.2. Basis functions and projection operators. As suggested in [22, 24], one should choose compact combinations of orthogonal polynomials as basis functions to minimize the bandwidth and the condition number of the coefficient matrix corresponding to (2.5). Let $\{p_k\}$ be a sequence of orthogonal polynomials. As a general rule, for one-dimensional differential equations with m boundary conditions, one should look for basis functions in the form

$$(2.6) \quad \phi_k(x) = p_k(x) + \sum_{j=1}^m a_j^{(k)} p_{k+j}(x),$$

where $a_j^{(k)}$ ($j = 1, \dots, m$) are chosen so that $\phi_k(x)$ satisfy the m homogeneous boundary conditions.

Using (1.4), one verifies readily that

$$(2.7) \quad \begin{aligned} \phi_k(x) &= L_k(x) - \frac{2k+3}{2k+5}L_{k+1}(x) - L_{k+2}(x) + \frac{2k+3}{2k+5}L_{k+3}(x) \in V_{k+3}, \\ \psi_k(x) &= L_k(x) + \frac{2k+3}{2k+5}L_{k+1}(x) - L_{k+2}(x) - \frac{2k+3}{2k+5}L_{k+3}(x) \in V_{k+3}^*. \end{aligned}$$

Therefore, for $N \geq 3$, we have

$$(2.8) \quad \begin{aligned} V_N &= \text{span}\{\phi_0, \phi_1, \dots, \phi_{N-3}\}; \\ V_N^* &= \text{span}\{\psi_0, \psi_1, \dots, \psi_{N-3}\}. \end{aligned}$$

Next, we discuss the properties of $\{\phi_k\}$ and $\{\psi_k\}$ and related projection operators in $L_{\omega^{-2,-1}}^2$ and $L_{\omega^{-1,-2}}^2$. Since the procedures for $L_{\omega^{-2,-1}}^2$ and $L_{\omega^{-1,-2}}^2$ are completely parallel, we shall describe only the results for $L_{\omega^{-2,-1}}^2$. One can obtain the corresponding results for $L_{\omega^{-1,-2}}^2$ by making a change of variable $x \rightarrow -x$.

LEMMA 2.1. *Let $\{\phi_k\}$ be defined as in (2.7). Then*

$$(2.9) \quad \int_I \phi_k \phi_j \omega^{-2,-1} dx = 0, \quad k \neq j,$$

and $\{\phi_k\}$ form a complete orthogonal basis in $L_{\omega^{-1,1}}^2$.

Furthermore, ϕ_k satisfies the following Sturm-Liouville equation:

$$(2.10) \quad A\phi_k := -(1-x)^2(1+x)\partial_x \{(1-x)^{-1}\partial_x \phi_k(x)\} = (k+1)(k+3)\phi_k(x).$$

Proof. By construction, $p_k(x) := \phi_k(x)\omega^{-2,-1}$ is a polynomial of degree $\leq k$. Thanks to the orthogonality of the Legendre polynomials,

$$\int_I \phi_k \phi_j \omega^{-2,-1} dx = \int_I p_k \phi_j dx = 0 \quad \forall k < j.$$

Hence $\{\phi_k\}$ is a sequence of orthogonal polynomials in $L_{\omega^{-2,-1}}^2$. One can verify that $\phi_k(x)$ is proportional to $(1-x)^2(1+x)J_k^{2,1}(x)$. Thus $\{\phi_k\}$ forms a complete orthogonal basis in $L_{\omega^{-2,-1}}^2$ since $\{J_k^{2,1}\}$ forms a complete orthogonal basis in $L_{\omega^{2,1}}^2$.

It is clear that $A\phi_k(x)$ is a polynomial of degree $\leq k+3$ and $\partial_x \{(1-x)^{-1}\partial_x \phi_k(x)\}$ is a polynomial of degree $\leq k$. Hence

$$\begin{aligned} \int_I A\phi_k(x)\phi_j(x)\omega^{-2,-1} dx &= - \int_I \partial_x \{(1-x)^{-1}\partial_x \phi_k(x)\}\phi_j dx \\ &= - \int_I \partial_x \{(1-x)^{-1}\partial_x \phi_j(x)\}\phi_k dx = 0 \quad \forall j < k. \end{aligned}$$

Therefore, $A\phi_k$ must be proportional to ϕ_k ; i.e., $A\phi_k = \lambda_k\phi_k$. By comparing the coefficients of x^{k+3} , we find that $\lambda_k = (k + 1)(k + 3)$. \square

Now, let π_N be the $L^2_{\omega^{-2,-1}}$ -orthogonal projector $L^2_{\omega^{-2,-1}} \rightarrow V_N$ defined by

$$(2.11) \quad (u - \pi_N u, v_N)_{\omega^{-2,-1}} = 0 \quad \forall v_N \in V_N.$$

We also define

$$(2.12) \quad B^m_{\omega^{-2,-1}}(I) = \{u \in L^2_{\omega^{-2,-1}}(I) : \partial_x^l u \in L^2_{\omega^{l-2,l-1}}(I), 1 \leq l \leq m\}.$$

Then, we have the following error estimates.

THEOREM 2.1.

$$\|\partial_x^l (u - \pi_N u)\|_{\omega^{l-2,l-1}} \lesssim N^{l-m} \|\partial_x^m u\|_{\omega^{m-2,m-1}} \quad \forall u \in B^m_{\omega^{-2,-1}}, 0 \leq l \leq m.$$

Proof. We recall that for $a, b > -1$, the Jacobi polynomials satisfy the following relations:

$$(2.13) \quad \partial_x^l J_k^{a,b}(x) = \kappa_{k,l}^{a,b} J_{k-l}^{a+l,b+l}(x), \quad a, b > -1, \quad k \geq l,$$

where

$$\kappa_{k,l}^{a,b} = \frac{\Gamma(k+l+a+b+1)}{2^l \Gamma(k+a+b+1)};$$

$$(2.14) \quad \int_I J_k^{a,b}(x) J_j^{a,b}(x) \omega^{a,b} dx = \gamma_k^{a,b} \delta_{kj},$$

where

$$\gamma_k^{a,b} = \frac{2^{a+b+1} \Gamma(k+a+1) \Gamma(k+b+1)}{(2k+a+b+1) \Gamma(k+1) \Gamma(k+a+b+1)}.$$

We shall extend the definition of the Jacobi polynomials to $(a, b) = (-2, -1)$ such that the relations (2.13)–(2.14) are still valid. To this end, we define

$$(2.15) \quad \begin{aligned} J_k^{-1,0}(x) &= -\frac{1}{2}(1-x)J_{k-1}^{1,0}(x), \quad k \geq 1, \\ J_k^{-2,-1}(x) &= \frac{1}{2}(k-2) \int_{-1}^x J_{k-1}^{-1,0}(t) dt, \quad k \geq 3. \end{aligned}$$

One derives immediately that $\{J_k^{-1,0}\}$ are mutually orthogonal in $L^2_{\omega^{-1,0}}$. Note that $\{L_{k-1} - L_k\}$ are also mutually orthogonal in $L^2_{\omega^{-1,0}}$. Hence $J_k^{-1,0}$ is proportional to $\{L_{k-1} - L_k\}$. One can also derive from the properties of Legendre polynomials that $J_k^{-2,-1}(\pm 1) = \partial_x J_k^{-2,-1}(1) = 0$. We then derive from (1.2) and (2.15) that $J_k^{-2,-1}$ must be proportional to ϕ_{k-3} . Hence $\{J_k^{-2,-1}\}$ are mutually orthogonal in $L^2_{\omega^{-2,-1}}$. Moreover, one can verify that $J_k^{-1,0}$ and $J_k^{-2,-1}$ satisfy the relations (2.13)–(2.14).

For any $u \in L^2_{\omega^{-2,-1}}$, we write

$$u(x) = \sum_{k=3}^{\infty} \tilde{u}_k J_k^{-2,-1}(x) \quad \text{with} \quad \tilde{u}_k = (u, J_k^{-2,-1})_{\omega^{-2,-1}} / \|J_k^{-2,-1}\|_{\omega^{-2,-1}}^2.$$

Hence $u - \pi_N u = \sum_{k=N+1}^{\infty} \tilde{u}_k J_k^{-2,-1}$. Let us define

$$(2.16) \quad C_{N,l,m} = \max_{k>N} \frac{(\kappa_{k,l}^{-2,-1})^2 \gamma_{k-l}^{l-2,l-1}}{(\kappa_{k,m}^{-2,-1})^2 \gamma_{k-m}^{m-2,m-1}}.$$

Then, by using (2.13)–(2.14), we find

$$(2.17) \quad \begin{aligned} \|\partial_x^l(u - \pi_N u)\|_{\omega^{l-2,l-1}}^2 &= \sum_{k=N+1}^{\infty} \tilde{u}_k^2 (\kappa_{k,l}^{-2,-1})^2 \|J_{k-l}^{l-2,l-1}\|_{\omega^{l-2,l-1}}^2 \\ &\leq C_{N,l,m} \sum_{k=N+1}^{\infty} \tilde{u}_k^2 (\kappa_{k,m}^{-2,-1})^2 \|J_{k-m}^{m-2,m-1}\|_{\omega^{m-2,m-1}}^2 \\ &\leq C_{N,l,m} \|\partial_x^m u\|_{\omega^{m-2,m-1}}^2. \end{aligned}$$

The desired results follow from the above inequality and the fact that

$$C_{N,l,m} \lesssim N^{2(l-m)}. \quad \square$$

2.3. Error estimates. Let us first prove the following generalized Poincaré inequalities.

LEMMA 2.2.

$$(2.18) \quad \begin{aligned} \int_I \frac{u^2}{(1-x)^4} dx &\leq \frac{4}{9} \int_I \frac{u_x^2}{(1-x)^2} dx \quad \forall u \in V_N, \\ \int_I \frac{u^2}{(1-x)^3} dx &\leq \int_I \frac{u_x^2}{1-x} dx \quad \forall u \in V_N. \end{aligned}$$

Proof. Let $u \in V_N$ and $h \leq 2$. Then, for any constant q , we have

$$\begin{aligned} 0 &\leq \int_I \left(\frac{u}{1-x} + qu_x \right)^2 \frac{1}{(1-x)^h} dx \\ &= \int_I \left(\frac{u^2}{(1-x)^{2+h}} + q \frac{(u^2)_x}{(1-x)^{1+h}} + q^2 \frac{u_x^2}{(1-x)^h} \right) dx \\ &= (1 - (1+h)q) \int_I \frac{u^2}{(1-x)^{2+h}} dx + q^2 \int_I \frac{u_x^2}{(1-x)^h} dx. \end{aligned}$$

We obtain the first inequality by taking $h = 2$ and $q = \frac{2}{3}$ and the second inequality with $h = 1$ and $q = 1$. \square

Remark 2.1. We note that with a change of variable $x \rightarrow -x$ in the above lemma, we have corresponding inequalities for $u \in V_N^*$.

LEMMA 2.3.

$$(2.19) \quad \frac{1}{3} \|u_x\|_{\omega^{-2,0}}^2 \leq (u_x, (u\omega^{-1,1})_{xx}) \leq 3 \|u_x\|_{\omega^{-2,0}}^2 \quad \forall u \in V_N.$$

Proof. For any $u \in V_N$, we have $u\omega^{-1,1} \in V_N^*$. Thanks to the homogeneous boundary conditions built into the spaces V_N and V_N^* , all the boundary terms from the integration by parts of the third-order term would vanish. Therefore, using the

identity $\partial_x^k \omega^{-1,1}(x) = \frac{2k!}{(1-x)^{k+1}}$ and Lemma 2.2, we find

$$\begin{aligned} (u_x, (u\omega^{-1,1})_{xx}) &= (u_x, u_{xx}\omega^{-1,1} + 2u_x\omega_x^{-1,1} + u\omega_{xx}^{-1,1}) \\ &= \frac{1}{2} \int_I ((u_x^2)_x \omega^{-1,1} + (u^2)_x \omega_{xx}^{-1,1} + 4u_x^2 \omega_x^{-1,1}) dx \\ &= \int_I \left(\frac{3}{2} u_x^2 \omega_x^{-1,1} - \frac{1}{2} u^2 \omega_{xxx}^{-1,1} \right) dx \\ &= 3 \int_I \frac{u_x^2}{(1-x)^2} dx - 6 \int_I \frac{u^2}{(1-x)^4} dx \geq \frac{1}{3} \int_I \frac{u_x^2}{(1-x)^2} dx. \end{aligned}$$

The desired results follow immediately from the above. \square

Before we proceed with the error estimates, we make the following simple but important observation.

LEMMA 2.4. *Let π_N be defined in (2.11). Then*

$$(\partial_x(u - \pi_N u), \partial_x^2 v_N) = 0 \quad \forall u \in V, v_N \in V_N^*.$$

Proof. The result is a direct consequence of (2.11), the identity

$$(\partial_x(u - \pi_N u), \partial_x^2 v_N) = -(u - \pi_N u, \omega^{2,1} \partial_x^3 v_N)_{\omega^{-2,-1}},$$

and the fact that $\omega^{2,1} \partial_x^3 v_N \in V_N$. \square

Let us denote $\hat{e}_N = \pi_N u - u_N$ and $e_N = u - u_N = (u - \pi_N u) + \hat{e}_N$.

THEOREM 2.2. *For any $\alpha, \beta \geq 0$ and $-\frac{1}{3} < \gamma < \frac{1}{6}$, there exists a unique solution for the system (2.4). Furthermore, for $u \in B_{\omega^{-2,-1}}^m$, we have*

$$\alpha \|e_N\|_{\omega^{-1,1}} + N^{-1} \|(e_N)_x\|_{\omega^{-1,0}} \lesssim (1 + |\gamma|N) N^{-m} \|\partial_x^m u\|_{\omega^{m-2,m-1}}, \quad m \geq 1.$$

Proof. We derive from (2.1), (2.5), and Lemma 2.4 that

$$\begin{aligned} (2.20) \quad \alpha(e_N, v_N)_{\omega^{-1,1}} - \beta(\partial_x e_N, v_N)_{\omega^{-1,1}} + \gamma(\partial_x e_N, \omega^{1,-1} \partial_x(v_N \omega^{-1,1}))_{\omega^{-1,1}} \\ + (\partial_x \hat{e}_N, \omega^{1,-1} \partial_x^2(v_N \omega^{-1,1}))_{\omega^{-1,1}} = 0 \quad \forall v_N \in V_N. \end{aligned}$$

Taking $v_N = \hat{e}_N$ in the above and using Lemma 2.3 and the identities

$$\begin{aligned} (2.21) \quad - (v_x, v)_{\omega^{-1,1}} &= -\frac{1}{2} \int_I (v^2)_x \omega^{-1,1} dx = \|v\|_{\omega^{-2,0}}^2 \quad \forall v \in V_N, \\ (v_x, (v\omega^{-1,1})_x) &= (v_x, v_x \omega^{-1,1} + 2v\omega^{-2,0}) = \|v_x\|_{\omega^{-1,1}}^2 - 2\|v\|_{\omega^{-3,0}}^2 \quad \forall v \in V_N, \end{aligned}$$

we obtain

$$\begin{aligned} \alpha \|\hat{e}_N\|_{\omega^{-1,1}}^2 + \beta \|\hat{e}_N\|_{\omega^{-2,0}}^2 + \gamma \|(\hat{e}_N)_x\|_{\omega^{-1,1}}^2 - 2\gamma \|\hat{e}_N\|_{\omega^{-3,0}}^2 + \frac{1}{3} \|(\hat{e}_N)_x\|_{\omega^{-2,0}}^2 \\ \leq -\alpha(u - \pi_N u, \hat{e}_N)_{\omega^{-1,1}} + \beta(\partial_x(u - \pi_N u), \hat{e}_N)_{\omega^{-1,1}} \\ - \gamma(\partial_x(u - \pi_N u), \partial_x(\hat{e}_N \omega^{-1,1})). \end{aligned}$$

The right-hand side can be bounded by using Lemma 2.2, the Cauchy-Schwarz inequality, and the fact that $\omega^{-1,2} \leq 2\omega^{-1,1} \leq 2\omega^{-2,0}$:

$$\begin{aligned} (u - \pi_N u, \hat{e}_N)_{\omega^{-1,1}} &\leq \|\hat{e}_N\|_{\omega^{-1,1}} \|u - \pi_N u\|_{\omega^{-1,1}} \leq 2\|\hat{e}_N\|_{\omega^{-1,1}} \|u - \pi_N u\|_{\omega^{-2,-1}}, \\ ((u - \pi_N u)_x, \hat{e}_N)_{\omega^{-1,1}} &= (u - \pi_N u, \partial_x \hat{e}_N \omega^{-1,1} + 2\hat{e}_N \omega^{-2,0}) \\ &\lesssim \|u - \pi_N u\|_{\omega^{-2,-1}} \|\partial_x \hat{e}_N\|_{\omega^{-2,0}}, \\ ((u - \pi_N u)_x, (\hat{e}_N \omega^{-1,1})_x) &= ((u - \pi_N u)_x, (\hat{e}_N)_x \omega^{-1,1} + 2\hat{e}_N \omega^{-2,0}) \\ &\leq \|(u - \pi_N u)_x\|_{\omega^{-1,0}} \|(\hat{e}_N)_x\|_{\omega^{-2,0}}. \end{aligned}$$

For $0 \leq \gamma < \frac{1}{6}$, we choose δ sufficiently small such that $\frac{1}{3} - 2\gamma - \delta > 0$. Combining the above inequalities, using the inequality

$$(2.22) \quad ab \leq \epsilon a^2 + \frac{1}{4\epsilon} b^2 \quad \forall \epsilon > 0,$$

and dropping some unnecessary terms, we get

$$\begin{aligned} \frac{\alpha}{2} \|\hat{e}_N\|_{\omega^{-1,1}}^2 + \left(\frac{1}{3} - 2\gamma - \delta\right) \|(\hat{e}_N)_x\|_{\omega^{-2,0}}^2 \\ \lesssim \|u - \pi_N u\|_{\omega^{-2,-1}}^2 + \gamma \|(u - \pi_N u)_x\|_{\omega^{-1,0}}^2 \\ \lesssim (1 + \gamma N^2) N^{-2m} \|\partial_x^m u\|_{\omega^{m-2,m-1}}. \end{aligned}$$

The last inequality follows from Theorem 2.1.

For $-\frac{1}{3} < \gamma < 0$, we choose δ sufficiently small such that $\frac{1}{3} + \gamma - \delta > 0$, and we derive similarly

$$\frac{\alpha}{2} \|\hat{e}_N\|_{\omega^{-1,1}}^2 + \left(\frac{1}{3} + \gamma - \delta\right) \|(\hat{e}_N)_x\|_{\omega^{-2,0}}^2 \lesssim (1 + |\gamma| N^2) N^{-2m} \|\partial_x^m u\|_{\omega^{m-2,m-1}}.$$

The desired results follow from the triangular inequality, Theorem 2.1, and the fact that $\|u\|_{\omega^{-1,0}} \leq 2\|u\|_{\omega^{-1,0}}$. \square

Remark 2.2. Note that the error estimate in the above theorem is optimal for $\gamma = 0$ but suboptimal for $\gamma \neq 0$.

2.4. Linear system and its coefficient matrices. Hence, by setting

$$(2.23) \quad \begin{aligned} u_N &= \sum_{k=0}^{N-3} \tilde{u}_k \phi_k, \quad \bar{u} = (\tilde{u}_0, \tilde{u}_1, \dots, \tilde{u}_{N-3})^t, \\ \tilde{f}_k &= (f, \psi_k), \quad \bar{f} = (\tilde{f}_0, \tilde{f}_1, \dots, \tilde{f}_{N-3})^t, \\ m_{ij} &= (\phi_j, \psi_i), \quad p_{ij} = -(\phi'_j, \psi_i), \quad q_{ij} = (\phi'_j, \psi'_i), \quad s_{ij} = (\phi'_j, \psi''_i), \end{aligned}$$

the linear system (2.4) becomes

$$(2.24) \quad (\alpha M + \beta P + \gamma Q + S)\bar{u} = \bar{f},$$

where M, P, Q , and S are $(N - 2) \times (N - 2)$ matrices with entries m_{ij}, p_{ij}, q_{ij} , and s_{ij} , respectively.

Thanks to the orthogonality of the Legendre polynomials, we have $m_{ij} = 0$ for $|i - j| > 3$. Therefore, M is a seven-diagonal matrix. We note that the homogeneous “dual” boundary conditions satisfied by ϕ_j and ψ_i allow us to integrate by parts freely without introducing additional boundary terms; namely, we have

$$s_{ij} = (\phi'_j, \psi''_i) = (\phi'''_j, \psi_i) = -(\phi_j, \psi'''_i).$$

Thanks to the compact form of ϕ_j and ψ_i , we have $s_{ij} = 0$ for $i \neq j$. So S is a diagonal matrix. Similarly, we see that P is a pentadiagonal matrix and Q is a tridiagonal matrix. It is an easy matter to derive that

$$(2.25) \quad s_{ii} = 2(2i + 3)^2.$$

Nonzero elements of M, P, Q can be easily determined from the properties of Legendre polynomials. Hence the linear system (2.24), under the condition of Theorem 2.2, can be easily formed and inverted.

3. Fifth-order equations. In this section, we shall consider an example of fifth-order equations. We shall follow essentially the same procedures as in the previous section and will omit some repetitive details.

3.1. Dual-Petrov-Galerkin method. Consider the model fifth-order equation:

$$(3.1) \quad \begin{aligned} \alpha u + \beta u_{xxx} - u_{xxxxx} &= f, \quad x \in I = (-1, 1), \\ u(\pm 1) = u_x(\pm 1) &= u_{xx}(1) = 0, \end{aligned}$$

where α and β are given constants. For the sake of simplicity and with the fifth-order KDV equation in mind, we included only zeroth- and third-order linear terms in the equation. Other linear terms as well as nonhomogeneous boundary conditions can be treated as in the previous section.

Similarly to the third-order equation, we define

$$(3.2) \quad \begin{aligned} W_N &= \{u \in P_N : u(\pm 1) = u_x(\pm 1) = u_{xx}(1) = 0\}, \\ W_N^* &= \{u \in P_N : u(\pm 1) = u_x(\pm 1) = u_{xx}(-1) = 0\}. \end{aligned}$$

We also define

$$(3.3) \quad \begin{aligned} W &= \{u : u \in H_0^2(I), u_{xx} \in L_{\omega^{-2,0}}^2\}, \\ W^* &= \{u : u \in H_0^2(I), u_{xx} \in L_{\omega^{0,-2}}^2\}. \end{aligned}$$

It is clear that $W_N \subset W$ and $W_N^* \subset W^*$.

We consider the following Legendre dual-Petrov-Galerkin approximation for (3.1): Find $u_N \in W_N$ such that

$$(3.4) \quad \alpha(u_N, v_N) - \beta(\partial_x^2 u_N, \partial_x v_N) + (\partial_x^2 u_N, \partial_x^3 v_N) = (f, v_N) \quad \forall v_N \in W_N^*.$$

Once again, the above dual-Petrov-Galerkin formulation is equivalent to the following weighted spectral-Galerkin approximation: Find $u_N \in W_N$ such that

$$(3.5) \quad \begin{aligned} \alpha(u_N, v_N)_{\omega^{-1,1}} - \beta(\partial_x^2 u_N, \omega^{1,-1} \partial_x (v_N \omega^{-1,1}))_{\omega^{-1,1}} \\ + (\partial_x^2 u_N, \omega^{1,-1} \partial_x^3 (v_N \omega^{-1,1}))_{\omega^{-1,1}} = (f, v_N)_{\omega^{-1,1}} \quad \forall v_N \in W_N. \end{aligned}$$

Note in particular that $W \subset V$ and $W^* \subset V^*$. Hence the results proved in the previous section are still valid here.

3.2. Basis functions and projection operators. The construction of suitable basis functions for W_N and W_N^* follows the general principle (2.6); i.e., we look for

$$(3.6) \quad \Phi_k = L_k + a_1^{(k)} L_{k+1} + a_2^{(k)} L_{k+2} + a_3^{(k)} L_{k+3} + a_4^{(k)} L_{k+4} + a_5^{(k)} L_{k+5}$$

such that $\Phi_k \in W$. Using (1.4) and after some simplifications, we find that $\{a_j^{(k)}\}$ satisfy the following relations:

$$(3.7) \quad \begin{aligned} a_2^{(k)} + a_4^{(k)} &= -1, \\ (k+2)(k+3)a_2^{(k)} + (k+4)(k+5)a_4^{(k)} &= -k(k+1), \end{aligned}$$

and

$$(3.8) \quad \begin{aligned} a_1^{(k)} + a_3^{(k)} + a_5^{(k)} &= 0, \\ (k+1)(k+2)a_1^{(k)} + (k+3)(k+4)a_3^{(k)} + (k+5)(k+6)a_5^{(k)} &= 0, \\ (k+1)^2(k+2)^2 a_1^{(k)} + (k+3)^2(k+4)^2 a_3^{(k)} + (k+5)^2(k+6)^2 a_5^{(k)} &= g_k, \end{aligned}$$

where $g_k = -k^2(k+1)^2 - (k+2)^2(k+3)^2 a_2^{(k)} - (k+4)^2(k+5)^2 a_4^{(k)}$.

One derives immediately from (3.7) that

$$(3.9) \quad a_2^{(k)} = -\frac{2(2k+5)}{2k+7}, \quad a_4^{(k)} = \frac{2k+3}{2k+7}.$$

We can then determine $a_1^{(k)}, a_3^{(k)}, a_5^{(k)}$ by solving the 3×3 system (3.8).

It is easy to verify that

$$(3.10) \quad \Psi_k = L_k - a_1^{(k)}L_{k+1} + a_2^{(k)}L_{k+2} - a_3^{(k)}L_{k+3} + a_4^{(k)}L_{k+4} - a_5^{(k)}L_{k+5} \in W^*.$$

Hence, for $N \geq 5$, we have

$$(3.11) \quad \begin{aligned} W_N &= \text{span}\{\Phi_0, \Phi_1, \dots, \Phi_{N-5}\}; \\ W_N^* &= \text{span}\{\Psi_0, \Psi_1, \dots, \Psi_{N-5}\}. \end{aligned}$$

We define

$$(3.12) \quad L_{\omega^{-3,-2}}^2(I) = L_{\omega^{-3,-2}}^2(I) \cap W, \quad L_{\omega^{-2,-3}}^2(I) = L_{\omega^{-2,-3}}^2(I) \cap W^*$$

equipped with the norm of $\|\cdot\|_{\omega^{-3,-2}}$ and $\|\cdot\|_{\omega^{-2,-3}}$, respectively. We shall only summarize the results for $L_{\omega^{-3,-2}}^2$ below. The corresponding results for $L_{\omega^{-2,-3}}^2$ are obtained by using the transform $x \rightarrow -x$. The proofs are essentially the same as in the previous section.

LEMMA 3.1. *Let $\{\Phi_k\}$ be defined as in (3.6). Then*

$$(3.13) \quad \int_I \Phi_k \Phi_j \omega^{-3,-2} dx = 0, \quad k \neq j,$$

and $\{\Phi_k\}$ forms a complete orthogonal basis in $L_{\omega^{-3,-2}}^2$.

Furthermore, Φ_k satisfies the following Sturm–Liouville equation:

$$(3.14) \quad B\Phi_k := -(1-x)^3(1+x)^2 \partial_x \left\{ (1-x)^{-2}(1+x)^{-1} \partial_x \Phi_k(x) \right\} = (k+1)(k+5)\Phi_k(x).$$

Now, let Π_N be the $L_{\omega^{-3,-2}}^2$ -orthogonal projector $L_{\omega^{-3,-2}}^2 \rightarrow W_N$ defined by

$$(3.15) \quad (u - \Pi_N u, v_N)_{\omega^{-3,-2}} = 0 \quad \forall v_N \in W_N.$$

We define

$$(3.16) \quad H_{\omega^{-3,-2}}^m(I) = \{u \in L_{\omega^{-3,-2}}^2(I) : \partial_x^l u \in L_{\omega^{l-3,l-2}}^2(I), 1 \leq l \leq m\}.$$

THEOREM 3.1.

$$\|\partial_x^l (u - \Pi_N u)\|_{\omega^{l-3,l-2}} \lesssim N^{l-m} \|\partial_x^m u\|_{\omega^{m-3,m-2}} \quad \forall u \in H_{\omega^{-3,-2}}^m, \quad 0 \leq l \leq m.$$

Proof. Let us define

$$(3.17) \quad J_k^{-3,-2}(x) = \frac{1}{2}(k-4) \int_{-1}^x J_{k-1}^{-2,-1}(t) dt, \quad k \geq 5.$$

One can show that $J_k^{-3,-2}(\pm 1) = \partial_x J_k^{-3,-2}(\pm 1) = \partial_x^2 J_k^{-3,-2}(1) = 0$. Since $J_k^{-2,-1}$ is proportional to ϕ_{k-3} , we then derive from (1.2) and (3.17) that $J_k^{-3,-2}$ must be proportional to Φ_{k-5} so $\{J_k^{-3,-2}\}$ are mutually orthogonal in $L_{\omega^{-3,-2}}^2$. Moreover, one can verify that $J_k^{-3,-2}$ satisfies the relations (2.13)–(2.14). Thus the desired results follow from the same arguments as those in Theorem 2.1. \square

3.3. Error estimates. We first prove the following generalized Poincaré inequalities.

LEMMA 3.2.

$$(3.18) \quad \int_I \frac{u^2}{(1-x)^6} dx \leq \frac{4}{25} \int_I \frac{u_x^2}{(1-x)^4} dx \leq \frac{16}{225} \int_I \frac{u_{xx}^2}{(1-x)^2} dx \quad \forall u \in W_N,$$

and

$$(3.19) \quad \int_I \frac{u_x^2}{(1-x)^4} dx \leq \frac{1}{7} \int_I \frac{u_{xx}^2}{(1-x)^2} dx + 2 \int_I \frac{u^2}{(1-x)^6} dx \quad \forall u \in W_N.$$

Proof. The proof is similar to that of Lemma 2.2. Letting $u \in W_N$, for any constant q ,

$$\begin{aligned} 0 \leq \int_I \left(\frac{u}{1-x} + qu_x \right)^2 \frac{1}{(1-x)^4} dx &= \int_I \left(\frac{u^2}{(1-x)^6} + q \frac{(u^2)_x}{(1-x)^5} + q^2 \frac{u_x^2}{(1-x)^4} \right) dx \\ &= (1-5q) \int_I \frac{u^2}{(1-x)^6} dx + q^2 \int_I \frac{u_x^2}{(1-x)^4} dx. \end{aligned}$$

We obtain the first part of (3.18) by taking $q = \frac{2}{5}$. The second part is a direct consequence of Lemma 2.2 since $u_x \in V_N$.

For (3.19), we consider the following relation with any constants q and r :

$$\begin{aligned} 0 \leq \int_I \left(\frac{u}{(1-x)^2} + q \frac{u_x}{1-x} + ru_{xx} \right)^2 \frac{1}{(1-x)^2} dx \\ = (1-5q+20r) \int_I \frac{u^2}{(1-x)^6} dx + (q^2-2r-4qr) \int_I \frac{u_x^2}{(1-x)^4} dx + r^2 \int_I \frac{u_{xx}^2}{(1-x)^2} dx. \end{aligned}$$

We obtain (3.19) by taking $q = \frac{3}{2}$ and $r = \frac{1}{2}$. \square

LEMMA 3.3.

$$(3.20) \quad \frac{5}{7} \int_I \frac{u_{xx}^2}{(1-x)^2} dx \leq (\partial_x^2 u, \partial_x^3(u\omega^{-1,1})) \leq \frac{203}{15} \int_I \frac{u_{xx}^2}{(1-x)^2} dx \quad \forall u \in W_N.$$

Proof. For any $u \in W_N$, we set $u = \Phi(1-x)$ with $\Phi(\pm 1) = \Phi_x(\pm 1) = 0$. Then, by integrating by parts and using the fact that all boundary terms are zero, we find

$$\begin{aligned} (\partial_x^2 u, \partial_x^3(u\omega^{-1,1})) &= -(\partial_x^3 u, \partial_x^2(u\omega^{-1,1})) = -(\Phi_{xxx}(1-x) - 3\Phi_{xx}, \Phi_{xx}(1+x) + 2\Phi_x) \\ &= \int_I \left\{ -\frac{1}{2}(\Phi_{xx}^2)_x(1-x^2) + 3\Phi_{xx}^2(1+x) + 3(\Phi_x^2)_x + 2\Phi_{xx}(\Phi_x(1-x))_x \right\} dx \\ &= 5 \int_I \Phi_{xx}^2 dx = 5 \int_I \left\{ \partial_x^2 \left(\frac{u}{1-x} \right) \right\}^2 dx. \end{aligned}$$

Expanding $\partial_x^2(\frac{u}{1-x})$ and integrating by parts, we get

$$(3.21) \quad \int_I \left\{ \partial_x^2 \left(\frac{u}{1-x} \right) \right\}^2 dx = \int_I \frac{u_{xx}^2}{(1-x)^2} dx - 6 \int_I \frac{u_x^2}{(1-x)^4} dx + 24 \int_I \frac{u^2}{(1-x)^6} dx.$$

We conclude by applying (3.18) and (3.19) to the above. \square

Let Π_N be defined in (3.15). By definition, we have

$$(3.22) \quad (\partial_x^2(u - \Pi_N u), \partial_x^3 v_N) = (u - \Pi_N u, \omega^{3,2} \partial_x^5 v_N)_{\omega^{-3,-2}} = 0 \quad \forall u \in W, v_N \in W_N^*.$$

Letting u and u_N be, respectively, the solution of (3.1) and (3.4), we denote $\hat{e}_N = \Pi_N u - u_N$ and $e_N = u - u_N = (u - \Pi_N u) + \hat{e}_N$.

THEOREM 3.2. *For any $\alpha, \beta \geq 0$, there exists a unique solution for the system (3.4). Furthermore, for $u \in H_{\omega^{-2,-1}}^m$, we have*

$$\begin{aligned} \alpha \|e_N\|_{\omega^{-1,1}} + \beta N^{-1} \|(e_N)_x\|_{\omega^{-2,0}} + N^{-2} \|(e_N)_{xx}\|_{\omega^{-1,0}} \\ \lesssim (1 + \beta N) N^{-m} \|\partial_x^m u\|_{\omega^{m-3,m-2}}, \quad m \geq 2. \end{aligned}$$

Proof. We derive from (3.1), (3.5), and (3.22) that

$$\alpha (e_N, v_N)_{\omega^{-1,1}} - \beta (\partial_x^2 e_N, \partial_x (v_N \omega^{-1,1})) + (\partial_x^2 \hat{e}_N, \partial_x^3 (v_N \omega^{-1,1})) = 0 \quad \forall v_N \in W_N.$$

Taking $v_N = \hat{e}_N$ in the above and using Lemmas 2.3 and 3.3, we obtain

$$\begin{aligned} \alpha \|\hat{e}_N\|_{\omega^{-1,1}}^2 + \frac{\beta}{9} \|(\hat{e}_N)_x\|_{\omega^{-2,0}}^2 + \frac{5}{7} \|(\hat{e}_N)_{xx}\|_{\omega^{-2,0}}^2 \\ \leq -\alpha (u - \Pi_N u, \hat{e}_N)_{\omega^{-1,1}} + \beta (\partial_x^2 (u - \Pi_N u), \partial_x (\hat{e}_N \omega^{-1,1})) \\ = -\alpha (u - \Pi_N u, \hat{e}_N)_{\omega^{-1,1}} \\ - \beta (\partial_x (u - \Pi_N u), (\partial_x^2 \hat{e}_N \omega^{-1,1} + 4 \partial_x \hat{e}_N \omega^{-2,0} + 4 \hat{e}_N \omega^{-3,0})). \end{aligned}$$

Using the Cauchy–Schwarz inequality, we bound the right-hand side as follows:

$$(u - \Pi_N u, \hat{e}_N)_{\omega^{-1,1}} \leq \|\hat{e}_N\|_{\omega^{-1,1}} \|u - \Pi_N u\|_{\omega^{-1,1}} \lesssim \|\hat{e}_N\|_{\omega^{-1,1}} \|u - \Pi_N u\|_{\omega^{-3,-2}};$$

and thanks to Lemma 2.2,

$$\begin{aligned} ((u - \Pi_N u)_x, (\partial_x^2 \hat{e}_N \omega^{-1,1} + 4 \partial_x \hat{e}_N \omega^{-2,0} + 4 \hat{e}_N \omega^{-3,0})) \\ \lesssim \|(u - \Pi_N u)_x\|_{\omega^{-2,-1}} (\|\partial_x^2 \hat{e}_N\|_{\omega^{0,3}} + \|\partial_x \hat{e}_N\|_{\omega^{-2,1}} + \|\hat{e}_N\|_{\omega^{-4,1}}) \\ \lesssim \|(u - \Pi_N u)_x\|_{\omega^{-2,-1}} (\|\partial_x^2 \hat{e}_N\|_{\omega^{-2,0}} + \|\partial_x \hat{e}_N\|_{\omega^{-2,0}}). \end{aligned}$$

Using (2.22) and combining the above inequalities, we arrive at

$$\begin{aligned} \frac{\alpha}{2} \|\hat{e}_N\|_{\omega^{-1,1}}^2 + \frac{\beta}{18} \|\partial_x \hat{e}_N\|_{\omega^{-2,0}}^2 + \frac{5}{14} \|\partial_x^2 \hat{e}_N\|_{\omega^{-2,0}}^2 \\ \lesssim \|u - \Pi_N u\|_{\omega^{-3,-2}}^2 + \beta \|(u - \Pi_N u)_x\|_{\omega^{-2,-1}}^2 \\ \lesssim (1 + \beta N^2) N^{-2m} \|\partial_x^m u\|_{\omega^{m-3,m-2}}. \end{aligned}$$

The last inequality follows from Theorem 3.1.

We can then conclude from the above inequality and the triangular inequality. \square

Remark 3.1. Similarly as in the previous section, the error estimate in the above theorem is optimal for $\beta = 0$ but suboptimal for $\beta \neq 0$.

3.4. Linear system and its coefficient matrices. Similarly to the third-order equation, we set

$$(3.23) \quad \begin{aligned} u_N &= \sum_{k=0}^{N-5} \tilde{u}_k \Phi_k, \quad \tilde{u} = (\tilde{u}_0, \tilde{u}_1, \dots, \tilde{u}_{N-5})^t, \\ \tilde{f}_k &= (f, \Psi_k), \quad \tilde{u} = (\tilde{f}_0, \tilde{f}_1, \dots, \tilde{f}_{N-5})^t, \\ m_{ij} &= (\Phi_j, \Psi_i), \quad M = (m_{ij})_{i,j=0,1,\dots,N-5}, \\ p_{ij} &= -(\Phi_j'', \Psi_i'), \quad P = (p_{ij})_{i,j=0,1,\dots,N-5}, \\ s_{ij} &= (\Phi_j'', \Psi_i'''), \quad S = (s_{ij})_{i,j=0,1,\dots,N-5} \end{aligned}$$

so that the linear system (3.4) becomes

$$(3.24) \quad (\alpha M + \beta P + S)\bar{u} = \bar{f}.$$

Obviously, M is an eleven nonzero diagonal matrix, and with integration by parts, we find that P is a pentadiagonal matrix and S is a diagonal matrix. Repeatedly using (1.2), we can derive that

$$(3.25) \quad s_{ii} = 2(2i + 3)(2i + 5)(2i + 7)(2i + 9)a_5^{(i)}.$$

Nonzero elements of M and P can be determined accordingly using (1.2) and (1.3). Hence the linear system (3.24), under the condition of Theorem 3.1, can also be easily inverted.

4. Application to the KDV equation. There is a vast body of literature on various aspects of the KDV equation. Although most of these studies are concerned with initial value problems, the initial-boundary value problems also received considerable attention. The most natural initial-boundary value KDV equation is set in a quarter-plane (see, for instance, [28, 16, 3, 2, 12, 13, 5] and the references therein). The KDV equation on a finite spatial interval has also been considered by several authors [20, 9, 4]. Here, as an example of application to nonlinear equations, we consider the KDV equation on a finite interval:

$$(4.1) \quad \begin{aligned} \alpha v_t + \beta v_x + \gamma v v_x + v_{xxx} &= 0, \quad x \in (-1, 1), \quad t \in (0, T], \\ v(-1, t) = g(t), \quad v(1, t) = v_x(1, t) &= 0, \quad t \in [0, T], \\ v(x, 0) = v_0(x), \quad x &\in (-1, 1). \end{aligned}$$

The positive constants α , β , and γ are introduced to accommodate the scaling of the spatial interval. The existence and uniqueness of the solution for (4.1) can be established as in [9, 4]. Beside its own interests, (4.1) can also be viewed as a legitimate approximate model for the KDV equation on a quarter-plane before the wave reaches the right boundary.

Let us first reformulate (4.1) as an equivalent problem with homogeneous boundary conditions. To this end, let $\hat{v}(x, t) = \frac{(1-x)^2}{4}g(t)$, and write $v(x, t) = u(x, t) + \hat{v}(x, t)$. Then u satisfies the following equation with homogeneous boundary conditions:

$$(4.2) \quad \begin{aligned} \alpha u_t + a(x, t)u + b(x, t)u_x + \gamma u u_x + u_{xxx} &= f, \quad x \in (-1, 1), \quad t \in (0, T], \\ u(\pm 1, t) = u_x(1, t) = 0; \quad t &\in [0, T], \\ u(x, 0) = u_0(x) = v_0(x) - \hat{v}(x, 0), \quad x &\in (-1, 1), \end{aligned}$$

where $a(x, t) = \frac{\gamma}{2}(x - 1)g(t)$, $b(x, t) = \beta + \gamma\hat{v}(x, t)$, and $f(x, t) = -\alpha\hat{v}_t(x, t)$.

For a given Δt , we set $t_k = k\Delta t$ and let $u_N^0 = \pi_N u_0$ and $u_N^1 \in V_N$ be an appropriate approximation of $u(\cdot, t_1)$, for instance; we can compute u_N^1 using one step of a semi-implicit first-order scheme so that for $u \in C^3(0, T; L^2_{\omega^{2,2}}(I)) \cap C(0, T; B^m_{\omega^{-2,-1}})$, we have

$$(4.3) \quad \|u_N^1 - \pi_N u(\cdot, t_1)\|_{\omega^{-1,1}} \lesssim \Delta t^2 + N^{-m}.$$

Let M be such that $|u(x, t)| \leq M$ for $x \in [-1, 1]$ and $t \in [0, T]$. We define a cut-off function

$$(4.4) \quad H(x) = \begin{cases} x, & |x| \leq 2M, \\ 2M, & x > 2M, \\ -2M, & x < -2M. \end{cases}$$

It is easy to verify that

$$(4.5) \quad |H(x) - H(y)| \leq |x - y| \quad \forall x, y.$$

We consider first a modified Crank–Nicolson leap-frog dual-Petrov–Galerkin approximation:

$$(4.6) \quad \begin{aligned} & \frac{\alpha}{2\Delta t} (u_N^{k+1} - u_N^{k-1}, v_N)_{\omega^{-1,1}} + \frac{1}{2} (\partial_x (u_N^{k+1} + u_N^{k-1}), \partial_x^2 (v_N \omega^{-1,1})) \\ &= (f(\cdot, t_k), v_N)_{\omega^{-1,1}} + \frac{\gamma}{2} (H(u_N^k) u_N^k, \partial_x (v_N \omega^{-1,1})) \\ & \quad - (a u_N^k, v_N)_{\omega^{-1,1}} + (u_N^k, \partial_x (b v_N \omega^{-1,1})) \quad \forall v_N \in V_N. \end{aligned}$$

We denote $\hat{e}_N^k = \pi_N u(\cdot, t_k) - u_N^k$, $\tilde{e}_N^k = u(\cdot, t_k) - \pi_N u(\cdot, t_k)$, and $e_N^k = u(\cdot, t_k) - u_N^k$.

THEOREM 4.1. *We assume $u \in C^3(0, T; L^2_{\omega^{2,2}}(I)) \cap C(0, T; B^m_{\omega^{-2,-1}})$ with $m > 1$. Then the scheme (4.6) is unconditionally stable, and the following error estimates hold:*

$$\begin{aligned} & \|e_N^{n+1}\|_{\omega^{-1,1}} \lesssim \Delta t^2 + N^{-m}, \quad 0 \leq n \leq [T/\Delta t] - 1, \\ & \left(\Delta t \sum_{k=1}^n \|\partial_x (e_N^{k+1} + e_N^{k-1})\|_{\omega^{-1,0}}^2 \right)^{\frac{1}{2}} \lesssim \Delta t^2 + N^{1-m}, \quad 1 \leq n \leq [T/\Delta t] - 1. \end{aligned}$$

Proof. Let E^k ($k = 1, 2, \dots$) be the truncation error defined by

$$(4.7) \quad \begin{aligned} & \frac{\alpha}{2\Delta t} (u(\cdot, t_{k+1}) - u(\cdot, t_{k-1})) + a(\cdot, t_k) u(\cdot, t_k) + b(\cdot, t_k) u_x(\cdot, t_k) + \gamma u(\cdot, t_k) \partial_x u(\cdot, t_k) \\ & \quad + \frac{1}{2} \partial_x^3 (u(\cdot, t_{k+1}) + u(\cdot, t_{k-1})) - f(\cdot, t_k) = E^k(\cdot). \end{aligned}$$

Comparing (4.6) with (4.7) and using Lemma 2.4, we have

$$(4.8) \quad \begin{aligned} & \frac{\alpha}{2\Delta t} (\hat{e}_N^{k+1} - \hat{e}_N^{k-1}, v_N)_{\omega^{-1,1}} + \frac{1}{2} (\partial_x (\hat{e}_N^{k+1} + \hat{e}_N^{k-1}), \partial_x^2 (v_N \omega^{-1,1})) \\ &= \frac{\gamma}{2} (u(\cdot, t_k)^2 - H(u_N^k) u_N^k, \partial_x (v_N \omega^{-1,1})) \\ & \quad - (a (\hat{e}_N^k + \tilde{e}_N^k), v_N)_{\omega^{-1,1}} + \frac{1}{2} ((\hat{e}_N^k + \tilde{e}_N^k), \partial_x (b v_N \omega^{-1,1})) \\ & \quad + (E^k, v_N)_{\omega^{-1,1}} - \frac{\alpha}{2\Delta t} (\tilde{e}_N^{k+1} - \tilde{e}_N^{k-1}, v_N)_{\omega^{-1,1}} \quad \forall v_N \in V_N. \end{aligned}$$

Let $A = \max_{x \in [-1,1], t \in [0,T]} |a(x, t)|$ and $B = \max_{x \in [-1,1], t \in [0,T]} (|b(x, t)| + |\partial_x b(x, t)|)$.

We now take $v_N = 2\Delta t(\hat{e}_N^{k+1} + \hat{e}_N^{k-1})$ in (4.8) and bound the right-hand side terms using repeatedly the Cauchy–Schwarz inequality and Lemma 2.2 as follows:

$$-2\Delta t (a (\hat{e}_N^k + \tilde{e}_N^k), \hat{e}_N^{k+1} + \hat{e}_N^{k-1})_{\omega^{-1,1}} \leq \Delta t A (\|\tilde{e}_N^k\|_{\omega^{-1,1}}^2 + \|\hat{e}_N^k\|_{\omega^{-1,1}}^2 + \|\hat{e}_N^{k+1} + \hat{e}_N^{k-1}\|_{\omega^{-1,1}}^2),$$

$$\begin{aligned} 2\Delta t (E^k, \hat{e}_N^{k+1} + \hat{e}_N^{k-1})_{\omega^{-1,1}} & \leq 2\Delta t \|E^k\|_{\omega^{2,2}} \|\hat{e}_N^{k+1} + \hat{e}_N^{k-1}\|_{\omega^{-4,0}} \\ & \leq c\Delta t \|E^k\|_{\omega^{2,2}}^2 + \frac{\Delta t}{36} \|\partial_x (\hat{e}_N^{k+1} + \hat{e}_N^{k-1})\|_{\omega^{-2,0}}^2. \end{aligned}$$

Let us denote $\tilde{E}_N^k = \frac{1}{2\Delta t} (\tilde{e}_N^{k+1} - \tilde{e}_N^{k-1})$. Similarly as above, we have

$$-\frac{\alpha}{2\Delta t} (\tilde{e}_N^{k+1} - \tilde{e}_N^{k-1}, \hat{e}_N^{k+1} + \hat{e}_N^{k-1})_{\omega^{-1,1}} \leq c\Delta t \|\tilde{E}_N^k\|_{\omega^{2,2}}^2 + \frac{\Delta t}{36} \|\partial_x (\hat{e}_N^{k+1} + \hat{e}_N^{k-1})\|_{\omega^{-2,0}}^2.$$

By the assumption $|u(x, t)| \leq M$ for all x and t , we have

$$|u(\cdot, t_k)^2 - H(u_N^k)u_N^k| = |u(\cdot, t_k)(H(u(\cdot, t_k)) - H(u_N^k)) + H(u_N^k)(u(\cdot, t_k) - u_N^k)| \leq 2M|u(\cdot, t_k) - u_N^k| \leq 2M(|\tilde{e}_N^k| + |\hat{e}_N^k|).$$

Hence,

$$\begin{aligned} & \gamma \Delta t (u(\cdot, t_k)^2 - H(u_N^k)u_N^k, \partial_x(\hat{e}_N^{k+1} + \hat{e}_N^{k-1})\omega^{-1,1}) \\ & \leq \gamma \Delta t \|u(\cdot, t_k)^2 - H(u_N^k)u_N^k\|_{\omega^{0,2}} \|\partial_x(\hat{e}_N^{k+1} + \hat{e}_N^{k-1})\|_{\omega^{-2,0}} \\ & \leq c \Delta t (\|\tilde{e}_N^k\|_{\omega^{-2,-1}}^2 + \|\hat{e}_N^k\|_{\omega^{-1,1}}^2) + \frac{\Delta t}{36} \|\partial_x(\hat{e}_N^{k+1} + \hat{e}_N^{k-1})\|_{\omega^{-2,0}}^2, \end{aligned}$$

To handle the remaining terms, we recall the following Hardy inequalities:

$$(4.9) \quad \begin{aligned} & \int_a^b \left[\frac{1}{t-a} \int_a^t \psi(s) ds \right]^2 (t-a)^\alpha dt \leq \frac{4}{1-\alpha} \int_a^b \psi^2(t) (t-a)^\alpha dt, \\ & \int_a^b \left[\frac{1}{b-t} \int_t^b \psi(s) ds \right]^2 (b-t)^\alpha dt \leq \frac{4}{1-\alpha} \int_a^b \psi^2(t) (b-t)^\alpha dt, \end{aligned}$$

which hold for all measurable functions ϕ on (a, b) with $a < b$ and $\alpha < -1$.

Let $(a, b) = (-1, 0)$, $\alpha = 0$, and $\phi(t) = \int_{-1}^t \psi(s) ds$. We find

$$\int_{-1}^0 \frac{1}{(1-t)^3(1+t)} \phi^2(t) dt \leq \int_{-1}^0 \frac{1}{(t+1)^2} \phi^2(t) dt \leq 4 \int_{-1}^0 (\phi_t)^2 dt.$$

Let $(a, b) = (0, 1)$, $\alpha = -1$, and $\phi(t) = \int_t^1 \psi(s) ds$. We find

$$\int_0^1 \frac{1}{(1-t)^3(1+t)} \phi^2(t) dt \leq \int_0^1 \frac{1}{(1-t)^3} \phi^2(t) dt \leq 2 \int_0^1 (\phi_t)^2 \frac{1}{1-t} dt.$$

Combining the above two inequalities, we obtain

$$(4.10) \quad \int_I \phi^2 \omega^{-3,-1} dx \leq 4 \int_I (\phi_x)^2 \omega^{-1,0} dx,$$

which holds for all ϕ such that $\phi(\pm 1) = 0$ and $\int_I (\phi_x)^2 \omega^{-1,0} dx < \infty$.

Thanks to Lemma 2.2 and (4.10),

$$\begin{aligned} & \gamma \Delta t (u(\cdot, t_k)^2 - H(u_N^k)u_N^k, (\hat{e}_N^{k+1} + \hat{e}_N^{k-1}) \partial_x \omega^{-1,1}) \\ & \leq 2\gamma \Delta t \|u(\cdot, t_k)^2 - H(u_N^k)u_N^k\|_{\omega^{-1,1}} \|(\hat{e}_N^{k+1} + \hat{e}_N^{k-1})\|_{\omega^{-3,-1}} \\ & \leq c \Delta t (\|\tilde{e}_N^k\|_{\omega^{-2,-1}}^2 + \|\hat{e}_N^k\|_{\omega^{-1,1}}^2) + \frac{\Delta t}{36} \|\partial_x(\hat{e}_N^{k+1} + \hat{e}_N^{k-1})\|_{\omega^{-2,0}}^2. \end{aligned}$$

Similarly, we have

$$\begin{aligned} & \Delta t ((\hat{e}_N^k + \tilde{e}_N^k), \partial_x(b(\hat{e}_N^{k+1} + \hat{e}_N^{k-1})\omega^{-1,1})) = \Delta t ((\hat{e}_N^k + \tilde{e}_N^k), b \partial_x(\hat{e}_N^{k+1} + \hat{e}_N^{k-1})\omega^{-1,1}) \\ & \quad + \Delta t ((\hat{e}_N^k + \tilde{e}_N^k), b(\hat{e}_N^{k+1} + \hat{e}_N^{k-1}) \partial_x \omega^{-1,1}) + \Delta t (\hat{e}_N^k + \tilde{e}_N^k, \partial_x b(\hat{e}_N^{k+1} + \hat{e}_N^{k-1})\omega^{-1,1}) \\ & \leq cB \Delta t (\|\tilde{e}_N^k\|_{\omega^{-2,-1}}^2 + \|\hat{e}_N^k\|_{\omega^{-1,1}}^2) + \frac{\Delta t}{36} \|\partial_x(\hat{e}_N^{k+1} + \hat{e}_N^{k-1})\|_{\omega^{-2,0}}^2. \end{aligned}$$

Combining the above inequalities into (4.8) and using Lemma 2.3, we obtain

$$\begin{aligned} & \alpha (\|\hat{e}_N^{k+1}\|_{\omega^{-1,1}}^2 - \|\hat{e}_N^{k-1}\|_{\omega^{-1,1}}^2) + \frac{\Delta t}{36} \|\partial_x(\hat{e}_N^{k+1} + \hat{e}_N^{k-1})\|_{\omega^{-2,0}}^2 \\ & \leq c \Delta t (\|E^k\|_{\omega^{2,2}}^2 + \|\tilde{E}_N^k\|_{\omega^{2,2}}^2 + \|\hat{e}_N^{k+1}\|_{\omega^{-1,1}}^2 + \|\hat{e}_N^k\|_{\omega^{-1,1}}^2 + \|\hat{e}_N^{k-1}\|_{\omega^{-1,1}}^2 + \|\tilde{e}_N^k\|_{\omega^{-2,-1}}^2). \end{aligned}$$

We can then apply the standard discrete Gronwall lemma to the above inequality to

get

$$\begin{aligned} \|\hat{e}_N^{n+1}\|_{\omega^{-1,1}}^2 + \frac{\Delta t}{36} \sum_{k=1}^n \|\partial_x(\hat{e}_N^{k+1} + \hat{e}_N^{k-1})\|_{\omega^{-2,0}}^2 &\lesssim \|\hat{e}_N^0\|_{\omega^{-1,1}}^2 + \|\hat{e}_N^1\|_{\omega^{-1,1}}^2 \\ &+ \Delta t \sum_{k=1}^n (\|\tilde{e}_N^k\|_{\omega^{-2,-1}}^2 + \|E^k\|_{\omega^{2,2}}^2 + \|\tilde{E}_N^k\|_{\omega^{2,2}}^2), \quad 1 \leq n \leq [T/\Delta t] - 1. \end{aligned}$$

Using the triangular inequality, we derive that

$$\begin{aligned} \|e_N^{n+1}\|_{\omega^{-1,1}}^2 &\lesssim \|\hat{e}_N^0\|_{\omega^{-1,1}}^2 + \|\hat{e}_N^1\|_{\omega^{-1,1}}^2 + \|\tilde{e}_N^{n+1}\|_{\omega^{-1,1}}^2 \\ &+ \Delta t \sum_{k=1}^n (\|\tilde{e}_N^k\|_{\omega^{-2,-1}}^2 + \|E^k\|_{\omega^{2,2}}^2 + \|\tilde{E}_N^k\|_{\omega^{2,2}}^2) \end{aligned}$$

and

$$\begin{aligned} \frac{\Delta t}{36} \sum_{k=1}^n \|\partial_x(e_N^{k+1} + e_N^{k-1})\|_{\omega^{-1,0}}^2 &\lesssim \|\hat{e}_N^0\|_{\omega^{-1,1}}^2 + \|\hat{e}_N^1\|_{\omega^{-1,1}}^2 \\ &+ \Delta t \sum_{k=1}^n (\|\tilde{e}_N^k\|_{\omega^{-2,-1}}^2 + \|E^k\|_{\omega^{2,2}}^2 + \|\tilde{E}_N^k\|_{\omega^{2,2}}^2 + \|\partial_x(\tilde{e}_N^{k+1} + \tilde{e}_N^{k-1})\|_{\omega^{-1,0}}^2). \end{aligned}$$

We can then conclude from the assumptions and Theorem 2.1. \square

Next, we consider the standard Crank–Nicolson leap-frog weighted Galerkin approximation:

$$\begin{aligned} (4.11) \quad &\frac{\alpha}{2\Delta t} (u_N^{k+1} - u_N^{k-1}, v_N)_{\omega^{-1,1}} + \frac{1}{2} (\partial_x(u_N^{k+1} + u_N^{k-1}), \partial_x^2(v_N \omega^{-1,1})) \\ &= -(a u_N^k, v_N)_{\omega^{-1,1}} + (u_N^k, \partial_x(bv_N \omega^{-1,1})) \\ &+ (f(\cdot, t_k), v_N)_{\omega^{-1,1}} + \frac{\gamma}{2} ((u_N^k)^2, \partial_x v_N \omega^{-1,1} + v_N \partial_x \omega^{-1,1}) \quad \forall v_N \in V_N. \end{aligned}$$

COROLLARY 4.1. *Under the conditions of Theorem 4.1, there exists c_0 such that for $\Delta t N \leq c_0$, the two schemes (4.6) and (4.11) are equivalent.*

Proof. We need only to show that the scheme (4.6) reduces to (4.11) under the condition that $\Delta t N \leq c_0$. Indeed, using the estimate in Theorem 4.1, the inverse inequality $\|u\|_{L^\infty} \lesssim N^2 \|u\|_{\omega^{0,1}}$ for all $u \in P_N$ (see Lemma 4.1 below), and the assumptions on u , we find that there exists c_0 such that for $\Delta t N \leq c_0$ we have

$$\begin{aligned} \|u_N^k\|_{L^\infty} &\leq \|u(\cdot, t_k)\|_{L^\infty} + \|\tilde{e}_N^k\|_{L^\infty} + \|\hat{e}_N^k\|_{L^\infty} \\ &\leq M + \|\tilde{e}_N^k\|_{L^\infty} + N^2 \|\hat{e}_N^k\|_{\omega^{0,1}} \\ &\leq M + \|\tilde{e}_N^k\|_{L^\infty} + cN^2 (\Delta t^2 + N^{-m}) \leq 2M. \end{aligned}$$

Hence (4.6) and (4.11) are equivalent. \square

The following lemma is a special case of Theorem 2.1 in [15]. For the reader's convenience, we provide an elementary proof below.

LEMMA 4.1.

$$\|u\|_{L^\infty} \lesssim N^2 \|u\|_{\omega^{0,1}} \quad \forall u \in P_N.$$

Proof. Let $J_k^{0,1}$ be the k th degree Jacobi polynomial of index $(0, 1)$. We recall that (cf. [27])

$$(4.12) \quad \|J_k^{0,1}\|_{L^\infty} = k + 1, \quad \|J_k^{0,1}\|_{\omega^{0,1}} = \sqrt{\frac{2}{k + 1}}.$$

For any $u \in P_N$, we write $u(x) = \sum_{k=0}^N u_k J_k^{0,1}(x)$. Then

$$\begin{aligned} \|u\|_{L^\infty}^2 &\leq (N + 1)^2 \left(\sum_{k=0}^N |u_k| \right)^2 \leq (N + 1)^3 \sum_{k=0}^N |u_k|^2 \\ &\leq \frac{1}{2} (N + 1)^4 \sum_{k=0}^N |u_k|^2 \frac{2}{k + 1} = \frac{1}{2} (N + 1)^4 \|u\|_{\omega^{0,1}}^2. \quad \square \end{aligned}$$

Remark 4.1. In practice, the nonlinear term is usually computed using the pseudospectral approach (cf. [11, 7]), which is discussed in the next section. It is not difficult to show that the results in Theorem 4.1 apply to the pseudospectral scheme (see [20] for a similar result).

Remark 4.2. We note that the result obtained here for the third-order KDV equation (4.2) can be extended to the fifth-order KDV equation, which has also attracted considerable attention (see, for instance, [18]).

5. Miscellaneous issues. We discuss in this section several extensions and practical issues related to the dual-Petrov-Galerkin method. We note, in particular, that the dual-Petrov-Galerkin method can be used with any spatial discretization method based on a variational formulation such as the finite-element method.

5.1. Other higher odd-order equations. We have discussed in detail the Legendre dual-Petrov-Galerkin method for third- and fifth-order equations. It is evident that the method can be directly applied to other higher odd-order equations of the form

$$(5.1) \quad \sum_{j=0}^{2m} a_j u^{(j)}(x) + u^{(2m+1)}(x) = 0, \quad m \geq 3,$$

with the boundary conditions

$$(5.2) \quad u(\pm 1) = u'(\pm 1) = \dots = u^{(m-1)}(\pm 1) = 0, \quad u^{(m)}(1) = 0.$$

Other boundary conditions and/or variable coefficients can be treated following the discussion below.

5.2. Other boundary conditions. It must be noted that our dual-Petrov-Galerkin approach is quite flexible and can be used for other unconventional boundary conditions. For instance, Colin and Ghidaglia [9] studied the KDV equation

$$(5.3) \quad u_t + \frac{2}{L}(u_x + uu_x) + \frac{8}{L^3}u_{xxx} = 0, \quad x \in (-1, 1), \quad t > 0,$$

with the boundary conditions

$$(5.4) \quad u(-1) = g(t), \quad u_x(1) = u_{xx}(1) = 0.$$

Note that we have scaled the interval from $(0, L)$, which was used in [9], to $(-1, 1)$.
Let us denote

$$(5.5) \quad X_N = \{u \in P_N : u(-1) = 0, u_x(1) = u_{xx}(1) = 0\}.$$

Then the “dual” space is

$$(5.6) \quad X_N^* = \{v \in P_N : v_x(-1) = v(-1) = 0, v_{xx}(1) = 0\}.$$

There exist unique coefficients $\{a_j^{(k)}, \tilde{a}_j^{(k)}\}$ such that

$$(5.7) \quad \begin{aligned} \phi_k &= L_k + \sum_{j=1}^3 a_j^{(k)} L_{k+j} \in X_N, \quad k = 0, 1, \dots, N-3, \\ \psi_k &= L_k + \sum_{j=1}^3 \tilde{a}_j^{(k)} L_{k+j} \in X_N^*, \quad k = 0, 1, \dots, N-3. \end{aligned}$$

Then, the Legendre dual-Petrov–Galerkin method for (5.3)–(5.4) is to find $u_N = v_N + \frac{(1-x)^3}{8}g(t)$ with $v_N \in X_N$ such that

$$(5.8) \quad (\partial_t u_N, \psi_j) + \frac{2}{L}(\partial_x u_N + u_N \partial_x u_N, \psi_j) + (\partial_x u_N, \partial_x^2 \psi_j) = 0, \quad j = 0, 1, \dots, N-3.$$

One can prove results which are similar to Theorem 4.1 for this problem.

5.3. Problems with variable coefficients: Pseudospectral method in modal basis. Let us consider, as an example, the following third-order equation:

$$(5.9) \quad \begin{aligned} a(x)u - b(x)u_x + u_{xxx} &= f, \quad x \in I = (-1, 1), \\ u(\pm 1) &= u_x(1) = 0. \end{aligned}$$

The pseudospectral dual-Petrov–Galerkin method for (5.9) is to find $u_N \in V_N$ such that

$$(5.10) \quad (a(x)u_N, v_N)_N - (b(x)u'_N, v_N)_N + (u'_N, v''_N)_N = (f, v_N)_N \quad \forall v_N \in V_N^*,$$

where

$$(5.11) \quad (u, v)_N = \sum_{k=0}^N u(x_k)v(x_k)\omega_k$$

is the discrete inner product of u and v associated with the Legendre–Gauss–Lobatto quadrature (cf. [7]). We recall that

$$(5.12) \quad (u, v)_N = (u, v) \quad \forall uv \in P_{2N-1}.$$

Let us denote $\tilde{\psi}_i = \frac{1}{(\phi'_i, \psi''_i)}\psi_i = \frac{1}{2(2i+3)^2}\psi_i$. Then we have

$$(\phi'_j, \tilde{\psi}''_i)_N = (\phi'_j, \psi''_i) = \delta_{ij}, \quad 0 \leq i, j \leq N-3.$$

Hence, by setting

$$\begin{aligned}
 (5.13) \quad u_N &= \sum_{k=0}^{N-3} \tilde{u}_k \phi_k, \quad \bar{u} = (\tilde{u}_0, \tilde{u}_1, \dots, \tilde{u}_{N-3})^t, \\
 \tilde{f}_k &= (f, \tilde{\psi}_k)_N, \quad \bar{f} = (\tilde{f}_0, \tilde{f}_1, \dots, \tilde{f}_{N-3})^t, \\
 m_{ij} &= (a(x)\phi_j, \tilde{\psi}_i)_N, \quad p_{ij} = -(b(x)\phi'_j, \tilde{\psi}_i)_N,
 \end{aligned}$$

the linear system (5.10) becomes

$$(5.14) \quad (M + P + I)\bar{u} = \bar{f}.$$

It is clear that the matrices M and P are full and their formation involves N^3 operations as well as the inversion of (5.14). Hence a direct approach is advisable only if one uses a small or moderate number of modes. Otherwise, an iterative method can be efficiently implemented as follows:

- Note that a conjugate gradient type iterative method does not require the explicit formation of the matrix; only the action of the matrix upon a given vector is needed at each iteration. Although the formation of M and P involves N^3 operations, their action on a given vector \bar{u} , i.e., $M\bar{u}$ and $P\bar{u}$, can be computed in $O(N^2)$ operations.
- The number of operations can be further reduced to a quasi-optimal $O(N \log N)$ if we use the following Chebyshev–Legendre dual-Petrov–Galerkin method (cf. [10, 24]): Find $u_N \in V_N$ such that

$$(I_N^c(a(x)u_N), v_N)_N - (I_N^c(b(x)u'_N), v_N)_N + (u'_N, v''_N)_N = (f, v_N)_N \quad \forall v_N \in V_N^*,$$

where I_N^c is the interpolation operator based on the Chebyshev–Gauss–Lobatto points, while $(\cdot, \cdot)_N$ is still the discrete inner product of u and v associated with the Legendre–Gauss–Lobatto quadrature. Hence the only difference between (5.3) and (5.10) is that $a(x)u_N$ and $b(x)u'_N$ in (5.10) are replaced by $I_N^c(a(x)u_N)$ and $I_N^c(b(x)u'_N)$. Thanks to the fast Fourier transform (FFT) and the fast Chebyshev–Legendre transform [1, 24], the Legendre coefficients of $I_N^c(a(x)u_N)$ and $I_N^c(b(x)u'_N)$ can be computed in $O(N \log N)$ operations given the Legendre coefficients of u_N (see [24] for details).

- Under reasonable assumptions on $a(x)$ and $b(x)$, the matrix $M + P + I$ is well conditioned; i.e, its condition number is independent of N . We now provide a heuristic argument for this statement.

Since $\phi_k \omega^{-1,1} \in V^*$, there exists unique $\{h_{kj}\}$ such that

$$\phi_k \omega^{-1,1} = \sum_{j=0}^{N-3} h_{kj} \psi_j, \quad k = 0, 1, \dots, N - 3.$$

Hence we have

$$\begin{aligned}
 \langle H\bar{u}, \bar{u} \rangle &= \left(\sum_{j=0}^{N-3} \tilde{u}_j \phi'_j, \sum_{k,j=0}^{N-3} \tilde{u}_j h_{kj} \psi''_j \right)_N \\
 &= \left(\sum_{j=0}^{N-3} \tilde{u}_j \phi'_j, \sum_{j=0}^{N-3} \tilde{u}_j (\phi_j \omega^{-1,1})'' \right) = (\partial_x u_N, \partial_x^2 (u_N \omega^{-1,1}))
 \end{aligned}$$

and

$$\begin{aligned} \langle HM\bar{u}, \bar{u} \rangle &= \left(a(x) \sum_{j=0}^{N-3} \tilde{u}_j \phi_j, \sum_{k,j=0}^{N-3} \tilde{u}_j h_{kj} \psi_j \right)_N \\ &= \left(a(x) \sum_{j=0}^{N-3} \tilde{u}_j \phi_j, \sum_{j=0}^{N-3} \tilde{u}_j \phi_j \omega^{-1,1} \right)_N = (a(x)u_N, u_N \omega^{-1,1})_N, \end{aligned}$$

where $\langle \bar{v}, \bar{v} \rangle := \sum_{j=0}^{N-3} v_j^2$ is the inner product in l^2 .

Let us assume that $0 \leq a(x) \leq a_1$. Then, thanks to (2.2) and (2.3), we derive from the above that there exists a constant a_2 independent of N such that

$$\langle H\bar{u}, \bar{u} \rangle \leq \langle H(M+I)\bar{u}, \bar{u} \rangle \leq a_2 \langle H\bar{u}, \bar{u} \rangle.$$

Hence the condition number of $M+I$, in the norm $\|\bar{v}\|_H := \langle H\bar{v}, \bar{v} \rangle^{\frac{1}{2}}$, is independent of N . Under assumptions similar to those in Theorem 2.2, one can also establish that the condition number of $M+P+I$ is independent of N . This statement is confirmed by our numerical results (see the next section).

Therefore, a conjugate gradient type iterative method like BICGSTAB or CGS for (5.14) will converge in a small and fixed number (i.e., independent of N) of steps. In short, the Chebyshev–Legendre dual-Petrov–Galerkin method for (5.9) can be solved in a quasi-optimal $O(N \log N)$ operation.

Since the unknowns are coefficients of the spectral expansion, instead of the nodal values of the approximate solution at the collocation points, we refer to the above as the pseudospectral dual-Petrov–Galerkin method in *modal* basis. The modal basis presents at least three distinct advantages compared with the nodal basis:

- As demonstrated in sections 2 and 3 (see also [22, 23, 24]), for problems with constant coefficients, using an appropriate modal basis leads to sparse matrices.
- With the nodal basis, the choice of quadrature rules/collocation points plays an important role and should be made in accordance with the underlying differential equations and boundary conditions (see, for instance, [17] and the references therein). For example, the Gauss–Lobatto points are not suitable for third-order equations (cf. [21]). With the modal basis, since the use of the quadrature rule is merely to approximate the integrals in the variational formulation, the choice of quadrature rules/collocation points is not important. Therefore, for the third-order equation (5.9), we can still use the usual Gauss–Lobatto quadrature.
- Most importantly, using an appropriate modal basis leads to well-conditioned matrices as we explained above.

6. Numerical results. We present in this section some numerical results illustrating the nice properties of our dual-Petrov–Galerkin method.

6.1. Third- and fifth-order linear equations. Let us first look at the conditioning of our dual-Petrov–Galerkin method. We list in Table 6.1 the condition numbers of $M+P+I$ in (5.14) with various $a(x)$ and $b(x)$. Notice that in all cases, the condition numbers are small and, more importantly, independent of N .

We list in Table 6.2 the condition numbers of $\alpha M + \beta P + S$ in (3.24) scaled by the diagonal matrix S with various α and β . Once again, the condition numbers are small and independent of N .

TABLE 6.1
Condition numbers of (5.14).

N	$a(x) = 1$ $b(x) = 0$	$a(x) = 10$ $b(x) = 0$	$a(x) = 50$ $b(x) = 0$	$a(x) = \sin x$ $b(x) = 2x - 1$	$a(x) = 10 \exp(x)$ $b(x) = \cos x$
16	1.119	2.218	7.219	1.188	2.393
64	1.119	2.218	7.219	1.188	2.393
128	1.119	2.218	7.219	1.188	2.393

TABLE 6.2
Condition numbers of (3.24) scaled by S .

N	$\alpha = 1$ $\beta = 0$	$\alpha = 100$ $\beta = 0$	$\alpha = 100$ $\beta = 100$	$\alpha = 1$ $\beta = -100$	$\alpha = 0$ $\beta = 1$
16	1.006	2.421	2.005	3.342	1.009
64	1.006	2.421	2.005	3.342	1.009
128	1.006	2.421	2.005	3.342	1.009

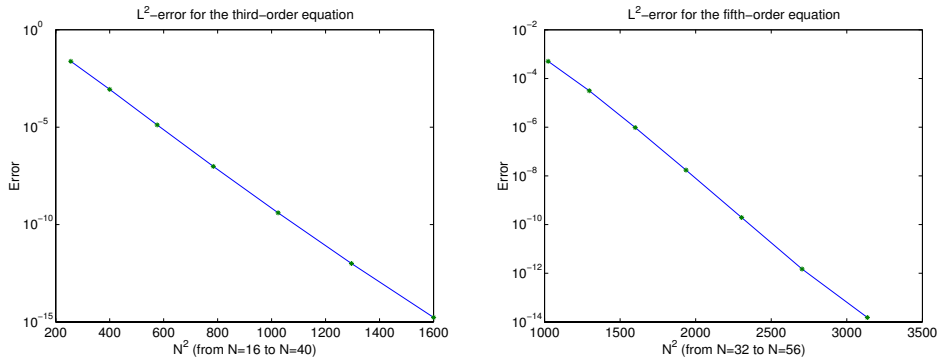


FIG. 6.1. L^2 -errors for the third- and fifth-order equations.

Next, we look at the accuracy of our dual-Petrov-Galerkin method. We take $a(x) = \sin x$ and $b(x) = 2x - 1$ in (2.4) and let the exact solution of (2.4) be $\cos(16\pi x)$. We plot in Figure 6.1 (left) the \log_{10} of the L^2 -error against N^2 . For the fifth-order equation (3.1), we take $\alpha = \beta = 1$ and the exact solution to be $\sin^3(3\pi x)$. The \log_{10} of the L^2 -error against N^2 is presented in Figure 6.1 (right). The straight lines in these plots indicate that the L^2 -errors converge like $\exp(-cN^2)$, a typical supergeometric convergence for analytic functions by spectral methods (cf. [6]).

6.2. KDV equation. Now, we present some numerical tests for the KDV equation. We first consider the initial value KDV problem

$$(6.1) \quad u_t + uu_x + u_{xxx} = 0, \quad u(x, 0) = u_0(x),$$

with the exact soliton solution

$$u(x, t) = 12\kappa^2 \operatorname{sech}^2(\kappa(x - 4\kappa^2 t - x_0)).$$

Since $u(x, t)$ converges to 0 exponentially as $|x| \rightarrow \infty$, we can approximate the initial value problem (6.1) by an initial-boundary value problem for $x \in (-M, M)$ as long as the soliton does not reach the boundaries.

We take $\kappa = 0.3$, $x_0 = -20$, $M = 50$, and $\Delta t = 0.001$ so that for $N \lesssim 160$, the time discretization error is negligible compared with the spatial discretization error.

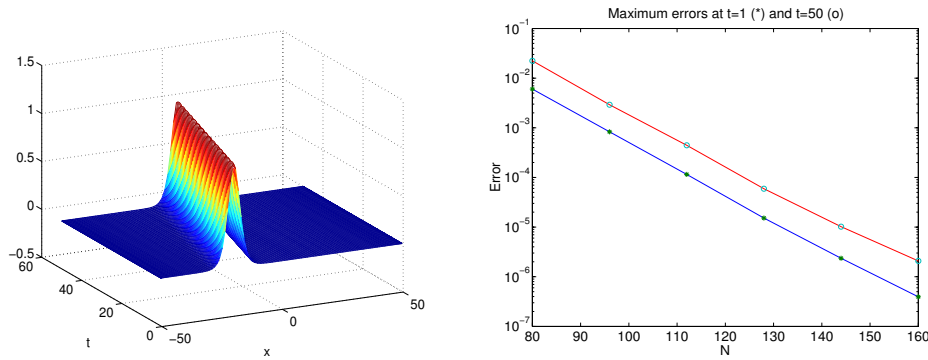


FIG. 6.2. Exact solution for the KDV equation: Left, time evolution; right, maximum error vs. N .

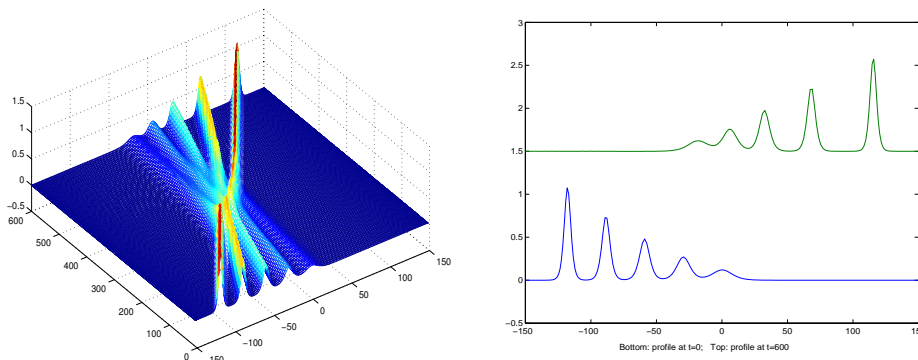


FIG. 6.3. Interaction of five solitary waves.

On the left of Figure 6.2, we plot the time evolution of the approximate solution, and on the right, we plot the maximum errors in the semi-log scale at $t = 1$ and $t = 50$. Note that the straight lines indicate that the errors converge like $\exp(-cN)$, which is typical for solutions that are infinitely differentiable but not analytic. The excellent accuracy for this known exact solution indicates that the KDV equation on a finite interval can be used to effectively simulate the KDV equation on a semi-infinite interval before the wave reaches the boundary.

In the following tests, we fix $M = 150$, $\Delta t = 0.02$, and $N = 256$.

Example 1: Interaction of five solitons. We start with the initial condition

$$u_0(x) = \sum_{i=1}^5 12\kappa_i^2 \operatorname{sech}^2(\kappa_i(x - x_i))$$

with

$$\begin{aligned} \kappa_1 &= .3, \quad \kappa_2 = .25, \quad \kappa_3 = .2, \quad \kappa_4 = .15, \quad \kappa_5 = .1, \\ x_1 &= -120, \quad x_2 = -90, \quad x_3 = -60, \quad x_4 = -30, \quad x_5 = 0. \end{aligned}$$

In Figure 6.3 (left), we plot the time evolution of the solution in the (x, t) plane. We also plot the initial profile and the profile at the final step ($t = 600$) in Figure 6.3 (right). We observe that the soliton with higher amplitude travels with faster speed,

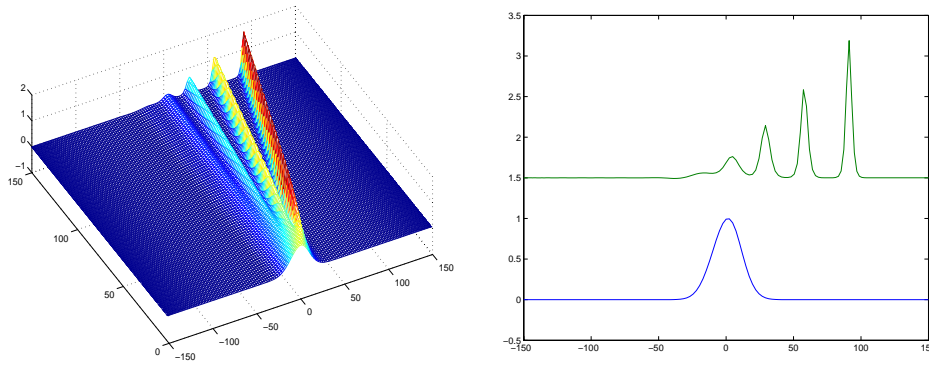


FIG. 6.4. Solitary waves generated by an initial Gaussian profile.

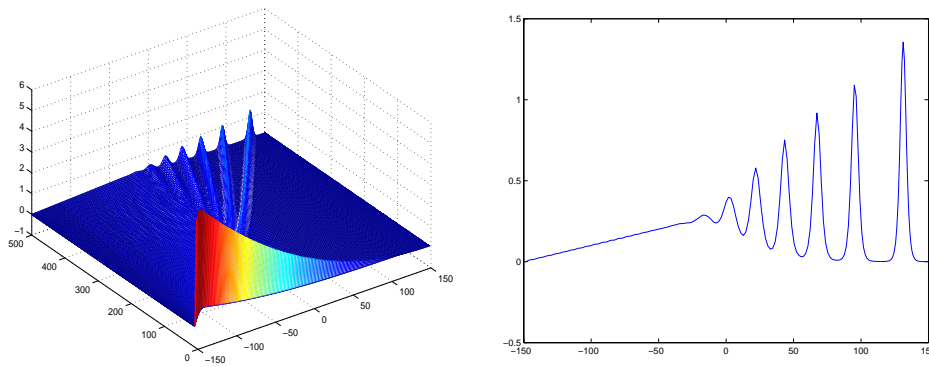


FIG. 6.5. Solitary waves generated by a square pulse (in time) at the left boundary.

and the amplitudes of the five solitary waves are well preserved at the final time. This indicates that our scheme has an excellent conservation property.

Example 2: Solitary waves generated by an initial Gaussian profile. We start with the initial condition $u_0(x) = \exp(-1.5x^2)$. We plot the time evolution of the solution on the left in Figure 6.4 and the profile at the final step ($t = 150$) on the right. The initial Gaussian profile has evolved into four separated solitary waves by the time $t = 150$.

Example 3: Solitary waves generated by a square pulse (in time) at the left boundary. In this example, we take $u_0(x) = 0$ but set

$$u(0, t) = \begin{cases} 5, & 0 \leq t \leq 5, \\ 0, & t > 5. \end{cases}$$

One may think of this situation as a dam of height five unit length that releases water for five unit time and is then shut off. We plot the time evolution of the solution on the left in Figure 6.5 and the profile at the final step ($t = 500$) on the right. The square pulse (in time) at the boundary generates a cascade of solitary waves as time evolves. This interesting phenomenon was first observed by Chu, Xiang, and Baransky in [8] (see also [12]).

7. Concluding remarks. We presented in this paper a new dual-Petrov–Galerkin method for third and higher odd-order equations. The key idea is to use test functions satisfying boundary conditions which are the “dual” of those for the trial functions. The resulting linear systems are sparse for problems with constant coefficients and well conditioned for problems with variable coefficients. By exploring the orthogonal properties of the test and trial basis functions in weighted Sobolev spaces, we were able to establish optimal error estimates for typical third-order and fifth-order linear equations and for a KDV equation on a finite interval. Obviously, the technique can be extended to other higher odd-order equations.

When combined with a Chebyshev–Legendre approach, our dual-Petrov–Galerkin method has a quasi-optimal computational complexity and is extremely accurate and efficient as illustrated by our numerical examples. Hence the method is most suitable for the study of complex dynamics of higher odd-order equations.

Finally, we note that the orthogonal polynomials $\{\phi_k\}$ and $\{\Phi_k\}$ introduced in this paper can be viewed as extensions of the Jacobi polynomials $J_k^{a,b}(x)$ with $(a, b) = (-2, -1)$ and $(a, b) = (-3, -2)$, respectively. They appear to be the most natural basis functions for, respectively, the third- and fifth-order equations (with the specified boundary conditions) considered in this paper. The extension of the Jacobi polynomials to more general (a, b) with $a, b < -1$ and their applications to spectral methods for other types of partial differential equations, including, in particular, hyperbolic systems, will be investigated in a forthcoming paper.

Acknowledgment. The author would like to thank Professor Benyu Guo for fruitful discussions and for carefully reading the first draft of the paper.

REFERENCES

- [1] B. K. ALPERT AND V. ROKHLIN, *A fast algorithm for the evaluation of Legendre expansions*, SIAM J. Sci. Statist. Comput., 12 (1991), pp. 158–179.
- [2] J. BONA AND R. WINTHER, *The Korteweg–de Vries equation, posed in a quarter-plane*, SIAM J. Math. Anal., 14 (1983), pp. 1056–1106.
- [3] J. L. BONA, W. G. PRITCHARD, AND L. R. SCOTT, *An evaluation of a model equation for water waves*, Philos. Trans. Roy. Soc. London Ser. A, 302 (1981), pp. 457–510.
- [4] J. L. BONA, S. M. SUN, AND B. Y. ZHANG, *A non-homogeneous boundary-value problem for the Korteweg–de Vries equation posed on a finite domain*, Comm. Partial Differential Equations, 28 (2003), pp. 1391–1436.
- [5] J. L. BONA, S. M. SUN, AND B. Y. ZHANG, *A non-homogeneous boundary-value problem for the Korteweg–de Vries equation in a quarter plane*, Trans. Amer. Math. Soc., 354 (2002), pp. 427–490.
- [6] J. P. BOYD, *Chebyshev and Fourier Spectral Methods*, 2nd ed., Dover Publications, Mineola, NY, 2001.
- [7] C. CANUTO, M. Y. HUSSAINI, A. QUARTERONI, AND T. A. ZANG, *Spectral Methods in Fluid Dynamics*, Springer-Verlag, New York, 1987.
- [8] C. K. CHU, L. W. XIANG, AND Y. BARANSKY, *Solitary waves induced by boundary motion*, Comm. Pure Appl. Math., 36 (1983), pp. 495–504.
- [9] T. COLIN AND J.-M. GHIDAGLIA, *An initial-boundary value problem for the Korteweg–de Vries equation posed on a finite interval*, Adv. Differential Equations, 6 (2001), pp. 1463–1492.
- [10] W. S. DON AND D. GOTTLIEB, *The Chebyshev–Legendre method: Implementing Legendre methods on Chebyshev points*, SIAM J. Numer. Anal., 31 (1994), pp. 1519–1534.
- [11] D. GOTTLIEB AND S. A. ORSZAG, *Numerical Analysis of Spectral Methods: Theory and Applications*, CBMS-NSF Regional Conf. Ser. in Appl. Math. 26, SIAM, Philadelphia, 1977.
- [12] B.-Y. GUO, *Numerical solution of an initial-boundary value problem of the Korteweg–de Vries equation*, Acta Math. Sci. (English Ed.), 5 (1985), pp. 337–348.

- [13] B.-Y. GUO AND J. SHEN, *On spectral approximations using modified Legendre rational functions: Application to the Korteweg-de Vries equation on the half line*, Indiana Univ. Math. J., 50 (2001), pp. 181–204. Dedicated to Professors Ciprian Foias and Roger Temam (Bloomington, IN, 2000).
- [14] B.-Y. GUO, *Spectral Methods and Their Applications*, World Scientific, River Edge, NJ, 1998.
- [15] B.-Y. GUO, *Jacobi approximations in certain Hilbert spaces and their applications to singular differential equations*, J. Math. Anal. Appl., 243 (2000), pp. 373–408.
- [16] J. L. HAMMACK AND H. SEGUR, *The Korteweg-de Vries equation and water waves. II, Comparison with experiments*, J. Fluid Mech., 65 (1974), pp. 289–313.
- [17] W. Z. HUANG AND D. M. SLOAN, *The pseudospectral method for third-order differential equations*, SIAM J. Numer. Anal., 29 (1992), pp. 1626–1647.
- [18] S. KICHENASSAMY AND P. J. OLVER, *Existence and nonexistence of solitary wave solutions to higher-order model evolution equations*, SIAM J. Math. Anal., 23 (1992), pp. 1141–1166.
- [19] H. MA AND W. SUN, *A Legendre–Petrov–Galerkin method and Chebyshev collocation method for the third-order differential equations*, SIAM J. Numer. Anal., 38 (2000), pp. 1425–1438.
- [20] H. MA AND W. SUN, *Optimal error estimates of the Legendre–Petrov–Galerkin method for the Korteweg–de Vries equation*, SIAM J. Numer. Anal., 39 (2001), pp. 1380–1394.
- [21] W. J. MERRYFIELD AND B. SHIZGAL, *Properties of collocation third-derivative operators*, J. Comput. Phys., 105 (1993), pp. 182–185.
- [22] J. SHEN, *Efficient spectral-Galerkin method I: Direct solvers of second- and fourth-order equations using Legendre polynomials*, SIAM J. Sci. Comput., 15 (1994), pp. 1489–1505.
- [23] J. SHEN, *Efficient spectral-Galerkin method II: Direct solvers of second- and fourth-order equations using Chebyshev polynomials*, SIAM J. Sci. Comput., 16 (1994), pp. 74–87.
- [24] J. SHEN, *Efficient Chebyshev-Legendre Galerkin methods for elliptic problems*, in Proceedings of ICOSAHOM’95, A. V. Ilin and R. Scott, eds., Houston J. Math., Houston, TX, 1996, pp. 233–240.
- [25] J. SHEN, *Efficient spectral-Galerkin methods III: Polar and cylindrical geometries*, SIAM J. Sci. Comput., 18 (1997), pp. 1583–1604.
- [26] J. SHEN, *Efficient spectral-Galerkin methods IV: Spherical geometries*, SIAM J. Sci. Comput., 20 (1999), pp. 1438–1455.
- [27] G. SZEGÖ, *Orthogonal Polynomials*, Amer. Math. Soc. Colloq. Publ. 23, AMS, New York, 1939.
- [28] N. J. ZABUSKY AND C. J. GALVIN, *Shallow water waves, the Korteweg-de-Vries equation and solitons*, J. Fluid Mech., 47 (1971), pp. 811–824.