Contemporary Mathematics
https://ojs.wiserpub.com/index.php/CM/

Research Article

Spectral Galerkin Scheme for the Post-Critical Dynamics of a Geometrically Nonlinear, Inextensible, Cantilevered Beam Subjected to Non-Conservative Forces

Kazuho Ito

Department of Regional Social Management, University of Yamanashi, Kofu, Japan
E-mail: ikazuho@yamanashi.ac.jp

Received: 10 December 2024; **Revised:** 11 March 2025; **Accepted:** 21 March 2025

Abstract: A spectral Galerkin scheme is proposed for a nonlinear model describing the large deflection and planar motion of an inextensible cantilevered beam subjected to non-conservative forces, such as Beck's beam and a cantilever immersed in axial airflow. It is established that the proposed scheme achieves spectral accuracy in space and second-order accuracy in time, relative to the sizes of the spatial and temporal grids, respectively. Furthermore, numerical simulations using the scheme are performed to investigate the post-critical limit cycle oscillations induced when the non-conservative forces exceed certain critical thresholds. The results provide a detailed analysis of how Kelvin-Voigt type internal damping destabilizes the system or increases the amplitude of the oscillation.

***Keywords*:** inextensible beam, spectral method, nonlinear Beck's beam, flutter

MSC: 65M70, 65M12, 65M15, 74B20, 74H15, 74S25

1. Introduction

In recent years, nonlinear partial differential equations describing the large deflection and planar motion of inextensible cantilevered beams have attracted considerable attention as models for analyzing the effects of non-conservative forces acting on such beams.

When a cantilevered beam is subjected to a follower force P (a compressive force at the free end in the tangential direction), it is referred to as Beck's beam, and has been widely studied as a simple model of structures subjected to the thrust, such as fluid-conveying pipes and rockets. It is known that there exists a critical value $P = P_{\text{crit}}$, beyond which the beam becomes unstable (see, e.g., [1, 2]), resulting in limit cycle oscillations (LCOs) [3, 4]. An intriguing aspect of Beck's beam is the destabilizing effect of internal damping, known as the Ziegler's paradox: within a certain range, an increase in damping reduces the critical value P_{crit} (see, e.g., [2, 5, 6]).

Similarly, when a cantilevered plate is immersed in axial airflow, there exists a critical flow velocity above which fluid pressure induces flutter, leading to LCOs [1, 7, 8]. Research interests in this context include understanding the onset of flutter (see, e.g., [9] and references therein) and examining the post-flutter regime, particularly at flow velocities where the piston theory is applicable, such as in supersonic regimes [10–12].

DOI: https://doi.org/10.37256/cm.6320256228

In both cases of non-conservative forces mentioned above, eigenvalue analysis based on linear structural models is commonly used to determine critical values and to examine their dependence on physical parameters such as damping. However, to investigate post-critical dynamics and capture the characteristics of LCOs, a robust nonlinear model capable of representing the large deflection motion of inextensible cantilevers is essential.

For models that describe the large deflection of slender beams, it is common to characterize the beam's shape in terms of the longitudinal and lateral deflections, u and v, from the undeformed state, and/or the rotation ϕ of the beam's centerline. In particular, for inextensible beams, since u and v can be expressed as functions of ϕ, the beam's motion can be formulated as a differential equation with ϕ as the unknown variable. This equation is known as the dynamic Euler elastica (see e.g., [3, 13–16]), which we refer to as the ϕ-model in this paper.

On the other hand, a nonlinear model for inextensible beams that has been widely used in the literature expresses the motion in terms of a single unknown variable, v. This model was originally derived by Crespo da Silva [17] by introducing certain approximations into the ϕ-model, and we refer to it as the v-model. While the v-model is commonly employed to study the effects of non-conservative forces (see e.g., [4, 8, 10, 11, 18–21]), as will be shown in Section 2, its derivation imposes geometric constraints on the beam's shape, meaning that the model is valid only within a limited range of $|\phi|$. Therefore, using the ϕ-model instead of the v-model is expected to yield more accurate and robust results, particularly in cases involving large non-conservative forces that induce large rotations ϕ.

This paper focuses on the ϕ-model with non-conservative forces and proposes a spectral Galerkin scheme to approximate solutions of the initial-boundary value problem (IBVP) for the model. For conservative problems, Ito [22] developed a numerical scheme with convergence rates of $O(N^{-m+\varepsilon}) + O(\Delta t^2)$, where N is the spatial grid size, Δt is the temporal grid spacing, and the exact solution of the IBVP is assumed to belong to the C^m class for some m. The present paper extends these results to non-conservative problems, establishing the same convergence properties. To the best of the author's knowledge, numerical studies on both the ϕ-model and the v-model have been conducted without providing a theoretical examination of computational accuracy-that is, a rigorous error analysis-of the numerical schemes employed, most of which are based on mode expansions using eigenfunctions of the linearized model (e.g., [3] for the ϕ-model, and [4, 8, 10] for the v-model). In contrast, our results of the convergence rate ensures that the proposed scheme achieves sufficient accuracy and numerical stability, thereby making the simulation results obtained with it reliable. This theoretical guarantee of accuracy represents a key contribution of the present work.

Furthermore, the proposed scheme facilitates a detailed examination of the post-critical dynamics of Beck's beam and cantilevers in axial flow over a wide range of speeds where the piston theory is valid, including high Mach numbers. Numerical simulations indicate that the proposed scheme is robust, as it successfully computes solutions even under large non-conservative forces and high initial velocities. This stands in stark contrast to conventional numerical methods based on the v-model, which fail to yield correct solutions due to issues such as numerical instability. As a result, the proposed scheme enables the elucidation of post-critical behavior in the regime of large deformations where large ϕ appears, which was previously inaccessible. For instance, it reveals how the amplitude of LCOs depends on the magnitude of internal damping and initial velocity.

2. Model equations

The ϕ-model under consideration consists of the system of partial differential equations

$$\rho A u_{tt} = \lambda_x + \bar{f}, \tag{1}$$

$$\rho A v_{tt} = \kappa_x + \bar{g}, \tag{2}$$

$$\rho I \phi_{tt} - EI\phi_{xx} - \alpha_0 EI\phi_{xxt} = -\lambda \sin\phi + \kappa\cos\phi + \bar{\mu}, \tag{3}$$

$$1+u_x=\cos\phi,\quad v_x=\sin\phi, \tag{4}$$

$$0<x<L,\quad t>0,$$

where $(\cdot)_t=\partial/\partial t$ and $(\cdot)_x=\partial/\partial x$. The unknowns of the equations are $u(x,\,t)$, $v(x,\,t)$, $\phi(x,\,t)$, $\lambda(x,\,t)$ and $\kappa(x,\,t)$, which are illustrated in Figure 1a. The system, known as the dynamic Euler elastica, describes the motion of an elastic beam undergoing geometrically large deflection (see e.g., [3, 15, 23, 24] for further descriptions and theoretical analyses).

In this model, it is assumed that:

• (planar motion) the centerline of the beam is lying in a plane with the orthogonal coordinate system O-XY, and, occupies, in the reference state, the interval $0\le X\le L$ of the X-axis; thus, denoting by $u(x,\,t)$ and $v(x,\,t)$ the longitudinal and transverse displacements, at time t, of the point that is at $X=x$ when the beam is in the reference state, the centerline is expressed as the parametric curve $x\mapsto\big(x+u(x,\,t),\,v(x,\,t)\big)$ in O-XY.

• (the Euler-Bernoulli assumption) a cross-section that is perpendicular to the X-axis in the reference state remains unchanged in shape and stays perpendicular to the centerline even after deformation; thus, the tangent $\big(1+u_x(x,\,t),v_x(x,\,t)\big)$ of the centerline has the same direction as the normal of $\Gamma(x,\,t)$, where $\Gamma(x,\,t)$ denotes the cross-section that is at $X=x$ when the beam is in the reference state.

• the beam is inextensible, i.e., $(1+u_x)^2+v_x^2=1$; thus, denoting by $\phi(x,\,t)$ the angle between the normal of $\Gamma(x,\,t)$ and the X-axis, it follows that $(1+u_x,\,v_x)=(\cos\phi,\,\sin\phi)$.

The variables $\lambda(x,\,t)$ and $\kappa(x,\,t)$ represent the components of the axial force in the X- and Y-directions acting on the cross-section $\Gamma(x,\,t)$, so that their derivatives, $\lambda_x(x,\,t)$ and $\kappa_x(x,\,t)$, represent the magnitudes of the forces acting per unit length of the beam at the corresponding point.

The parameters in the model are the mass density per unit volume ρ, the cross-sectional area A, the moment of inertia of the cross section I, the Young's modulus E, the viscous coefficient α_0 of Kelvin-Voigt type internal damping, and the beam length L. It is assumed that the cross-sectional shape of the beam is uniform and the material is homogeneous, so that all these parameters are constants.

The external distributed forces $\bar f(x,\,t)$, $\bar g(x,\,t)$, and torque $\bar\mu(x,\,t)$, illustrated in Figure 1b, are specified as

$$\bar f=-\bar n\sin\phi,\quad \bar g=\bar n\cos\phi,\quad \bar\mu=0, \tag{5}$$

where $\bar n(x,\,t)$ is the magnitude of the distributed external force in the normal direction to the beam. When the beam is immersed in the airflow with the velocity U_∞ directed along X-axis as shown in Figure 1b, the pressure across the beam causes a normal force, which we express as

$$\bar n=-\gamma_0(U_\infty\sin\phi-u_t\sin\phi+v_t\cos\phi). \tag{6}$$

In the case that ϕ is sufficiently small such that $\cos\phi\approx1$ and $u_t\sin\phi$ is negligible, the expression (6) corresponds to the first order piston theory (see e.g., [9, 10]):

$$\bar n=-\gamma_0(U_\infty v_x+v_t),$$

where γ_0 is a non-negative coefficient depending on the density, velocity and Mach number of the flow. Though $U_\infty = 0$ implies $\gamma_0 = 0$ in the piston theory, we shall consider the case $U_\infty = 0$ and $\gamma_0 > 0$ as well in our numerical simulation, for which $\bar{f}$ and $\bar{g}$ in (5) together with (6) represent the external damping proportional to the normal component of the beam's velocity.

The beam is clamped at one end and subjected to boundary forces $\bar{\lambda}_L$, $\bar{\kappa}_L$, and a torque $\bar{\mu}_L$ at the free end:

$$\lambda_x(0,\,t) + \bar{f}(0,\,t) = \kappa_x(0,\,t) + \bar{g}(0,\,t) = \phi(0,\,t) = 0, \tag{7}$$

$$\lambda(L,\,t) = \bar{\lambda}_L(t), \quad \kappa(L,\,t) = \bar{\kappa}_L(t), \quad EI\phi_x(L,\,t) + \alpha_0 EI\phi_{xt}(L,\,t) = \bar{\mu}_L(t), \tag{8}$$

where

$$\bar{\lambda}_L(t) = -P_0\cos\phi(L,\,t), \quad \bar{\kappa}_L(t) = -P_0\sin\phi(L,\,t), \quad \bar{\mu}_L(t) = 0, \tag{9}$$

and P_0 represents the intensity of a constant compressive follower force at the free end, as illustrated in Figure 1b.

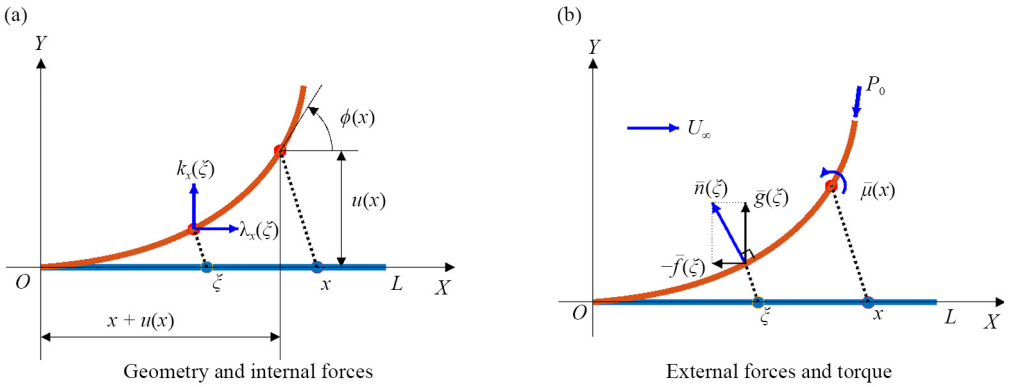

Figure 1. Schematic of inextensible cantilevered beam (time variable t being suppressed)

Let us now outline the relationship between the ϕ-model and the v-model, the latter being frequently employed in the literature to investigate post-critical dynamics or LCOs. The ϕ-model (1)-(3), along with boundary conditions (7) and (8), can be derived via Hamilton's principle:

$$\delta\int_{t_1}^{t_2}(T-V)\;\mathrm{d}t + \int_{t_1}^{t_2}\delta W\;\mathrm{d}t = 0 \quad \text{with (4)}, \tag{10}$$

where

$$T = \frac{1}{2}\int_0^L \left(\rho I\phi_t^2 + \rho A u_t^2 + \rho A v_t^2\right)\;\mathrm{d}x, \quad V = \frac{1}{2}\int_0^L EI\phi_x^2\;\mathrm{d}x, \tag{11}$$

$$\delta W = \int_0^L \left(-\alpha_0 EI\phi_{xt}\delta\phi_x + \bar{f}\delta u + \bar{g}\delta v + \bar{\mu}\delta\phi\right)\,\mathrm{d}x + \left[\bar{\lambda}_L\delta u + \bar{\kappa}_L\delta v + \bar{\mu}_L\delta\phi\right]_{x=L}. \tag{12}$$

On the other hand, the v-model is derived by introducing the following assumptions into the principle for the ϕ-model above.

(A1) The rotational kinetic energy $(1/2)\int_0^L \rho I\phi_t^2\,\mathrm{d}x$ is negligible. This is common when modeling a slender beam, where the coefficient ρI of the kinetic energy associated with angular velocity ϕ_t is much smaller than the coefficient ρA associated with linear velocities u_t and v_t.

(A2) Variables are expanded in terms of v_x, retaining terms up to the third order:

(a) $u_x \approx -v_x^2/2$, derived from the inextensibility condition:

$$1 + u_x = \sqrt{1 - v_x^2} \approx 1 - v_x^2/2,$$

(b) $\phi \approx v_x + v_x^3/6$, derived from $\phi = \arctan\left(v_x/\sqrt{1-v_x^2}\right)$,

(c) $\phi_x^2 \approx v_{xx}^2(1+v_x^2)$, derived from (b).

Applying these approximations to (11), (12), the principle (10) changes as

$$\delta\int_{t_1}^{t_2}(T - V + G)\,\mathrm{d}t + \int_{t_1}^{t_2}\delta W\,\mathrm{d}t = 0,$$

where

$$T = \frac{1}{2}\int_0^L \left(\rho A u_t^2 + \rho A v_t^2\right)\,\mathrm{d}x, \quad V = \frac{1}{2}\int_0^L EIv_{xx}^2\left(1+v_x^2\right)\,\mathrm{d}x, \quad G = \frac{1}{2}\int_0^L \sigma\left(2u_x + v_x^2\right)\,\mathrm{d}x,$$

$$\delta W = \int_0^L \Big(-\alpha_0 EI\left\{v_{xx}\left(1+v_x^2/2\right)\right\}_t\ \delta\left(v_{xx}\left(1+v_x^2/2\right)\right) + \bar{f}\delta u + \bar{g}\delta v + \bar{\mu}\left(1+v_x^2/2\right)\ \delta v_x\Big)\,\mathrm{d}x$$

$$+ \left[\bar{\lambda}_L\delta u + \bar{\kappa}_L\delta v + \bar{\mu}_L\left(1+v_x^2/2\right)\delta v_x\right]_{x=L},$$

and G accounts for the constraint (A2) (a) with σ as the Lagrange multiplier. This yields the motion equation:

$$\rho A v_{tt} + EIv_{xxxx} + EIA_0(v)_x + \alpha_0 EIA_1(v)_x + \left(v_x\int_x^L \left(\rho A u_{tt} - \bar{f}\right)\,\mathrm{d}\xi - \bar{\lambda}_L v_x\right)_x - \bar{g} + \left\{\bar{\mu}\left(1+v_x^2/2\right)\right\}_x = 0, \tag{13}$$

$$u = -\frac{1}{2}\int_0^x v_x^2\,\mathrm{d}\xi,$$

with natural boundary conditions

$$\Big[EIv_{xxx}+EIA_0(v)+\alpha_0 EIA_1(v)-\bar{\lambda}_L v_x+\bar{\kappa}_L+\bar{\mu}\big(1+v_x^2/2\big)\Big]_{x=L}=0,$$

$$\Big[-EIv_{xx}\big(1+v_x^2\big)-\alpha_0 EI\big\{v_{xx}\big(1+v_x^2/2\big)\big\}_t\big(1+v_x^2/2\big)+\bar{\mu}_L\big(1+v_x^2/2\big)\Big]_{x=L}=0,$$

where

$$A_0(v)=-v_{xx}^2v_x+\big(v_{xx}v_x^2\big)_x,\quad A_1(v)=\big\{v_{xx}\big(1+v_x^2/2\big)\big\}_{xt}\big(1+v_x^2/2\big).$$

This is the v-model, which is widely used in the literature (see [5, 8, 17, 25, 26] for further descriptions of the model's derivation). However, for the internal damping term, one finds several variations, other than $\alpha_0 EIA_1(v)_x$, in the literature, such as (i) $\alpha_0\big(EIv_{xxxx}+EIA_0(v)_x\big)_t$ (e.g., [5, 19]), which is Kelvin-Voigt damping based on the stiffness terms $EIv_{xxxx}+EIA_0(v)_x$, (ii) the linear part $\alpha_0 EIv_{xxxxt}$ of (i), which is the common form for the linear Euler-Bernoulli beam.

Note that the assumption (A2) (a) and (b) can be expressed in terms of the rotation ϕ as

$$1-\cos\phi\approx\frac{1}{2}\sin^2\phi,\quad \phi\approx\sin\phi+\frac{1}{6}\sin^3\phi,$$

respectively. Therefore, the model (13) is valid only within a limited range of beam motion or $|\phi|$ where these approximations remains effective. In contrast, the ϕ-model does not impose such restrictions. Thus, by combining it with an efficient numerical scheme, it becomes possible to extend the range of beam deformation motions that can be handled in post-critical analysis to cases where large ϕ appears.

We conclude this section with presenting a normalized form of the ϕ model. Introducing the parameters

$$\beta=\frac{AL^2}{I},\quad \alpha=\alpha_0\sqrt{\frac{E}{\rho L^2}},\quad \gamma=\frac{\gamma_0 L}{A\sqrt{\rho E}},\quad U=U_\infty\sqrt{\frac{\rho}{E}},\quad P=\frac{P_0L^2}{EI},\tag{14}$$

and setting as

$$\tilde{x}=\frac{x}{L},\quad \tilde{t}=t\sqrt{\frac{E}{\rho L^2}},\quad \tilde{u}=\frac{u}{L},\quad \tilde{v}=\frac{v}{L},\quad \tilde{\phi}=\phi,\quad \tilde{\lambda}=\frac{\lambda}{EA},\quad \tilde{\kappa}=\frac{\kappa}{EA},$$

$$\tilde{\bar{f}}=\frac{L\bar{f}}{EA},\quad \tilde{\bar{g}}=\frac{L\bar{g}}{EA},\quad \tilde{\bar{\mu}}=\frac{L^2\bar{\mu}}{EI},\quad \tilde{\bar{\lambda}}_L=\frac{\bar{\lambda}_L}{EA},\quad \tilde{\bar{\kappa}}_L=\frac{\bar{\kappa}_L}{EA},\quad \tilde{\bar{\mu}}_L=\frac{L\,\bar{\mu}_L}{EI},$$

the IBVP (1)-(9) is rewritten as follows (tilde being removed):

$$u_{tt} = \lambda_x + \bar{f}, \tag{15}$$

$$v_{tt} = \kappa_x + \bar{g}, \tag{16}$$

$$\phi_{tt} - \phi_{xx} - \alpha\phi_{xxt} = \beta\left(-\lambda\sin\phi + \kappa\cos\phi\right) + \bar{\mu}, \tag{17}$$

$$1 + u_x = \cos\phi, \quad v_x = \sin\phi, \tag{18}$$

$$0 < x < 1, \quad t > 0,$$

$$\bar{f} = -\bar{n}\sin\phi, \quad \bar{g} = \bar{n}\cos\phi, \quad \bar{\mu} = 0, \tag{19}$$

where

$$\bar{n} = -\gamma(U\sin\phi - u_t\sin\phi + v_t\cos\phi). \tag{20}$$

The boundary conditions are

$$\lambda_x(0,\,t) + \bar{f}(0,\,t) = \kappa_x(0,\,t) + \bar{g}(0,\,t) = \phi(0,\,t) = 0, \tag{21}$$

$$\lambda(1,\,t) = \bar{\lambda}_L(t), \quad \kappa(1,\,t) = \bar{\kappa}_L(t), \quad \phi_x(1,\,t) + \alpha\phi_{xt}(1,\,t) = \bar{\mu}_L(t), \tag{22}$$

where

$$\bar{\lambda}_L(t) = -(P/\beta)\cos\phi(1,\,t), \quad \bar{\kappa}_L(t) = -(P/\beta)\sin\phi(1,\,t), \quad \bar{\mu}_L(t) = 0. \tag{23}$$

3. Variational form of the IBVP

The system (15)-(20) with the boundary conditions (21)-(23) can be formulated in a variational form. Integrating (15) and (16) twice in x, and using the boundary conditions on λ and κ in (21), (22), we have

$$\lambda(x,\,t) = -\int_x^1\int_0^\eta (\cos\phi)_{tt}\ \mathrm{d}\xi\mathrm{d}\eta + \int_x^1 \bar{f}\,\mathrm{d}\eta + \bar{\lambda}_L,$$

$$\kappa(x,\,t) = -\int_x^1\int_0^\eta (\sin\phi)_{tt}\ \mathrm{d}\xi\mathrm{d}\eta + \int_x^1 \bar{g}\,\mathrm{d}\eta + \bar{\kappa}_L.$$

Let

$$\mathscr{X} = \left\{ \psi \in C^m([0,\ 1]) \mid \psi(0) = 0 \right\}$$

for some $m \geq 1$ which we shall specify latter on, where $C^m([0,\ 1])$ denotes the set of continuous functions on $[0,\ 1]$ that have m continuous derivatives. Multiplying (17) with $\psi \in \mathscr{X}$, integrating the result in x from 0 to 1, using the above representation of λ and κ, and applying outer forces (19), (20), (23), we can arrive at the variational form of the problem:

$$\begin{aligned} &\int_0^1 \phi_{tt}\psi \ \mathrm{d}x + \int_0^1 \left(\phi_x + \alpha\phi_{xt}\right)\psi_x \ \mathrm{d}x \\ =&\beta\left\{ a\big((\cos\phi)_{tt},\ \psi\sin\phi\big) - a\big((\sin\phi)_{tt},\ \psi\cos\phi\big) - \gamma\int_0^1 \big(U\sin\phi + B_\phi\phi_t\big)B_\phi\psi \ \mathrm{d}x \right\} - Pb_\phi(\psi) \end{aligned} \tag{24}$$

for any test function ψ belonging to $\mathscr{X}$. In the above, a is a bilinear form defined by

$$a(f,\ g) = \int_0^1 \left(\int_0^x f(\xi)\ \mathrm{d}\xi\right)\left(\int_0^x g(\xi)\ \mathrm{d}\xi\right)\ \mathrm{d}x \quad \text{for } f, g \in L^2(0,\ 1),$$

B_ϕ a linear operator defined by

$$[B_\phi f](x,\ t) = \sin\phi(x,\ t)\int_0^x f(\xi)\sin\phi(\xi,\ t)\ \mathrm{d}\xi + \cos\phi(x,\ t)\int_0^x f(\xi)\cos\phi(\xi,\ t)\ \mathrm{d}\xi \quad \text{for } f \in L^2(0,\ 1),$$

and b_ϕ a linear functional defined by

$$[b_\phi(f)](t) = -\cos\phi(1,\ t)\int_0^1 f(x)\sin\phi(x,\ t)\ \mathrm{d}x + \sin\phi(1,\ t)\int_0^1 f(x)\cos\phi(x,\ t)\ \mathrm{d}x. \quad \text{for } f \in L^2(0,\ 1),$$

where $L^2(0,\ 1)$ denotes the space of square integrable functions on $(0,\ 1)$ with the inner product $(f,\ g)_{L^2}$ and the norm $\|f\|_{L^2}$.

The system we are concerned with in the rest of the paper is the Equation (24) subjected to the boundary conditions

$$\phi(0,\ t) = 0, \quad \phi_x(1,\ t) + \alpha\phi_{xt}(1,\ t) = 0 \tag{25}$$

and the initial conditions

$$\phi(x,\ 0) = \phi^0(x), \quad \phi_t(x,\ 0) = \psi^0(x) \tag{26}$$

for those data ϕ^0, ψ^0 that are sufficiently smooth, and satisfy some compatibility conditions including (25).

We conclude this section with the energy equality. Define the total energy $\mathscr{E}(t)$ of the system by

$$\mathscr{E}(t) = \frac{1}{2}\int_0^1 \left\{(\phi_t)^2 + (\phi_x)^2\right\}\,\mathrm{d}x + \frac{\beta}{2}a\big((\cos\phi)_t,\,(\cos\phi)_t\big) + \frac{\beta}{2}a\big((\sin\phi)_t,\,(\sin\phi)_t\big) \tag{27}$$

$$\left(\int_0^1 u_t^2\,\mathrm{d}x = a\big((\cos\phi)_t,\,(\cos\phi)_t\big),\quad \int_0^1 v_t^2\,\mathrm{d}x = a\big((\sin\phi)_t,\,(\sin\phi)_t\big)\right),$$

then, by setting $\psi = \phi_t$ in (24) and using the boundary conditions (25), we can see that the solution of the problem (24) with (25) satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathscr{E}(t) = -\alpha\int_0^1 \phi_{xt}^2\,\mathrm{d}x - \beta\gamma\int_0^1 (B_\phi\phi_t)^2\,\mathrm{d}x - \beta\gamma U\int_0^1 (B_\phi\phi_t)\sin\phi\,\mathrm{d}x - Pb_\phi(\phi_t), \tag{28}$$

and we see

$$B_\phi\phi_t = -u_t\sin\phi + v_t\cos\phi,\quad b_\phi(\phi_t) = \big[u_t\cos\phi + v_t\sin\phi\big]_{x=1}.$$

On the right-hand side of (28), the first and second terms represent the dissipation due to internal and external damping, while the third and fourth terms reflect non-conservativity caused by an axial flow and a follower force, respectively.

4. Discretization of the IBVP

4.1 *Galerkin scheme with Legendre grid*

Let $\mathbb{P}_k$ be the space of polynomials of degree k on $0 \le x \le 1$. We denote by $L_k(x) \in \mathbb{P}_k$ the k-th Legendre polynomial, and by

$$0 = x_0 < x_1 < \cdots < x_{N-1} < x_N = 1$$

and $\{\omega_j : j = 0,\,1,\,\ldots,\,N\}$ the Legendre-Gauss-Lobatto quadrature nodes and weights, respectively (see e.g., [27, 28]): these correspond to the Legendre polynomial $\hat{L}_k(\xi)$ defined usually on $-1 \le \xi \le 1$ and the associated Gauss-Lobatto quadrature node ξ_j and weight $\hat{\omega}_j$ through

$$L_k(x) = \hat{L}_k(2x-1),\quad x_j = (1+\xi_j)/2,\quad \omega_j = \hat{\omega}_j/2.$$

For $f,\,g \in C([0,\,1])$ (= the set of continuous functions on $[0,\,1]$), we introduce the discrete inner product and norms by

$$(f,\,g)_N = \sum_{j=0}^{N} f(x_j)g(x_j)\omega_j,\quad \|f\|_N = \sqrt{(f,\,f)_N},\quad \|f\|_{N,\,\infty} = \max_j |f(x_j)|.$$

It is known that

$$\|p\|_{L^2}^2 \le \|p\|_N^2 \le \left(2+\frac{1}{N}\right)\|p\|_{L^2}^2,$$

$$(p,\, L_k)_N = (p,\, L_k)_{L^2} \times \begin{cases} 1: & k = 0,\, 1,\, \ldots,\, N-1 \\ \\ 2+1/N: & k = N \end{cases}, \quad p \in \mathbb{P}_N,$$

$$\|L_k\|_{L^2}^2 = \frac{1}{2k+1}, \quad |L_k(x)| \le 1, \quad k = 0,\, 1,\, \ldots. \tag{29}$$

We employ $\{x_j : j = 0,\, 1,\, \ldots,\, N\}$ as a grid of the spatial domain $0 \le x \le 1$. Let $\{\chi_j(x) : j = 0,\, 1,\, \ldots,\, N\} \subset \mathbb{P}_N$ be the Lagrange basis associated with $\{x_j\}$, i.e.,

$$\chi_j(x_l) = \delta_{jl} := \begin{cases} 1 \ : \ j = l, \\ \\ 0 \ : \ j \ne l, \end{cases} \qquad j,\, l = 0,\, 1,\, \ldots,\, N.$$

For $f \in C([0,\, 1])$, let $I_N f$ denote the polynomial of degree N that interpolates f at the grid $\{x_j\}$, so that

$$[I_N f](x) = \sum_{j=0}^{N} f(x_j)\chi_j(x) \in \mathbb{P}_N.$$

The temporal domain $t \ge 0$ is discretized by placing grid points

$$t_n = n\Delta t, \quad n = 0,\, 1,\, \ldots$$

with a grid spacing Δt.

Throughout this paper, we make use of the following notation for any temporal sequence $\{\theta^n\}$:

$$\Delta_t^0\theta^n = \frac{\theta^{n+1}-\theta^{n-1}}{2\Delta t}, \qquad \Delta_t^2\theta^n = \frac{\theta^{n+1}-2\theta^n+\theta^{n-1}}{\Delta t^2},$$

$$\Delta_t^+\theta^n = \frac{\theta^{n+1}-\theta^n}{\Delta t}, \qquad \Delta_t^-\theta^n = \frac{\theta^n-\theta^{n-1}}{\Delta t}, \qquad \bar{\theta}^n = \frac{\theta^{n+1}+\theta^{n-1}}{2}.$$

Furthermore, we define

$$C_\theta^n = C\big(\theta^{n+1},\, \theta^{n-1}\big), \qquad S_\theta^n = S\big(\theta^{n+1},\, \theta^{n-1}\big),$$

$$\tilde{C}_\theta^n = \tilde{C}(\theta^{n+1},\ \theta^n,\ \theta^{n-1}), \qquad \tilde{S}_\theta^n = \tilde{S}(\theta^{n+1},\ \theta^n,\ \theta^{n-1}),$$

where

$$C(\theta_1,\ \theta_2) = \frac{\sin\theta_1 - \sin\theta_2}{\theta_1 - \theta_2} \quad (= \cos\theta_1 \text{ if } \theta_1 = \theta_2),$$

$$S(\theta_1,\ \theta_2) = \frac{-(\cos\theta_1 - \cos\theta_2)}{\theta_1 - \theta_2} \quad (= \sin\theta_1 \text{ if } \theta_1 = \theta_2),$$

and

$$\tilde{C}\big(\theta_1,\ \theta_2,\ \theta_3\big) = \begin{cases} \cos\theta_2 & : \theta_1 = \theta_2 = \theta_3, \\ \dfrac{\sin\theta_1 - S\big(\theta_1,\ \theta_2\big)}{\big(\theta_1 - \theta_2\big)/2} & : \theta_1 = \theta_3 \neq \theta_2, \\ \dfrac{S\big(\theta_1,\ \theta_2\big) - S\big(\theta_3,\ \theta_2\big)}{\big(\theta_1 - \theta_3\big)/2} & : \text{otherwise}, \end{cases}$$

$$\tilde{S}\big(\theta_1,\ \theta_2,\ \theta_3\big) = \begin{cases} \sin\theta_2 & : \theta_1 = \theta_2 = \theta_3, \\ -\dfrac{\cos\theta_1 - C\big(\theta_1,\ \theta_2\big)}{\big(\theta_1 - \theta_2\big)/2} & : \theta_1 = \theta_3 \neq \theta_2, \\ -\dfrac{C\big(\theta_1,\ \theta_2\big) - C\big(\theta_3,\ \theta_2\big)}{\big(\theta_1 - \theta_3\big)/2} & : \text{otherwise}. \end{cases}$$

The representations C_θ^n and S_θ^n for approximating $\cos\theta^n$ and $\sin\theta^n$, respectively, were introduced by Falk and Xu [29]. Combined with the expressions C_θ^n and S_θ^n, the central difference scheme for discretizing the time derivatives $(\cos\phi)_{tt}$ and $(\sin\phi)_{tt}$ makes our scheme energy-conservative when the system has no external excitations, i.e., $U = P = 0$, as shown in (37) below. This property not only imparts numerical stability to our scheme, but also is one of the keys to ensuring that the obtained numerical solution converges to the true solution, as will be seen in the main theorem of the present paper.

A direct calculation leads to

$$\begin{aligned} \Delta_t^2 \cos\theta^n &= -(\Delta_t^2\theta^n)S_\theta^n - (\Delta_t^+\theta^n)(\Delta_t^-\theta^n)\tilde{C}_\theta^n, \\ \Delta_t^2 \sin\theta^n &= (\Delta_t^2\theta^n)C_\theta^n - (\Delta_t^+\theta^n)(\Delta_t^-\theta^n)\tilde{S}_\theta^n. \end{aligned} \tag{30}$$

For $f, g \in C([0,\ 1])$, we define $J_N f \in \mathbb{P}_{N+1}$ by

$$[J_N f](x) = \int_0^x [I_N f](\xi)\ \mathrm{d}\xi,$$

and the bilinear form a_N by

$$a_N(f,\ g) = \left(J_N f,\ J_N g\right)_N$$

with the abbreviation

$$a_N(f) = a_N(f,\ f).$$

We shall approximate by φ_j^n the solution $\phi(x_j,\ t_n)$ of the IBVP (24)-(26), and denote

$$\varphi^n(x) = \sum_{j=0}^{N} \varphi_j^n \chi_j(x) \in \mathscr{X}_N, \tag{31}$$

where

$$\mathscr{X}_N = \{p \in \mathbb{P}_N \mid p(0) = 0\}.$$

The numerical scheme we propose for solving the IBVP is

$$\begin{aligned}
&\left(\Delta_t^2 \varphi^n,\ \psi\right)_N + \left(\bar{\varphi}_x^n + \alpha(\Delta_t^0 \varphi^n)_x, \psi\right)_N \\
=&\beta a_N\left(\Delta_t^2 \cos\varphi^n,\ S_\varphi^n \psi\right) - \beta a_N\left(\Delta_t^2 \sin\varphi^n,\ C_\varphi^n \psi\right) \\
&- \beta\gamma\left(U \sin\varphi^n + B_{N,\ \varphi}^n(\Delta_t^0 \varphi^n),\ B_{N,\ \varphi}^n \psi\right)_N - P b_{N,\ \varphi}^n(\psi), \quad n = 1,\ 2,\ \ldots
\end{aligned} \tag{32}$$

for any $\psi \in \mathscr{X}_N$, where

$$B_{N,\ \varphi}^n f = J_N(f S_\varphi^n) \sin\varphi^n + J_N(f C_\varphi^n) \cos\varphi^n,$$

$$b_{N,\ \varphi}^n(f) = -\left(f,\ S_\varphi^n\right)_N \cos\varphi^n(1) + \left(f,\ C_\varphi^n\right)_N \sin\varphi^n(1), \quad f \in C([0,\ 1]).$$

The initial data φ^0 and φ^1 are constructed as follows:

$$\varphi_j^0 = \phi^0(x_j), \tag{33}$$

$$\varphi_j^1 = \phi^0(x_j) + \Delta t\,\psi^0(x_j) + \frac{\Delta t^2}{2} w_N(x_j), \quad j = 0,\ 1,\ \ldots,\ N, \tag{34}$$

in terms of the solution $w_N \in \mathscr{X}_N$ of the variational equation

$$\begin{aligned}
&\big(w_N,\ \psi\big)_N + \beta a_N\big(w_N \cos\phi^0,\ \psi\cos\phi^0\big) + \beta a_N\big(w_N \sin\phi^0,\ \psi\sin\phi^0\big) \\
&= -\big((I_N\phi^0)_x,\ \psi_x\big)_N - \alpha\big((I_N\psi^0)_x,\ \psi_x\big)_N + \beta a_N\big((\psi^0)^2\sin\phi^0,\ \psi\cos\phi^0\big) - \beta a_N\big((\psi^0)^2\cos\phi^0,\ \psi\sin\phi^0\big) \\
&\quad - \beta\gamma\big(U\sin\psi^0 + B^0_{N,\,\phi}\psi^0, B^0_{N,\,\phi}\psi\big)_N - Pb^0_{N,\,\phi}(\psi)
\end{aligned} \tag{35}$$

for any $\psi \in \mathscr{X}_N$, where

$$B^0_{N,\,\phi} f = J_N(f\sin\phi^0)\sin\phi^0 + J_N(f\cos\phi^0)\cos\phi^0,$$

$$b^0_{N,\,\phi}(f) = -\big(f,\ \sin\phi^0\big)_N \cos\phi^0(1) + \big(f,\ \cos\phi^0\big)_N \sin\phi^0(1), \quad f \in C([0,\ 1]).$$

Define the discrete energy $\mathscr{E}^n$ by

$$\mathscr{E}^n = \frac{1}{2}\left\{\big|\big|\Delta_t^+\varphi^n\big|\big|_N^2 + \int_0^1 \frac{(\varphi_x^{n+1})^2 + (\varphi_x^n)^2}{2}\,\mathrm{d}x + \beta a_N\big(\Delta_t^+\cos\varphi^n\big) + \beta a_N\big(\Delta_t^+\sin\varphi^n\big)\right\}, \tag{36}$$

then, setting $\psi = \Delta_t^0\varphi^n$ in (32), we obtain the energy equality (see [22])

$$\begin{aligned}
\frac{\mathscr{E}^n - \mathscr{E}^{n-1}}{\Delta t} &= -\alpha||(\Delta_t^0\varphi^n)_x||_N^2 - \beta\gamma||B^n_{N,\,\varphi}(\Delta_t^0\varphi^n)||_N^2 \\
&\quad - \beta\gamma U\big(\sin\varphi^n, B^n_{N,\,\varphi}(\Delta_t^0\varphi^n)\big)_N - Pb^n_{N,\,\varphi}(\Delta_t^0\varphi^n), \quad n = 1,\ 2,\ \ldots
\end{aligned} \tag{37}$$

and we see

$$B^n_{N,\,\varphi}(\Delta_t^0\varphi^n) = -(\Delta_t^0 u^n)\sin\varphi^n + (\Delta_t^0 v^n)\cos\varphi^n,$$

$$b^n_{N,\,\varphi}(\Delta_t^0\varphi^n) = \big[(\Delta_t^0 u^n)\cos\varphi^n + (\Delta_t^0 v^n)\sin\varphi^n\big]_{x=1},$$

where

$$x+u^n(x)=[J_N(\cos\varphi^n)](x),\quad v^n(x)=[J_N(\sin\varphi^n)](x).$$

Note that (37) is the discrete counterpart of (28).

We conclude this section with the matrix form and the implementation of our scheme. Substituting (30) with $\theta=\varphi$ into (32), and using (31), we can see that the equation (32) with $\psi=\chi_i$ is written as

$$\sum_{j=0}^{N}P_{ij}^n(\Delta_t^2\varphi_j^n)+\sum_{j=0}^{N}Q_{ij}^n(\Delta_t^+\varphi_j^n)+\sum_{j=0}^{N}R_{ij}^n(\Delta_t^0\varphi_j^n)+\sum_{j=0}^{N}G_{ij}\big(\bar{\varphi}_j^n+\alpha\Delta_t^0\varphi_j^n\big)+T_i^n=0,\quad i=0,\ \ldots,\ N, \tag{38}$$

(the equation at $i=0$ is replaced with $\varphi_0^n=0$)

where

$$P_{ij}^n=\delta_{ij}\omega_j+\beta\Big\{a_N(C_\varphi^n\chi_i,\ C_\varphi^n\chi_j)+a_N(S_\varphi^n\chi_i,\ S_\varphi^n\chi_j)\Big\}\ \Big(\text{note that }\delta_{ij}\omega_j=(\chi_i,\ \chi_j)_N\Big),$$

$$Q_{ij}^n=\beta\Big\{a_N\big(S_\varphi^n\chi_i,\ (\Delta_t^-\varphi^n)\tilde{C}_\varphi^n\chi_j\big)-a_N\big(C_\varphi^n\chi_i,\ (\Delta_t^-\varphi^n)\tilde{S}_\varphi^n\chi_j\big)\Big\},$$

$$R_{ij}^n=\beta\gamma\big(B_{N,\ \varphi}^n\chi_i,B_{N,\ \varphi}^n\chi_j\big)_N,$$

$$T_i^n=\beta\gamma U\big(B_{N,\ \varphi}^n\chi_i,\sin\varphi^n\big)_N+P\big(-\cos\varphi_N^n\sin\varphi_i^n+\sin\varphi_N^n\cos\varphi_i^n\big)\cdot\omega_i,$$

$$G_{ij}=(\chi_i',\ \chi_j')_N.$$

The scheme (38) constitutes a system of nonlinear algebraic equations of the form

$$\mathbf{A}\left(\boldsymbol{\varphi}^{n+1},\ \boldsymbol{\varphi}^n,\ \boldsymbol{\varphi}^{n-1}\right)\boldsymbol{\varphi}^{n+1}=\mathbf{b}\left(\boldsymbol{\varphi}^{n+1},\ \boldsymbol{\varphi}^n,\ \boldsymbol{\varphi}^{n-1}\right),\quad n=1,\ 2,\ \ldots \tag{39}$$

with the unknown vector $\boldsymbol{\varphi}^{n+1}$, where

$$\boldsymbol{\varphi}^n=[\varphi_0^n,\ \ldots,\ \varphi_N^n]^T,\quad n=1,\ 2,\ \ldots\quad \left((\cdot)^T:\ \text{transpose}\ \right),$$

$$\mathbf{A}=\mathbf{P}^n+\Delta t\mathbf{Q}^n+\frac{\Delta t}{2}\mathbf{R}^n+\frac{1}{2}\left(\Delta t^2+\alpha\Delta t\right)\mathbf{G},$$

$$\mathbf{b}=\left(2\mathbf{P}^n+\Delta t\mathbf{Q}^n\right)\boldsymbol{\varphi}^n-\left(\mathbf{P}^n-\frac{\Delta t}{2}\mathbf{R}^n+\frac{1}{2}\left(\Delta t^2-\Delta t\boldsymbol{\alpha}\right)\mathbf{G}\right)\boldsymbol{\varphi}^{n-1}-\Delta t^2\mathbf{T}^n$$

(the 1st row of $\mathbf{A}$ is replaced with $[1,\ 0,\ \ldots,\ 0]$, and the 1st element of $\mathbf{b}$ with 0),

and

$$\mathbf{P}^n = [P_{ij}^n], \quad \mathbf{Q}^n = [Q_{ij}^n], \quad \mathbf{R}^n = [R_{ij}^n], \quad \mathbf{G} = [G_{ij}], \quad \mathbf{T}^n = [T_i^n].$$

Similarly, the initial equation (35) can be written as follows: define

$$P_{ij}^0 = \delta_{ij}\omega_j + \beta a_N(\chi_i \cos\phi^0,\ \chi_j \cos\phi^0) + \beta a_N(\chi_i \sin\phi^0,\ \chi_j \sin\phi^0),$$

$$Q_{ij}^0 = \beta a_N(\chi_i \sin\phi^0,\ \chi_j \cos\phi^0) - \beta a_N(\chi_i \cos\phi^0,\ \chi_j \sin\phi^0),$$

$$R_i^0 = \beta\gamma\left(B_{N,\,\phi}^0 \chi_i,\ B_{N,\,\phi}^0 \psi^0\right)_N,$$

$$T_i^0 = \beta\gamma U\left(B_{N,\,\phi}^0 \chi_i,\ \sin\phi^0\right)_N + P\left(-\cos\phi^0(1)\sin\phi_i^0 + \sin\phi^0(1)\cos\phi_i^0\right)\cdot\omega_i,$$

then, (35) is expressed as a system of $N+1$ linear equations:

$$\mathbf{A}^0\mathbf{w}_N = \mathbf{b}^0 \tag{40}$$

with the unknown vector $\mathbf{w}_N = [w_N(x_0)\ \ldots\ w_N(x_N)]^T$, where

$$\mathbf{A}^0 = \mathbf{P}^0, \quad \mathbf{b}^0 = -\mathbf{G}\left(\phi^0 + \alpha\psi^0\right) - \mathbf{Q}^0(\psi^0)^2 - \mathbf{R}^0 - \mathbf{T}^0,$$

(the 1st row of $\mathbf{A}^0$ is replaced with $[1,\ 0,\ \ldots,\ 0]$, and the 1st element of $\mathbf{b}^0$ with 0),

and

$$\mathbf{P}^0 = [P_{ij}^0], \quad \mathbf{Q}^0 = [Q_{ij}^0], \quad \mathbf{R}^0 = [R_i^0], \quad \mathbf{T}^0 = [T_i^0], \quad \mathbf{P}hi^0 = [\phi^0(x_i)], \quad \mathbf{P}si^0 = [\psi^0(x_i)], \quad (\mathbf{P}si^0)^2 = [\psi^0(x_i)^2].$$

Algorithm 1 Implementation of the scheme

procedure *Solve_IBVP*($\boldsymbol{\phi}^0,\ \boldsymbol{\psi}^0,\ \Delta t,\ N,\ final, tol$) ▷ *final*: final time step, *tol*: step tolerance for iteration
  $\boldsymbol{\varphi}^0 \leftarrow \boldsymbol{\phi}^0, \quad \boldsymbol{\varphi}^1 \leftarrow \boldsymbol{\phi}^0 + \Delta t\,\boldsymbol{\psi}^0 + \Delta t^2\mathbf{w}_N/2$ ▷ using the solution $\mathbf{w}_N$ of (40)
  for $n \leftarrow 1,\ 2,\ \ldots,$ final **do**
    $\mathbf{u} \leftarrow 2\boldsymbol{\varphi}^n - \boldsymbol{\varphi}^{n-1}$ ▷ initial guess of the next iteration
    repeat ▷ iteration to solve (39) for $\boldsymbol{\varphi}^{n+1}$
      $\mathbf{v} \leftarrow \mathbf{u}$
      $\hat{\mathbf{A}} \leftarrow \mathbf{A}\left(\mathbf{v},\ \boldsymbol{\varphi}^n,\ \boldsymbol{\varphi}^{n-1}\right), \quad \hat{\mathbf{b}} \leftarrow \mathbf{b}\left(\mathbf{v},\ \boldsymbol{\varphi}^n,\ \boldsymbol{\varphi}^{n-1}\right)$
      $\mathbf{u} \leftarrow \hat{\mathbf{A}}^{-1}\hat{\mathbf{b}}$

9: **until** $\|\mathbf{u}-\mathbf{v}\|_\infty/\|\mathbf{u}\|_\infty < tol$ ▷ $\|\mathbf{x}\|_\infty$: maximum norm of vector $\mathbf{x}$
10: $\boldsymbol{\varphi}^{n+1} \leftarrow \mathbf{u}$
11: **end for**
12: **return** $\boldsymbol{\varphi}^0,\ \boldsymbol{\varphi}^1,\ \boldsymbol{\varphi}^2,\ \ldots,\ \boldsymbol{\varphi}^{final}$ ▷ numerical solution of the IBVP
13. **end procedure**.

The scheme is implemented as shown in Algorithm 1. To calculate the matrix $\mathbf{A}$ and the vector $\mathbf{b}$ in line 7 of Algorithm 1, we need to compose the matrices whose $(i,\ j)$-elements are $a_N(f\chi_i,\ g\chi_i)$ and $\big(J_N(f\chi_i)u,\ J_N(g\chi_i)v\big)_N$, and the vector whose i-th element is $\big(J_N(f\chi_i)u,\ v\big)_N$ for continuous functions $f(x)$, $g(x)$, $u(x)$, $v(x)$, which appear in the definitions of $\mathbf{P}^n$, $\mathbf{Q}^n$, $\mathbf{R}^n$, and $\mathbf{T}^n$. In terms of the Legendre polynomial $L_k(x)$, the value of $[J_N f](x)$ is calculated exactly as

$$[J_N f](x) = \Big[l_0(x)\ \cdots\ l_N(x)\Big]\mathbf{L}^{-1}\begin{bmatrix} f(x_0) \\ \vdots \\ f(x_N)\end{bmatrix}\left(\mathbf{L} = \begin{bmatrix} L_0(x_0) & \cdots & L_N(x_0) \\ \vdots & \ddots & \vdots \\ L_0(x_N) & \cdots & L_N(x_N)\end{bmatrix}\right),$$

where

$$l_k(x) := \int_0^x L_k(\xi)\ \mathrm{d}\xi = \frac{1}{4k+2}\big(L_{k+1}(x) - L_{k-1}(x)\big), \quad k = 0,\ 1,\ \ldots \quad (\text{being } L_{\ -1}(x) = -1).$$

Define

$$\mathbf{J} = \begin{bmatrix} l_0(x_0) & \cdots & l_N(x_0) \\ \vdots & \ddots & \vdots \\ l_0(x_N) & \cdots & l_N(x_N)\end{bmatrix}\mathbf{L}^{-1}, \quad \boldsymbol{\Omega} = \begin{bmatrix} \omega_0 & 0 & \cdots & 0 \\ 0 & \omega_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \omega_N\end{bmatrix},$$

then, we can see that

$$\begin{bmatrix} a_N\big(f\chi_0,\ g\chi_0\big) & \cdots & a_N\big(f\chi_0,\ g\chi_N\big) \\ \vdots & \ddots & \vdots \\ a_N\big(f\chi_N,\ g\chi_0\big) & \cdots & a_N\big(f\chi_N,\ g\chi_N\big)\end{bmatrix} = \mathrm{diag}(f)\cdot\mathbf{J}^T\cdot\boldsymbol{\Omega}\cdot\mathbf{J}\cdot\mathrm{diag}(g),$$

$$\begin{bmatrix} \big(J_N(f\chi_0)u,\ J_N(g\chi_0)v\big)_N & \cdots & \big(J_N(f\chi_0)u,\ J_N(g\chi_N)v\big)_N \\ \vdots & \ddots & \vdots \\ \big(J_N(f\chi_N)u,\ J_N(g\chi_0)v\big)_N & \cdots & \big(J_N(f\chi_N)u,\ J_N(g\chi_N)v\big)_N\end{bmatrix} = \mathrm{diag}(f)\cdot(\mathrm{diag}(u)\cdot\mathbf{J})^T\cdot\boldsymbol{\Omega}\cdot\mathrm{diag}(v)\cdot\mathbf{J}\cdot\mathrm{diag}(g),$$

$$\begin{bmatrix} \big(J_N(f\chi_0)u,\ v\big)_N \\ \vdots \\ \big(J_N(f\chi_N)u,\ v\big)_N\end{bmatrix} = \mathrm{diag}(f)\cdot(\mathrm{diag}(u)\cdot\mathbf{J})^T\cdot\boldsymbol{\Omega}\cdot\begin{bmatrix} v(x_0) \\ \vdots \\ v(x_N)\end{bmatrix},$$

where diag(h) for a continuous function $h(x)$ denotes the diagonal matrix with the diagonal elements $[h(x_0),\ \dots\ h(x_N)]$. Thus, the matrix $\mathbf{P}^n$ and the vector $\mathbf{T}^n$, for example, are composed as

$$\mathbf{P}^n = \mathbf{\Omega} + \beta\left(\mathrm{diag}(C_\varphi^n)\cdot\mathbf{J}^T\cdot\mathbf{\Omega}\cdot\mathbf{J}\cdot\mathrm{diag}(C_\varphi^n) + \mathrm{diag}(S_\varphi^n)\cdot\mathbf{J}^T\cdot\mathbf{\Omega}\cdot\mathbf{J}\cdot\mathrm{diag}(S_\varphi^n)\right),\quad n = 1,\ 2,\ \dots,$$

$$\mathbf{T}^n = \beta\gamma U\left(\mathrm{diag}(S_\varphi^n)\cdot(\mathrm{diag}(\sin\boldsymbol{\varphi}^n)\cdot\mathbf{J})^T + \mathrm{diag}(C_\varphi^n)\cdot(\mathrm{diag}(\cos\boldsymbol{\varphi}^n)\cdot\mathbf{J})^T\right)\cdot\mathbf{\Omega}\cdot\sin\boldsymbol{\varphi}^n$$

$$+ P\left(-\cos\varphi_N^n\sin\boldsymbol{\varphi}^n + \sin\varphi_N^n\cos\boldsymbol{\varphi}^n\right)\cdot\mathbf{\Omega},\quad n = 1,\ 2,\ \dots.$$

Moreover, we see

$$\mathbf{G} = \left(\mathbf{D}\cdot\mathbf{L}^{-1}\right)^T\cdot\mathbf{\Omega}\cdot\mathbf{D}\cdot\mathbf{L}^{-1}\quad\left(\mathbf{D} = \begin{bmatrix} L_0'(x_0) & \cdots & L_N'(x_0)\\ \vdots & \ddots & \vdots\\ L_0'(x_N) & \cdots & L_N'(x_N)\end{bmatrix}\right),$$

where $L_k'(x)$ is the derivative of $L_k(x)$.

4.2 *Preliminaries*

Let $L^2_\omega(0,\ 1)$ denote the weighted L^2-space of the weight function $\omega(x)$, whose inner product and norm are

$$(f,\ g)_\omega = \int_0^1 f(x)g(x)\omega(x)\ \mathrm{d}x,\quad \|f\|_\omega = \sqrt{(f,\ f)_\omega}.$$

For every integer $m \geq 1$, we define the function space

$$B^m(0,\ 1) = \left\{f \mid f^{(k)} \in L^2_{\omega^k}(0,\ 1),\ 0 \leq k \leq m\right\},\quad \omega^k(x) = x^k(1-x)^k$$

($f^{(k)}$ stands for the k-th derivative of $f(x)$), which becomes a Hilbert space with the inner product and norm

$$(f,\ g)_{B^m} = \sum_{k=0}^m (f^{(k)},\ g^{(k)})_{\omega^k},\quad \|f\|_{B^m} = \sqrt{(f,\ f)_{B^m}},$$

and to which we introduce a semi-norm

$$|f|_{\omega,\ k} = \|f^{(k)}\|_{\omega^k},\quad k = 0,\ 1,\ \dots,\ m.$$

Moreover, let $H^m(0,\ 1)$ be the usual Sobolev space with the inner product and norm

$$(f,\, g)_{H^m} = \sum_{k=0}^{m} (f^{(k)},\, g^{(k)})_{L^2}, \quad \|f\|_{H^m} = \sqrt{(f,\, f)_{H^m}}.$$

We cite here from Shen et al. [27, sect. 3.5] and Canuto et al. [28, sect. 5.4] some known results in the spectral theory, which are employed in the next section, and restate them as the next proposition in a compact form.

For integers m and l, we define

$$\mu(l) = \min\{m,\, N+l\}, \quad \nu_N^{m,\,l} = \sqrt{\frac{(N+l-\mu(l))!}{N!}} N^{-\{l+\mu(l)\}/2}.$$

We can see that

$$\nu_N^{m,\,l+1} \le \nu_N^{m,\,l} \le N\nu_N^{m,\,l+1} \quad \text{for } m \ge l+1. \tag{41}$$

Proposition 1 Let $m \ge 1$ be an integer.

(i) Let $u \in B^m(0,\, 1)$ and $\psi \in \mathbb{P}_N$. If $m \ge 2$, we have

$$|(u,\, \psi)_{L^2} - (u,\, \psi)_N| \le C\nu_N^{m,\,0}\big(|u|_{\omega,\,\mu(0)} + |u|_{\omega,\,\mu(1)}\big)\ \|\psi\|_{L^2} \quad \big(\le C\nu_N^{m,\,0}\|u\|_{B^{\mu(1)}}\ \|\psi\|_{L^2}\big),$$

where C is a positive constant independent of m, N, u and ψ. Moreover, combining with (ii) and (v) below, we have

$$|(u,\, v)_{L^2} - (u,\, v)_N| \le C\nu_N^{m,\,0}\big(\|u\|_{B^{\mu(1)}}\ \|v\|_{H^1} + \|u\|_{H^1}\ \|v\|_{B^{\mu(1)}}\big), \quad u,\, v \in B^m(0,\, 1).$$

(ii) Let $u \in B^m(0,\, 1)$. If $m \ge 2$, we have

$$|I_N u|_{\omega,\,k} \le C|u|_{\omega,\,k}, \quad k = 2,\, 3,\, \ldots,\, m,$$

and

$$|u - I_N u|_{\omega,\,k} \le C\nu_N^{m,\,1} N^k |u|_{\omega,\,\mu(1)}, \quad k = 0,\, 1,\, \ldots,\, \mu(1),$$

where C is a positive constant independent of m, N, k and u.

(iii) Let $u \in H^m(0,\, 1)$. If $m \ge 2$, we have

$$\|u - I_N u\|_{H^1} \le C\nu_N^{m,\,1} N |u^{(1)}|_{\omega,\,\mu(1)-1},$$

where C is a positive constant independent of m, N and u.

(iv) Let π_N denote the orthogonal projection with respect to the Legendre polynomial, i.e.,

$$[\pi_N f](x) = \sum_{k=0}^{N} \frac{(f,\, L_k)_{L^2}}{\|L_k\|_{L^2}^2} L_k(x) \in \mathbb{P}_N, \quad f \in L^2(0,\, 1),$$

then, for $u \in B^m(0,\, 1)$, we have

$$\|u - \pi_{N-l} u\|_{L^2} \leq \sqrt{\frac{\big(N-l+1-\mu(1-l)\big)!}{\big(N-l+1+\mu(1-l)\big)!}}\, |u|_{\omega,\, \mu(1-l)} \leq \nu_N^{m,\, 1-l} |u|_{\omega,\, \mu(1-l)} \quad l = 0,\, 1,\, 2.$$

(v) For $u \in H^1(0,\, 1)$, we have

$$\|(u - I_N u)^{(1)}\|_{L^2} + N\|u - I_N u\|_{\omega^{-1}} \leq C\|u\|_{H^1},$$

and hence,

$$\|I_N u\|_{L^2} \leq C\big(\|u\|_{L^2} + N^{-1}\|u^{(1)}\|_{L^2}\big), \quad \|(I_N u)^{(1)}\|_{L^2} \leq C\|u\|_{H^1},$$

where C is a positive constant independent of N and u.

Define the operators J and J_N by

$$[Jf](x) = \int_0^x f(\xi)\ \mathrm{d}\xi, \quad [J_N f](x) = \int_0^x [I_N f](\xi)\ \mathrm{d}\xi,$$

and their transposes ${}^tJ$ and ${}^tJ_N$ by

$$\big(f,\, Jg\big)_{L^2} = \big({}^tJf,\, g\big)_{L^2},$$

$$\big(f,\, J_N g\big)_N = \big({}^tJ_N f,\, g\big)_N, \quad f,\, g \in C([0,\, 1]).$$

Lemma 1 It follows that, for $x \in [0,\, 1]$,

$$(i)\ \big|[J_N f](x)\big| \leq \|f\|_N, \quad f \in C([0,\, 1])$$

$$(ii)\ \big|[J_N f - Jf](x)\big| \leq C\nu_N^{m,\, 1}\|f\|_{B^{\mu(1)}}, \quad f \in C^m([0,\, 1]).$$

Proof. The results follows immediately from the definitions of the operators and Proposition 1 (ii). □

While omitting the details, it can be shown that ${}^tJ$ and ${}^tJ_N$ are expressed with the Legendre polynomials as

$$[{}^tJf](x)\ \left(=\int_x^1 f(\xi)\ \mathrm{d}\xi\right)=\frac{1}{2}\sum_{l=0}^{\infty}(f,\ L_{l+1}-L_{l-1})_{L^2}L_l(x) \tag{42}$$

$$[{}^tJ_Nf](x)=\frac{1}{2}\sum_{l=0}^{N-1}(f,\ L_{l+1}-L_{l-1})_NL_l(x)=\frac{1}{2}\left\{\sum_{l=0}^{N-2}(I_Nf,\ L_{l+1}-L_{l-1})_{L^2}L_l(x)+(f,\ L_N-L_{N-2})_NL_{N-1}(x)\right\}$$

$$=\int_x^1[I_Nf](\xi)\ \mathrm{d}\xi+\frac{1}{2}\left\{\frac{N+1}{2N+1}(f,\ L_N)_NL_{N-1}(x)+(f,\ L_{N-1})_NL_N(x)+\frac{N}{2N+1}(f,\ L_N)_NL_{N+1}(x)\right\} \tag{43}$$

for $f\in C([0,\ 1])$.

Lemma 2 It follows that, for $x\in[0,\ 1]$,

$$(i)\ \left|[{}^tJ_Nf](x)\right|\le 2\|f\|_N,\quad f\in C([0,\ 1])$$

$$(ii)\ \left|[{}^tJ_Nf-{}^tJf](x)\right|\le C\nu_N^{m,\ 0}N^{1/2}\|f\|_{B^{\mu(1)}},\quad f\in C^m([0,\ 1]).$$

Proof. Part (i) follows from (43) and (29).
For Part (ii), from (42) and (43), we see that

$$\left|[{}^tJ_Nf-{}^tJf](x)\right|$$

$$\le\int_x^1\left|[(I_N-I)f](\xi)\right|\ \mathrm{d}\xi+\frac{1}{2}\left(|(f,\ L_{N-1})_N|+|(f,\ L_N)_N|\right)$$

$$\le\|(I_N-I)f\|_{L^2}+\frac{1}{2}\left(|(f,\ L_{N-1})_{L^2}|+\left|\left((I_N-I)f,\ L_{N-1}\right)_{L^2}\right|\right)+\frac{1}{2}\left(2+\frac{1}{N}\right)\left(|(f,\ L_N)_{L^2}|+\left|\left((I_N-I)f,\ L_N\right)_{L^2}\right|\right)$$

$$\le 3\|(I_N-I)f\|_{L^2}+2\|L_{N-1}\|_{L^2}\ \|f-\pi_{N-2}f\|_{L^2},$$

then, the result follows from Proposition 1 (ii) and (iv), (41) and (29). □

The following three lemmas are given in [30].

Lemma 3 Let $f,\ g,\ h\in C([0,\ 1])$ and $N\ge 4$. It follows that

$$|a_N(fg,\ h)|\le\|f\|_N\|g\|_N\|h\|_N.$$

Lemma 4 Let $f,\ g\in C^m([0,\ 1])$, $\psi\in\mathbb{P}_N$ and $N\ge 4$. If $m\ge 2$, then we have

$$\text{(i)}\ \left|[a_N-a](f,\ g)\right|\le C\nu_N^{m,\ 0}\left(\|f\|_{B^{\mu(1)}}\|g\|_{H^1}+\|f\|_{H^1}\|g\|_{B^{\mu(1)}}\right),$$

(ii) $\left|[a_N - a](f,\ g\psi)\right| \le C\nu_N^{m,\,0}\left(\|f\|_{B^{\mu(1)}}\|g\|_{N,\,\infty} + \|g({}^tJJf)\|_{B^{\mu(1)}}\right)\|\psi\|_N,$

where C is a constant independent of m, N, f, g and ψ.

Lemma 5 Let $\{\theta^n\}$ be a temporal sequence, and $\phi = \phi(t)$ a smooth function of $t \ge 0$. Denote $\phi^n = \phi(t_n)$, and define $\varepsilon^n = \theta^n - \phi^n$. The following inequalities hold:

$$(i)\ (S_\theta^n)^2 + (C_\theta^n)^2 \le 1,$$

$$(ii)\ |\tilde{S}_\theta^n|,\ \ |\tilde{C}_\theta^n| \le 1,$$

$$(iii.a)\ |S_\theta^n - S_\phi^n|,\ \ |C_\theta^n - C_\phi^n| \le |\varepsilon^{n+1}| + |\varepsilon^{n-1}|,$$

$$(iii.b)\ |S_\theta^n - S_\theta^{n-1}|,\ \ |C_\theta^n - C_\theta^{n-1}| \le \Delta t\ \left(|\Delta_t^+\theta^n| + |\Delta_t^-\theta^{n-1}|\right),$$

$$(iv)\ |\tilde{S}_\theta^n - \tilde{S}_\phi^n|,\ \ |\tilde{C}_\theta^n - \tilde{C}_\phi^n| \le \frac{4}{3}\left(|\varepsilon^{n+1}| + |\varepsilon^n| + |\varepsilon^{n-1}|\right),$$

$$(v)\ |\tilde{S}_\phi^n - S_\phi^n|,\ \ |\tilde{C}_\phi^n - C_\phi^n| \le \Delta t^2\left((\Delta_t^+\phi^n)^2 + (\Delta_t^-\phi^n)^2 + |\Delta_t^2\phi^n|\right),$$

$$(vi)\ |S_\phi^n - \sin\phi^n|,\ \ |C_\phi^n - \cos\phi^n| \le \Delta t^2\left(|\xi_2^n(\phi)| + (\Delta_t^0\phi^n)^2|\right).$$

In the above, $\xi_k^n(f)\Delta t^k$ is the sum of the residual terms of Taylor series for $f(\cdot,\ t_{n+1})$ and $f(\cdot,\ t_{n-1})$ at $t = t_n$, i.e.,

$$[\xi_k^n(f)](x) = \frac{1}{k!}\left\{\frac{\partial^k f}{\partial t^k}(x,\ t_n^+) + \frac{\partial^k f}{\partial t^k}(x,\ t_n^-)\right\}\ (t_n^\pm \text{ is between } t_n \text{ and } t_{n\pm1}).$$

5. Convergence of the scheme

In this section, we prove our main theorem, which provides an error estimate of the scheme. The accuracy of the spectral method generally depends on the regularity of the exact solution. However, regarding the regularity of the exact solution ϕ, which requires certain appropriate compatibility conditions, no results are known so far, except for an existence result in [23, 24] for conservative problems. Thus, we assume the order of the regularity as a parameter m:

Hypothesis 1 Let $m \ge 4$ be an integer. IBVP (24)-(26) has a unique solution $\phi \in C^4([0,\ 1] \times [0,\ \infty))$ for which

$$\phi(\cdot,\ t),\ \ \phi_t(\cdot,\ t),\ \ \phi_{tt}(\cdot,\ t) \in C^m([0,\ 1]) \quad \text{for any } t \ge 0$$

is satisfied, and the total energy $\mathscr{E}(t)$ defined by (27) is uniformly bounded in $t \ge 0$.

Accordingly, we also make an assumption on the discrete solution:

Hypothesis 2 For the solution φ^n of the scheme (32)-(34), the discrete total energy $\mathscr{E}^n$ defined by (36) is uniformly bounded in n, i.e., there exists an $M > 0$ such that $\mathscr{E}^n \le M$, $n = 1,\ 2,\ \ldots$.

Under Hypothesises 1 and 2, this paper focuses attention on estimating the error of the scheme. For the exact solution ϕ of the original problem and the discrete solution φ^n of the scheme, we define the errors of the scheme by

$$\varepsilon^n = \varphi^n - \phi^n, \qquad e^n = \varphi^n - I_N\phi^n\Big(\phi^n = \phi(\cdot,\, t_n)\Big).$$

In what follows, $C \geq 1$ is a constant independent of m, N, n, Δt and the functions. Moreover, we denote by $\zeta^n(\varphi)$ a polynomial of one or more terms in

$$\|\Delta_t^+\varphi^n\|_N, \quad \|\Delta_t^+\varphi^{n-1}\|_N, \quad \|\Delta_t^+\varphi^{n-2}\|_N, \quad \|\varphi_x^{n+1}\|_N, \quad \|\varphi_x^{n-1}\|_N,$$

and, by $\eta^n(\phi;\, m)$, that in

$$\|\cdot\|_{N,\,\infty}\text{-}, \ \|\cdot\|_N\text{-, and } \|\cdot\|_{B^{\mu(1)}}\text{-norms of functions comprised of } \phi^n,\ \phi_t^n,\ \phi_{tt}^n,$$

$$\|\xi_1^n(\phi)^2\|_N, \ \|\xi_2^n(\phi)\|_N, \ \|\xi_4^n(\cos\phi)\|_N, \ \|\xi_4^n(\sin\phi)\|_N$$

($\xi_k^n(f)$ is defined in Lemma 5).

5.1 *Error estimate*

By subtracting (24) from (32), we obtain the error equation

$$(\Delta_t^2 e^n,\ \psi)_N + \big(\bar{e}_x^n + \alpha(\Delta_t^0 e^n)_x,\ \psi_x\big)_{L^2} = \Theta_1 + \beta\big(\Theta_2 + \Theta_3 - \gamma U\Theta_4 - \gamma\Theta_5\big) - P\Theta_6, \qquad \psi \in \mathscr{X}_N, \tag{44}$$

where

$$\Theta_1 = (\phi_{tt}^n,\ \psi)_{L^2} - (\Delta_t^2\phi^n,\ \psi)_N + \big((\phi^n - I_N\bar{\phi}^n)_x,\ \psi_x\big)_{L^2} + \alpha\big((\phi_t^n - I_N\Delta_t^0\phi^n)_x,\ \psi_x\big)_{L^2},$$

$$\Theta_2 = a_N\big(\Delta_t^2\cos\varphi^n,\ S_\varphi^n\psi\big) - a_N\big((\cos\phi^n)_{tt},\ \psi\sin\phi^n\big) - a_N\big(\Delta_t^2\sin\varphi^n,\ C_\varphi^n\psi\big) + a_N\big((\sin\phi^n)_{tt},\ \psi\cos\phi^n\big),$$

$$\Theta_3 = [a_N - a]\big((\cos\phi^n)_{tt},\ \psi\sin\phi^n\big) - [a_N - a]\big((\sin\phi^n)_{tt},\ \psi\cos\phi^n\big),$$

$$\Theta_4 = \big(\sin\varphi^n,\ B_{N,\,\varphi}^n\psi\big)_N - \big(\sin\phi^n,\ B_\phi^n\psi\big)_{L^2},$$

$$\Theta_5 = \big(B_{N,\,\varphi}^n(\Delta_t^0\varphi^n),\ B_{N,\,\varphi}^n\psi\big)_N - \big(B_\phi^n\phi_t^n,\ B_\phi^n\psi\big)_{L^2},$$

$$\Theta_6 = b_{N,\,\varphi}^n(\psi) - b_\phi^n(\psi),$$

being $\phi_{tt}^n = \phi_{tt}(\cdot,\, t_n)$, $(\cos\phi^n)_{tt} = [(\cos\phi)_{tt}](\cdot,\, t_n)$ and so on.

Define the transposes ${}^tB^n_\phi$ and ${}^tB^n_{N,\,\varphi}$ in the similar way to ${}^tJ$ and ${}^tJ_N$, respectively.

Lemma 6 It follows that

$$\left\|B^n_{N,\,\varphi}f\right\|_N,\ \left\|{}^tB^n_{N,\,\varphi}f\right\|_N \le \sqrt{2}\ \|f\|_N,$$

$$\left\|B^n_{\phi}f\right\|_{L^2},\ \left\|{}^tB^n_{\phi}f\right\|_{L^2} \le \sqrt{2}\ \|f\|_{L^2}, \quad f\in C([0,\,1]).$$

Proof. It is easy show that

$$\|Jf\|_{L^2},\ \|{}^tJf|_{L^2} \le 2^{-1/2}\ \|f\|_{L^2}, \qquad \|J_Nf|_N,\ \|{}^tJ_Nf|_N \le 2^{-1/2}\ \|f\|_N, \tag{45}$$

from which the results follow immediately. □

Lemma 7 Let f be a function which is comprised of ϕ^n, such as $\sin\phi^n$, $B^n_\phi\Delta^0_t\phi^t$. Then, we have

$$\left\|\left(B^n_{N,\,\varphi}-B^n_\phi\right)f\right\|_N \le C\left(1+\eta^n(\phi;\,m)\right)\left(\|e^{n+1}\|_N+\|e^n\|_N+\|e^{n-1}\|_N++\nu_N^{m,\,1}+\Delta t^2\right),$$

$$\left\|\left({}^tB^n_{N,\,\varphi}-{}^tB^n_\phi\right)f\right\|_N \le C\left(1+\eta^n(\phi;\,m)\right)\left(\|e^{n+1}\|_N+\|e^n\|_N+\|e^{n-1}\|_N++\nu_N^{m,\,0}N^{1/2}+\Delta t^2\right).$$

Proof. We see that

$$\begin{aligned}\left\|\left(B^n_{N,\,\varphi}-B^n_\phi\right)f\right\|_N \le{}& \left\|J_N(fS^n_\varphi)(\sin\varphi^n-\sin\phi^n)\right\|_N+\left\|J_N\left(f(S^n_\varphi-S^n_\phi)\right)\right\|_N\\ &+\left\|J_N\left(f(S^n_\phi-\sin\phi^n)\right)\right\|_N+\left\|(J_N-J)(f\sin\phi^n)\right\|_N+\text{etc.},\end{aligned}$$

which, together with Lemmas 1 and 5, leads to the former inequality. Similarly,

$$\begin{aligned}\left\|\left({}^tB^n_{N,\,\varphi}-{}^tB^n_\phi\right)f\right\|_N \le{}& \left\|{}^tJ_N(f\sin\varphi^n)\left(S^n_\varphi-S^n_\phi\right)\right\|_N+\left\|{}^tJ_N(f\sin\varphi^n)\left(S^n_\phi-\sin\phi^n\right)\right\|_N\\ &+\left\|{}^tJ_N\left(f(\sin\varphi^n-\sin\phi^n)\right)\right\|_N+\left\|({}^tJ_N-{}^tJ)(f\sin\phi^n)\right\|_N+\text{etc.},\end{aligned}$$

together with Lemmas 2 and 5, leads to the latter one. □

Lemma 8 We have

$$|\Theta_1| \le (1+\alpha)C\eta^n(\phi;\,m)\left(\nu_N^{m,\,1}N+\Delta t^2\right)\left(\|\psi\|_{L^2}+\|\psi_x\|_{L^2}\right),$$

$$\Theta_2 = -a_N\left((\Delta^2_te^n)S^n_\varphi,\ \ S^n_\varphi\psi\right)-a_N\left((\Delta^2_te^n)C^n_\varphi,\ \ C^n_\varphi\psi\right)+\hat{\Theta}_2,$$

$$|\Theta_3| \le C\eta^n(\phi;\, m)\ \nu_N^{m,\,0}\ \|\psi\|_{L^2}$$

with the residual term $\hat{\Theta}_2$ estimated by

$$|\hat{\Theta}_2| \le C\,(1+\eta^n(\phi;\, m))\left(1+\|\Delta_t^- \varphi^n\|_N\right)\left(\|\Delta_t^+ e^n\|_N + \|\Delta_t^- e^n\|_N + \|e^{n+1}\|_N + \|e^n\|_N + \|e^{n-1}\|_N + \Delta t^2\right)\ \|\psi\|_N. \tag{46}$$

Proof. See Lemmas 5.1 and 5.2 in [22] for the results on Θ_1 and Θ_2. The estimate for Θ_3 follows from Lemma 4 (ii).□

Lemma 9 We have

$$|\Theta_4| \le C\big(1+\eta^n(\phi;\, m)\big)\Big(\|e^{n+1}\|_N + \|e^n\|_N + \|e^{n-1}\|_N + \nu_N^{m,\,0} N^{1/2} + \Delta t^2\Big)\|\psi\|_N.$$

Proof. We see that

$$|\Theta_4| \le \|\sin\varphi^n - \sin\phi^n\|_N\ \|B_{N,\,\varphi}^n \psi\|_N + \left\|{}^t B_{N,\,\varphi}^n \sin\phi^n - {}^t B_\phi^n \sin\phi^n\right\|_N\ \|\psi\|_N + \left|\left({}^t B_\phi^n \sin\phi^n,\ \psi\right)_N - \left({}^t B_\phi^n \sin\phi^n,\ \psi\right)_{L^2}\right|,$$

then, the result follows from Lemmas 6, 7 and Proposition 1 (i). □

Lemma 10

$$\Theta_5 = \left(B_{N,\,\varphi}^n(\Delta_t^0 e^n),\, B_{N,\,\varphi}^n \psi\right)_N + \hat{\Theta}_5$$

with

$$|\hat{\Theta}_5| \le C\big(1+\eta^n(\phi;\, m)\big)\Big(\|e^{n+1}\|_N + \|e^n\|_N + \|e^{n-1}\|_N + + \nu_N^{m,\,0} N^{1/2} + \Delta t^2\Big)\|\psi\|_N.$$

Proof. We see that

$$\begin{aligned}
|\hat{\Theta}_5| &= \left|\left(B_{N,\,\varphi}^n(\Delta_t^0\phi^n),\, B_{N,\,\varphi}^n\psi\right)_N - \left(B_\phi^n \phi_t^n,\, B_\phi^n \psi\right)_{L^2}\right| \\
&\le \left\|\left(B_{N,\,\varphi}^n - B_\phi^n\right)\Delta_t^0\phi^n\right\|_N\ \|B_{N,\,\varphi}^n\psi\|_N + \left\|\left({}^t B_{N,\,\varphi}^n - {}^t B_\phi^n\right) B_\phi^n \Delta_t^0 \phi^n\right\|_N\ \|\psi\|_N \\
&\quad + \left|\left({}^t B_\phi^n B_\phi^n \Delta_t^0\phi^n,\ \psi\right)_N - \left({}^t B_\phi^n B_\phi^n \Delta_t^0\phi^n,\ \psi\right)_{L_2}\right| + \left\|B_\phi^n\left(\Delta_t^0\phi^n - \phi_t^n\right)\right\|_{L^2}\ \|B_\phi^n\psi\|_{L^2},
\end{aligned}$$

then, the result follows from Lemmas 6, 7 and Proposition 1 (i). □

Lemma 11 We have

$$|\Theta_6| \le C\big(1+\eta^n(\phi;\, m)\big)\Big(\|e^{n+1}\|_N + \|e^{n-1}\|_N + |e^n(1)| + \nu_N^{m,\,0} + \Delta t^2\Big)\|\psi\|_N.$$

Proof. We see that

$$\begin{aligned}
|\Theta_6| &\le \big|\big(\psi,\, S_\varphi^n\big)_N \cos\varphi^n(1) - \big(\psi,\, \sin\phi^n\big)_{L^2}\cos\phi^n(1)\big| + \big|\big(\psi,\, C_\varphi^n\big)_N \sin\varphi^n(1) - \big(\psi,\, \cos\phi^n\big)_{L^2}\sin\phi^n(1)\big| \\
&\le \|\psi\|_N\|S_\varphi^n - S_\phi^n\|_N + \|\psi\|_N\|S_\phi^n - \sin\phi^n\|_N + \big|\big(\psi,\, \sin\phi^n\big)_N - \big(\psi,\, \sin\phi^n\big)_{L^2}\big| + \|\psi\|_{L^2}\big|\cos\varphi^n(1) - \cos\phi^n(1)\big| \\
&\quad + \|\psi\|_N\|C_\varphi^n - C_\phi^n\|_N + \|\psi\|_N\|C_\phi^n - \cos\phi^n\|_N + \big|\big(\psi,\, \cos\phi^n\big)_N - \big(\psi,\, \cos\phi^n\big)_{L^2}\big| + \|\psi\|_{L^2}\big|\sin\varphi^n(1) - \sin\phi^n(1)\big|,
\end{aligned}$$

then, the result follows from Lemmas 5 and Proposition 1 (i). □

We now give our main result on the error estimate of the scheme:

Theorem 1 Let $m \ge 4$ be an integer described in Hypothesis 1. If Hypothesis 1 and 2 is satisfied, then there exists a $\Delta t_0 \in (0,\, N^{-1}]$ such that for $0 < \Delta t \le \Delta t_0$,

$$\|\phi^n - \varphi^n\|_N + \|\Delta_t^+(\phi^n - \varphi^n)\|_N + \|(I_N\phi^n - \varphi^n)_x\|_{L^2} \le C_n\big(\Delta t^2 + \nu_N^{m,\,1}N^3\big), \quad n = 1,\, 2,\, \ldots,$$

where $C_n \ge 1$ is a constant that depends on ϕ, ϕ^0, ψ^0 and n, and is independent of N and Δt.

Proof. We take N and Δt so that $N\Delta t \le 1$.

We first set $\psi = e^{n+1} - e^{n-1}$ in the error Equation (44). Then, the left hand side of the equation becomes

$$\|\Delta_t^+ e^n\|_N^2 - \|\Delta_t^+ e^{n-1}\|_N^2 + \frac{1}{2}\big(\|e_x^{n+1}\|_{L^2}^2 + \|e_x^n\|_{L^2}^2\big) - \frac{1}{2}\big(\|e_x^n\|_{L^2}^2 + \|e_x^{n-1}\|_{L^2}^2\big) + 2\alpha\Delta t\|(\Delta_t^0 e^n)_x\|_{L^2}^2.$$

Moreover, Θ_2 can be decomposed as

$$\Theta_2 = -a_N\big(S_\varphi^n\Delta_t^+ e^n\big) + a_N\big(S_\varphi^{n-1}\Delta_t^+ e^{n-1}\big) + \Theta_{2,\,1} - a_N\big(C_\varphi^n\Delta_t^+ e^n\big) + a_N\big(C_\varphi^{n-1}\Delta_t^+ e^{n-1}\big) + \Theta_{2,\,2} + \hat{\Theta}_2,$$

where $\hat{\Theta}_2$ are that in Lemma 8, and

$$\Theta_{2,\,1} = a_N\big((S_\varphi^n + S_\varphi^{n-1})\Delta_t^- e^n,\, (S_\varphi^n - S_\varphi^{n-1})\Delta_t^- e^n\big),$$

$$\Theta_{2,\,2} = a_N\big((C_\varphi^n + C_\varphi^{n-1})\Delta_t^- e^n,\, (C_\varphi^n - C_\varphi^{n-1})\Delta_t^- e^n\big).$$

Thus, defining E^n by

$$E^n = \|\Delta_t^+ e^n\|_N^2 + \frac{1}{2}\big(\|e_x^{n+1}\|_{L^2}^2 + \|e_x^n\|_{L^2}^2\big) + \beta a_N\big(S_\varphi^n \Delta_t^+ e^n\big) + \beta a_N\big(C_\varphi^n \Delta_t^+ e^n\big),$$

we see that (44) is equivalent to

$$E^n - E^{n-1} + 2\alpha\Delta t\|(\Delta_t^0 e^n)_x\|_{L^2}^2 = \Theta_1 + \beta\big(\Theta_{2,\,1} + \Theta_{2,\,2} + \hat{\Theta}_2 + \Theta_3 - \gamma U\Theta_4 - 2\gamma\Delta t\|B_{N,\,\varphi}^n(\Delta_t^0 e^n)\|_N^2 - \gamma\hat{\Theta}_5\big) - P\Theta_6. \quad (47)$$

From Lemmas 3 and 5 (i) and (iii.b), we see that

$$|\Theta_{2,\,1}|,\ \ |\Theta_{2,\,2}| \le C\Delta t\ \ \|\Delta_t^- e^n\|_N^2\big(\|\Delta_t^+\varphi^n\|_N + \|\Delta_t^-\varphi^{n-1}\|_N\big) \le C\Delta t\ \ \zeta^n(\varphi)E^{n-1}.$$

Note that for $\psi = e^{n+1} - e^{n-1}$,

$$\|\psi\|_N \le \Delta t\big(\sqrt{E^n} + \sqrt{E^{n-1}}\big), \qquad \|\psi_x\|_{L^2} \le \sqrt{2}\big(\sqrt{E^n} + \sqrt{E^{n-1}}\big).$$

From these inequalities, Lemmas 8-11 and (41), the right hand side of (47) is estimated as follows:

$$\big\{|\Theta_1| + \beta\big(|\Theta_{2,\,1}| + |\Theta_{2,\,2}| + |\hat{\Theta}_2| + |\Theta_3| + \gamma U|\Theta_4| + \gamma|\hat{\Theta}_5|\big) + P|\Theta_6|\big\} \Big/ \big(\sqrt{E^n} + \sqrt{E^{n-1}}\big)$$

$$\le C\big(1 + \eta^n(\phi;\, m)\big)\big(1 + \zeta^n(\varphi)\big)$$

$$\times\Big\{\Delta t\,\Big(\sqrt{E^n} + \sqrt{E^{n-1}} + \|e^{n+1}\|_N + \|e^n\|_N + \|e^{n-1}\|_N + |e^n(1)|\Big) + \nu_N^{m,\,1}N + \Delta t^2\Big\}. \quad (48)$$

Letting

$$\Phi^n = \sqrt{E^n} + \|e^{n+1}\|_N + \|e^n\|_N,$$

it follows from (47) and (48) that

$$\sqrt{E^n} - \sqrt{E^{n-1}} \le C\big(1 + \eta^n(\phi;\, m)\big)\big(1 + \zeta^n(\varphi)\big)\Big\{\Delta t\ \ \big(\Phi^n + \Phi^{n-1} + |e^n(1)|\big) + \nu_N^{m,\,1}N + \Delta t^2\Big\}. \quad (49)$$

We next set $\psi = \Delta t^2 e^{n+1}$ in the error equation (44). Then, we see that

$$\|e^{n+1}\|_N^2 + \frac{\Delta t^2 + \alpha \Delta t}{2} \|e_x^{n+1}\|_{L^2}^2$$

$$\leq \left|(2e^n - e^{n-1},\ e^{n+1})_N\right| + \frac{\Delta t^2 + \alpha \Delta t}{2} \left|(e_x^{n-1},\ e_x^{n+1})_{L^2}\right| + |\Theta_1| + \beta\left(|\Theta_2| + |\Theta_3| + \gamma U|\Theta_4| + \gamma|\Theta_5|\right) + P|\Theta_6|$$

$$\leq C\big(1 + \eta^n(\phi;\, m)\big)\big(1 + \zeta^n(\varphi)\big)$$

$$\times \left[\Delta t\sqrt{E^n} + \sqrt{E^{n-1}} + \|e^n\|_N + \Delta t^2\left(\|e^{n+1}\|_N + \|e^{n-1}\|_N + |e^n(1)|\right) + \nu_N^{m,\,1} N\Delta t + \Delta t^3\right]\left(\|e^{n+1}\|_N + \Delta t\|e_x^{n+1}\|_{L^2}\right),$$

which implies

$$\|e^{n+1}\|_N - \|e^{n-1}\|_N \leq \|e^{n+1}\|_N \left(+\Delta t\|e_x^{n+1}\|_{L^2}\right)$$

$$\leq C\big(1 + \eta^n(\phi;\, m)\big)\big(1 + \zeta^n(\varphi)\big)\left[\Delta t\ \Phi^n + \Phi^{n-1} + \Delta t^2|e^n(1)| + \nu_N^{m,\,1} N\Delta t + \Delta t^3\right]. \tag{50}$$

Adding both sides of (49) and (50), and noting that

$$|e^n(1)| \leq CN\|e^n\|_N, \quad N\Delta t \leq 1$$

(the former follows from the inverse inequality [28, sect. 5.4.1]), we have

$$\Phi^n - \Phi^{n-1} \leq C\sigma^n\left[\Delta t\ \Phi^n + \Phi^{n-1} + \nu_N^{m,\,1} N + \Delta t^2\right],$$

where

$$\sigma^n = \big(1 + \eta^n(\phi;\, m)\big)\big(1 + \zeta^n(\varphi)\big).$$

Recall now that Hypothesises 1 and 2 state the uniform boundedness of the energies $\mathscr{E}(t)$ in t and $\mathscr{E}^n$ in n. These imply that the temporal sequence $\eta^n(\phi;\, m)$ and $\zeta^n(\varphi)$ are uniformly bounded in n. Thus, we can take a Δt such that

$$C\sigma^n\Delta t \leq 1/2, \quad n = 1,\ 2,\ \ldots.$$

Then, we have

$$\Phi^n - \Phi^{n-1} \le C\sigma^n \left(\Phi^{n-1} + \nu_N^{m,\,1} N + \Delta t^2 \right). \tag{51}$$

Hence, by Gronwall's lemma,

$$\Phi^n \le \left\{ \Phi^0 + C \left(\nu_N^{m,\,1} N + \Delta t^2 \right) \sum_{k=1}^{n} \sigma^k \right\} \exp\left(C \sum_{k=1}^{n} \sigma^k \right),$$

which, together with the estimate of Φ^0 given by Theorem 2 below, leads to the results. □

Theorem 2 Let $\phi^0,\ \psi^0 \in C^m([0,\ 1])$ for some $m \ge 2,$ and Δt be as in Theorem 1. The initial data φ^0 and φ^1 defined by (33) and (34) satisfy

$$\|e^1\|_N + \|\Delta_t^+ e^0\|_N + \|e_x^1\|_{L^2} \le C_0 \left(\Delta t^2 + \nu_N^{m,\,1} N^3 \right),$$

where $C_0 \ge 1$ is a constant that is independent of N and $\Delta t,$ and depends on $\phi^0,\ \psi^0$ and $\phi(\cdot,\ t)$ for $0 \le t \le \Delta t$.

Proof. Denote $w^0(x) = \phi_{tt}(x,\ 0)$. The Equation (24) of the original problem at $t = 0$ can be written as

$$\begin{aligned}
&(w^0,\ \phi)_{L^2} + \beta a\left(w^0 \sin\phi^0,\ \psi \sin\phi^0\right) + \beta a\left(w^0 \cos\phi^0,\ \psi \cos\phi^0\right) \\
&= -(\phi_x^0 + \alpha\psi_x^0,\ \psi_x)_{L^2} - \beta a\left((\psi^0)^2 \cos\phi^0,\ \psi \sin\phi^0\right) + \beta a\left((\psi^0)^2 \sin\phi^0,\ \psi \cos\phi^0\right) \\
&\quad - \beta\gamma\left(U \sin\phi^0 + B_\phi^0 \psi^0,\ B_\phi^0 \psi\right)_{L^2} - P b_\phi^0(\psi), \qquad \psi \in \mathscr{X}_N,
\end{aligned}$$

where $B_\phi^0 = B_{\phi^0}$ and $b_\phi^0 = b_{\phi^0}$. Subtracting this from (35) and setting $e = w_N - I_N w^0$, we obtain

$$\begin{aligned}
&(e,\ \psi)_N + \beta a_N\left(e \sin\phi^0,\ \psi \sin\phi^0\right) + \beta a_N\left(e \cos\phi^0,\ \psi \cos\phi^0\right) \\
&= R_1 + R_2 + \beta\left(R_3 + \gamma U R_4 + \gamma R_5\right) + P R_6,
\end{aligned} \tag{52}$$

where

$$\begin{aligned}
R_1 &= (w^0,\ \psi)_{L^2} - (w^0,\ \psi)_N, \\
R_2 &= \left((\phi^0 + \alpha\psi^0)_x,\ \psi_x\right)_{L^2} - \left((I_N(\phi^0 + \alpha\psi^0))_x,\ \psi_x\right)_N, \\
R_3 &= [a - a_N]\left(f^0,\ \psi \sin\phi^0\right) + [a - a_N]\left(g^0,\ \psi \cos\phi^0\right),
\end{aligned}$$

$$R_4 = \left(\sin\phi^0,\ B^0_\phi\psi\right)_{L^2} - \left(\sin\phi^0,\ B^0_{N,\,\phi}\psi\right)_N,$$

$$R_5 = \left(B^0_\phi\psi^0,\ B^0_\phi\psi\right)_{L^2} - \left(B^0_{N,\,\phi}\psi^0,\ B^0_{N,\,\phi}\psi\right)_N,$$

$$R_6 = b^0_\phi(\psi) - b^0_{N,\,\phi}(\psi),$$

being

$$f^0 = w^0\sin\phi^0 + (\psi^0)^2\cos\phi^0, \quad g^0 = w^0\cos\phi^0 - (\psi^0)^2\sin\phi^0.$$

The terms in the right-hand side of (52) are estimated as follows. By Proposition 1 (i),

$$|R_1| \le C\eta^0(\phi;\ m)\nu_N^{m,\,0}\|\psi\|_{L^2},$$

where $\eta^0(\phi;\ m)$ is a polynomial of $\|\cdot\|_{B^{\mu(1)}}$-norms of those functions that are comprised of ϕ^0, ψ^0 and w^0. Since $\left(I_N(\phi^0+\alpha\psi^0)\right)_x\cdot\psi_x \in \mathbb{P}_{2N-2}$, it follows from Proposition 1 (ii) and the inverse inequality

$$\|\psi_x\|_{L^2} \le (N+1)(N+2)\|\psi\|_{L^2}, \quad \psi\in\mathbb{P}_N,$$

that

$$|R_2| = \left|\left(\left(\phi^0+\alpha\psi^0\right)_x - \left(I_N(\phi^0+\alpha\psi^0)\right)_x,\ \psi_x\right)_{L^2}\right|$$

$$\le \left\|\left(\phi^0+\alpha\psi^0 - I_N(\phi^0+\alpha\psi^0)\right)_x\right\|_{L^2}\ \|\psi_x\|_{L^2} \le C\eta^0(\phi;\ m)\ \nu_N^{m,\,1}N\cdot N^2\|\psi\|_{L^2}.$$

By Lemma 4 (i),

$$|R_3| \le C\eta^0(\phi;\ m)\ \nu_N^{m,\,0}\ \|\psi\|_{L^2}.$$

On the right-hand side of

$$|R_4| \le \left|\left({}^tB^0_\phi\sin\phi^0,\ \psi\right)_{L^2} - \left({}^tB^0_\phi\sin\phi^0,\ \psi\right)_N\right| + \left|\left({}^tB^0_\phi\sin\phi^0 - {}^tB^0_{N,\,\phi}\sin\phi^0,\ \psi\right)_N\right|,$$

by Proposition 1 (i) and Proposition 2 (ii), we see that

(1st term) $\le C\eta^0(\phi;\, m)\ \nu_N^{m,\,0}\ \|\psi\|_{L^2}$,

(2nd term) $\le \left\|({}^tJ - {}^tJ_N)(\sin^2\phi^0)\right\|_N\ \|\psi\|_N + \left\|({}^tJ - {}^tJ_N)(\sin\phi^0\cos\phi^0)\right\|_N\ \|\psi\|_N \le C\eta^0(\phi;\, m)\ \nu_N^{m,\,0} N^{1/2}\ \|\psi\|_N$.

Similarly, by Propositions 1 (i), 2 (ii) and Lemma 1 (ii), we have

$$|R_5| \le \left|\left({}^tB_\phi^0 B_\phi^0\psi^0,\ \psi\right)_{L^2} - \left({}^tB_\phi^0 B_\phi^0\psi^0,\ \psi\right)_N\right| + \left|\left({}^tB_\phi^0 B_\phi^0\psi^0 - {}^tB_{N,\,\phi}^0 B_\phi^0\psi^0,\ \psi\right)_N\right| + \left|\left(B_\phi^0\psi^0 - B_{N,\,\phi}^0\psi^0,\ B_{N,\,\phi}^0\psi\right)_N\right|$$

$$\le C\eta^0(\phi;\, m)\ \nu_N^{m,\,0}\ \|\psi\|_{L^2} + C\eta^0(\phi;\, m)\ \nu_N^{m,\,0} N^{1/2}\ \|\psi\|_{L^2} + C\eta^0(\phi;\, m)\ \nu_N^{m,\,1}\ \|\psi\|_{L^2}.$$

By Proposition 1 (i),

$$|R_6| \le \left|\left(\psi,\ \sin\phi^0\right)_{L^2} - \left(\psi,\ \sin\phi^0\right)_N\right| + \left|\left(\psi,\ \cos\phi^0\right)_{L^2} - \left(\psi,\ \cos\phi^0\right)_N\right| \le C\eta^0(\phi;\, m)\ \nu_N^{m,\,0}\ \|\psi\|_N.$$

The estimations for $|R_i|$ above and (41) lead to

$$(e,\ \psi)_N + \beta a_N\left(e\sin\phi^0,\ \psi\sin\phi^0\right) + \beta a_N\left(e\cos\phi^0,\ \psi\cos\phi^0\right) \le C\eta^0(\phi;\, m)\ \nu_N^{m,\,1} N^3\ \|\psi\|_N.$$

We now set $\psi = e$ to obtain

$$\|w_N - w^0\|_N = \|e\|_N \le C\eta^0(\phi;\, m)\ \nu_N^{m,\,1} N^3.$$

Therefore, by the Definition (34) of φ^1 and Taylor's theorem, we obtain

$$\|e^1\|_N = \|\varphi^1 - \phi^1\|_N \le \frac{\Delta t^2}{2}\|w_N - w^0\|_N + \frac{\Delta t^3}{3!}\|\phi_{ttt}(\cdot,\, t_0^+)\|_N$$

$$\le C\left(\eta^0(\phi;\, m) + \|\phi_{ttt}(\cdot,\, t_0^+)\|_N\right)\left(\nu_N^{m,\,1} N^3\Delta t^2 + \Delta t^3\right)\ \ (0 \le t_0^+ \le \Delta t).$$

By the Definition (33) of φ^0, we see $e^0 = 0$, and thus, $\|\Delta_t^+ e^0\| = \|e^1\|/\Delta t$.
Moreover, the same argument shows that

$$\|e_x^1\|_N = \|\varphi_x^1 - (I_N\phi^1)_x\|_N \le \frac{\Delta t^2}{2}\left\|(w_N - I_N w^0)_x\right\|_N + \frac{\Delta t^3}{3!}\|\phi_{xttt}(\cdot,\, t_0^+)\|_N$$

$$\le C\left(\eta^0(\phi;\, m) + \|\phi_{xttt}(\cdot,\, t_0^+)\|_N\right)\left(\nu_N^{m,\,1} N^5\Delta t^2 + \Delta t^3\right)\quad (0 \le t_0^+ \le \Delta t).$$

Finally, we note that $N\Delta t \leq 1$. □

6. Numerical simulations

In this section, we present the results of numerical simulations conducted using the scheme described in (39), (33), (34).

At each time step, the algebraic Equation (39) for $\boldsymbol{\varphi}^{n+1}$ was solved iteratively using the update formula

$$\mathbf{A}(\boldsymbol{\varphi}^{(\nu)}, \boldsymbol{\varphi}^{n}, \boldsymbol{\varphi}^{n-1})\boldsymbol{\varphi}^{(\nu+1)} = \mathbf{b}(\boldsymbol{\varphi}^{(\nu)}, \boldsymbol{\varphi}^{n}, \boldsymbol{\varphi}^{n-1}) \tag{53}$$

with the termination criteria

$$\frac{||\boldsymbol{\varphi}^{(\nu+1)} - \boldsymbol{\varphi}^{(\nu)}||_{\infty}}{||\boldsymbol{\varphi}^{(\nu+1)}||_{\infty}} < tol \quad \text{or} \quad \nu > maxiter,$$

where $\|\boldsymbol{\varphi}\|_{\infty} = \max_j |\varphi_j|$ for a vector $\boldsymbol{\varphi}$ whose j-th element is φ_j. Unless otherwise specified, the tolerances were set as

$$tol = 4.0 \times 10^{-10}, \quad maxiter = 500.$$

The grid size and time spacing were fixed, respectively, as follows:

$$N = 100, \quad \Delta t = 1/1,000.$$

The initial displacement and velocity in (26) were set as

$$\phi^0(x) = 0, \quad \psi^0(x) = V^0 \times \frac{\arctan(10x)}{\pi/2},$$

where the velocity parameter V^0 was fixed to $V^0 = 1 \times 10^{-3}$ unless stated otherwise.

Before addressing the non-conservative problem, we verified the validity of the scheme implementation by monitoring the fluctuation of an invariant quantity in the conservative case. Specifically, we considered the system (15)-(23) with no external excitations ($P = U = 0$), no damping ($\alpha = \gamma = 0$), and physical parameters identical to those in (54) below. In this system, it follows from (37) that the discrete energy $\mathscr{E}^n$ must remain invariant, i.e., $\mathscr{E}^n = \mathscr{E}^0$ for all n. Figure 2 shows that the fluctuation of the invariant energy, calculated as $|\mathscr{E}^n - \mathscr{E}^0|/\mathscr{E}^0$, remains reasonably small and consistent across different initial velocity values V^0. This confirms the accuracy of the scheme's implementation.

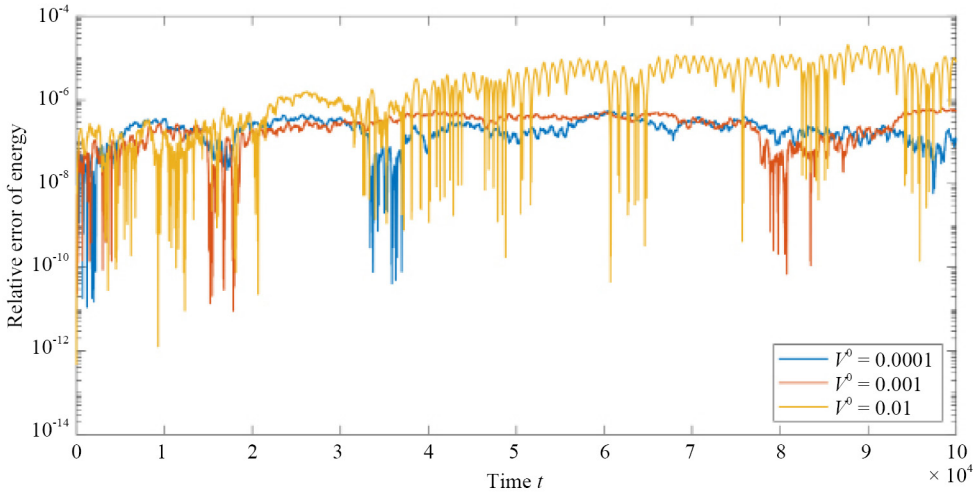

Figure 2. Relative errors $|\mathscr{E}^n - \mathscr{E}^0|/\mathscr{E}^0$ for conservative system ($P = U_\infty = \alpha = \gamma = 0$) with varying initial velocities V^0

6.1 *Viscoelastic beam subjected to follower force*

In this subsection, we examine the behavior of a viscoelastic beam subjected to a compressive follower force, characterized by

$$\alpha_0 > 0 \text{ in (3)}, \quad \gamma_0 > 0 \text{ in (6)}, \quad U_\infty = 0 \text{ in (6)}, \quad P_0 > 0 \text{ in (9)}.$$

6.1.1 *Physical parameters*

The material and geometric parameters of the beam, as specified in (1)-(4), were set as

$$\rho = 2,840\ [\mathrm{kg/m^3}], \quad A = bh, \quad I = bh^3/12, \quad E = 69\ [\mathrm{GPa}], \quad L = 0.508\ [\mathrm{m}], \tag{54}$$

$$\big(\text{base } b = 0.0254\ [\mathrm{m}], \quad \text{height } h = 0.0015\ [\mathrm{m}]\big).$$

The corresponding physical parameters in the normalized Equations (15)-(23) were calculated by (14): especially,

$$\beta = 1.38 \times 10^6.$$

The parameters in (54) are the same as those adopted in the simulation by McHugh and Dowell [4] based on the v-model (13). Due to the constraints of the v-model, they focused on motion where the vertical displacement of the beam's free end remained within approximately 20% of the beam length. The analyses in this section based on the ϕ-model does not require such restrictions.

Throughout this subsection, the axes in all figures represent normalized quantities as defined in (15)-(23). Their correspondence to unnormalized quantities from (1)-(9) is detailed in Table 1.

Table 1. Correspondence of normalized and actual quantities for follower force problem ($\nu = \sqrt{\rho L^2/E}$ in the table)

Normalized quantity		Actual quantity	
time	t	$1.03 \times 10^{-4}\ t$	$(= t\nu)$
energy	$\mathscr{E}$	$0.97\mathscr{E}$	$(= \mathscr{E}\ (EI/L))$
displacement	$(X,\ Y)$	$0.508\ (X,\ Y)$	$(= (LX, LY))$
frequency	F	$9.71 \times 10^3\ F$	$(= F/\nu)$
internal damping	α	$\alpha_0 = 0.121\ \alpha/\sqrt{\beta}$	$(= \alpha\nu)$
fluid density	γ	$\gamma_0 = 1050\ \gamma\sqrt{\beta}$	$(= \gamma A\sqrt{\rho E}/L)$
follower force	P	$P_0 = 1.91\ P$	$(= PEI/L^2)$
initial velocity	V^0	$9.71 \times 10^3\ V^0$	$(= V^0/\nu)$

6.1.2 *Snapshots, time histories of energy, and limit cycles*

The damping parameters were set to $\alpha/\sqrt{\beta} = 0.04$ and $\gamma\sqrt{\beta} = 0.05$.

For a follower load of $P = 22$, the time history of the total energy $\mathscr{E}^n$ (defined in (36)) reveals that the system converges to a limit cycle oscillation, as seen in Figure 3a. Figure 3b shows snapshots of the beam's shape at specific time points (indicated by asterisks in Figure 3a); each shape was calculated as the parametric curve $x \mapsto (X^n(x),\ Y^n(x))$ with

$$X^n(x) = \int_0^x \cos\varphi^n(\xi)\ \mathrm{d}\xi, \quad Y^n(x) = \int_0^x \sin\varphi^n(\xi)\ \mathrm{d}\xi, \quad 0 \le x \le 1. \tag{55}$$

The red curve in Figure 3b corresponds to the trajectory of the end point ($x = 1$) during the limit cycle phase (indicated by blue line in Figure 3a). Similar trajectories are drawn in Figure 4 for several values of follower force P.

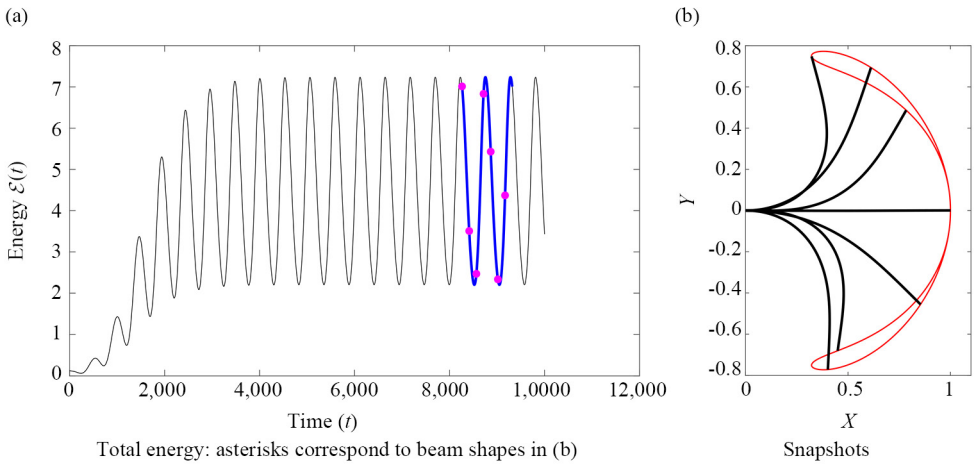

Figure 3. Time evolution of total energy and snapshots of beam motion for follower force $P = 22$ and damping $\alpha/\sqrt{\beta} = 0.04$, $\gamma\sqrt{\beta} = 0.05$

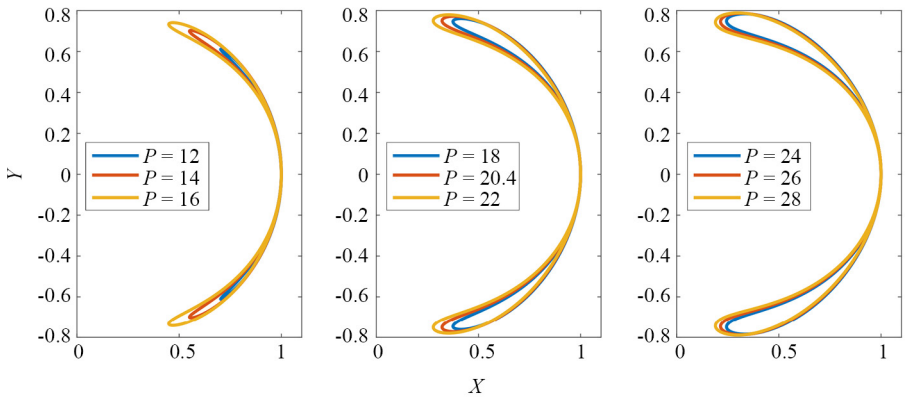

Figure 4. Locus curves of the free end in the limit cycle phase. $\alpha/\sqrt{\beta} = 0.04$, $\gamma\sqrt{\beta} = 0.05$

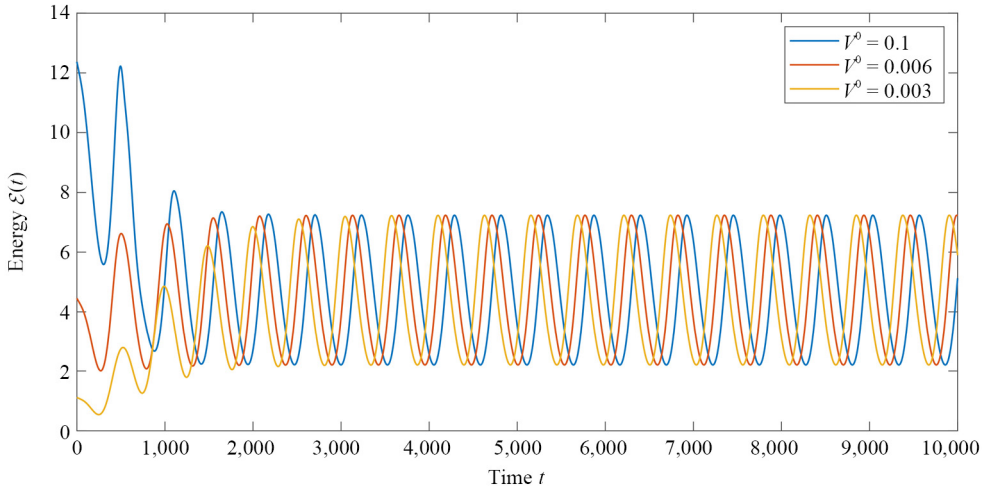

Figure 5. Total energies of beam with varying initial velocities. $P = 22$, $\alpha/\sqrt{\beta} = 0.04$, $\gamma\sqrt{\beta} = 0.05$

Furthermore, the effect of initial conditions on the limit cycle was investigated, demonstrating in Figure 5 that different initial velocities eventually converge to the same limit cycle oscillation.

6.1.3 *Stability of the system*

To study the stability of the system, we varied the internal damping coefficient α while fixing the external damping to $\gamma\sqrt{\beta} = 0.05$.

The stability changes with the magnitude of the follower force P. For the case of no internal damping ($\alpha = 0$), eigenvalue analysis of a linearized model estimates the critical value P_{crit} of P below which the system is stable as approximately $P_{\text{crit}} \approx 20.05$ (see e.g., [3, 4]). This result is consistent with our nonlinear model as shown in Figure 6, where the total energy decays exponentially for $P \leq 20.04$ and converges to limit cycle oscillations for $P \geq 20.06$.

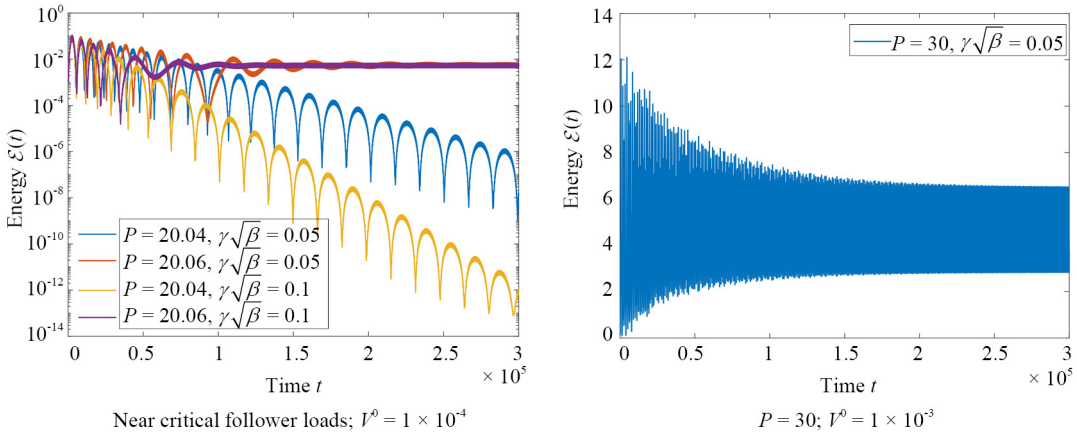

Figure 6. Total energies of undamped ($\alpha/\sqrt{\beta} = 0$) beam subjected to follower forces

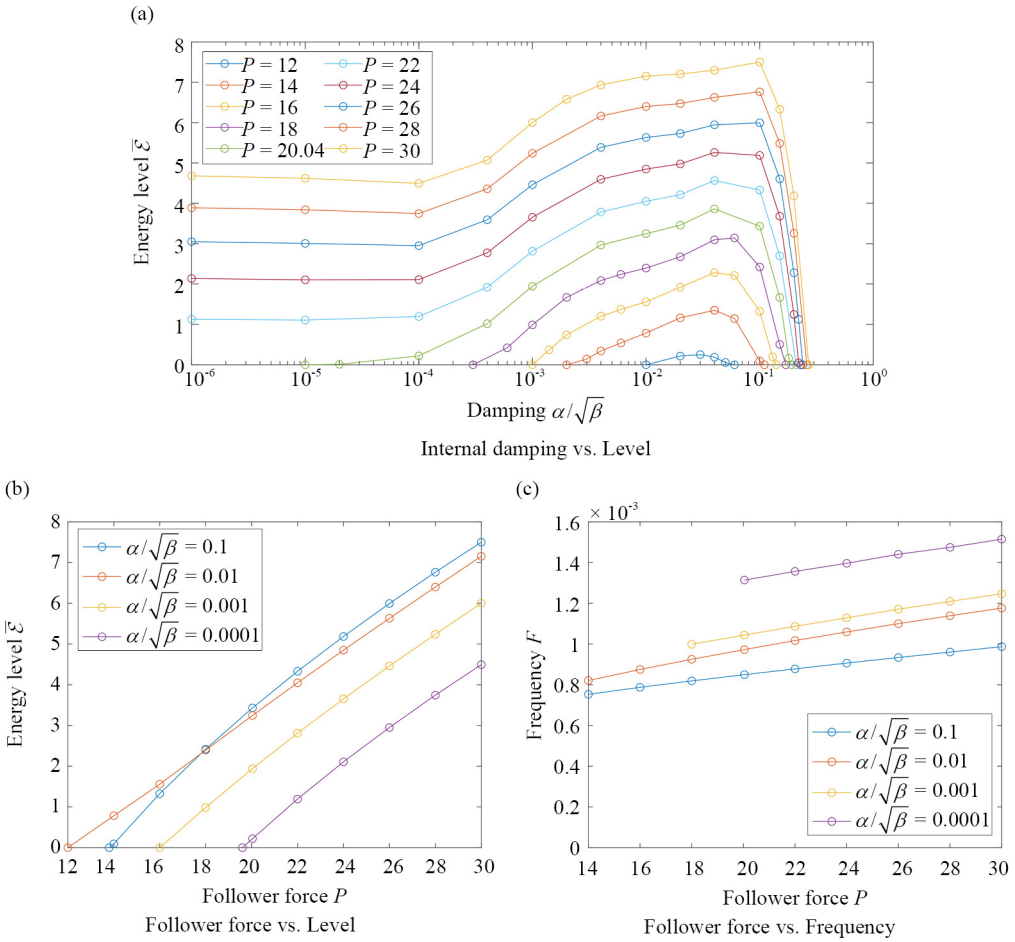

Figure 7. Energy levels and frequencies of limit cycle oscillations caused by follower force. $\gamma\sqrt{\beta} = 0.05$

Next, we examined the effect of internal damping on stability. Known as the the Ziegler's paradox, internal damping within a certain range destabilizes Beck's beam, reducing the critical force P_{crit} (see e.g., [3]). Numerical simulations

with our nonlinear model confirmed this destabilizing effect and revealed the post-critical behavior of the beam. As shown in Figure 7a, the energy level $\bar{\mathscr{E}}$ of the limit cycle, calculated as the time-averaged $\mathscr{E}(t)$ during the limit cycle phase, increases as the damping coefficient α increased over a certain range for all P values.

Additionally, the energy level of the limit cycle is observed in Figure 7b to depend linearly on the magnitude of P, with the critical value P_{crit} (estimated as the curve's intercept with the abscissa) varying with α as follows (Table 2):

Table 2. Estimates of critical values for follower force

$\alpha/\sqrt{\beta}$	0	0.0001	0.001	0.01	0.1
P_{crit}	$20.04 < P_{\text{crit}} < 20.06$	$19.6 < P_{\text{crit}} < 19.8$	$16.0 < P_{\text{crit}} < 16.2$	$12.0 < P_{\text{crit}} < 12.2$	$13.8 < P_{\text{crit}} < 14.0$

The frequency of the limit cycle oscillations also increases linearly with P while decreasing with higher α values, as seen in Figure 7c. The frequency was calculated from the time required for the beam's endpoint to complete one full loop of its trajectory.

6.2 *Viscoelastic beam subjected to axial flow-Simplified case with linear piston theory*

In this section, we analyze a viscoelastic beam subjected to axial flow. The governing parameters are specified as follows:

$$\alpha_0 > 0 \text{ in (3)}, \quad \gamma_0 > 0 \text{ in (6)}, \quad U_\infty \neq 0 \text{ in (6)}, \quad P_0 = 0 \text{ in (9)}.$$

6.2.1 *Physical parameters*

The material and geometric properties of the beam, as described in (1)-(3), were set as

$$\frac{EI}{\rho A} = \frac{\gamma_0}{\rho A} = 1, \quad L = 1 \text{ [m]}, \tag{56}$$

along with

$$\rho = 2,800 \text{ [kg/m}^3\text{]}, \quad A = 4 \times 10^{-5} \text{ [m}^2\text{]}, \quad E = 70 \text{ [GPa]}.$$

The parameters satisfying (56) were also adopted in the simulation by Deliyianni et al. [8] based on v-model (13). Using (14), the normalized parameters become

$$\beta \left(= E/\rho\right) = 2.5 \times 10^7, \quad \alpha = \alpha_0\sqrt{\beta}, \quad \gamma = 1/\sqrt{\beta}, \quad U = U_\infty/\sqrt{\beta}.$$

Throughout this subsection, unless otherwise specified, the axes in all figures represent normalized quantities as defined in (15)-(23). Their correspondence to unnormalized quantities from (1)-(9) is detailed in Table 3.

Table 3. Correspondence of normalized and actual quantities for beam s.t. axial flow

Normalized quantity		Actual quantity	
time	t	$2\times 10^{-4}\ t$	$\left(=t/\sqrt{\beta}\right)$
energy	$\mathscr{E}$	$0.112\mathscr{E}$	$\left(=\mathscr{E}\ (EI/L)\right)$
displacement	$(X,\ Y)$	$(X,\ Y)$	$\left(=(LX,\ LY)\right)$
frequency	F	$5{,}000\ F$	$\left(=F\sqrt{\beta}\right)$
internal damping	α	$\alpha_0=\alpha/\sqrt{\beta}$	
fluid density	γ	$\gamma_0=0.112$	$\left(=\gamma\rho A\sqrt{\beta}\right)$
flow speed	U	$U_\infty=5{,}000\ U$	$\left(=U\sqrt{\beta}\right)$
initial velocity	V^0	$5{,}000\ V^0$	$\left(=V^0\sqrt{\beta}\right)$

6.2.2 *Positive flow speed*

Snapshots, Time Histories of Energy, and Limit Cycles

First, we consider the case of positive flow velocity, $U_\infty > 0$, where the flow moves from the clamped end to the free end.

For flow speed of $U_\infty\ (=5{,}000\ U)=400$, the time history of the total energy $\mathscr{E}^n$ reveals that the system converges to a limit cycle oscillation, as seen in Figure 8a. Figure 8b shows snapshots of the beam's shape at specific time points (indicated by asterisks in Figure 8a); each shape was calculated as the parametric curve $x\mapsto(X^n(x),\ Y^n(x))$ defined by (55). The red curve in Figure 8b corresponds to the trajectory of the end point ($x=1$) during the limit cycle phase (indicated by blue line in Figure 8a).

Figure 9 illustrates that different initial velocities eventually converge to the same limit cycle oscillation.

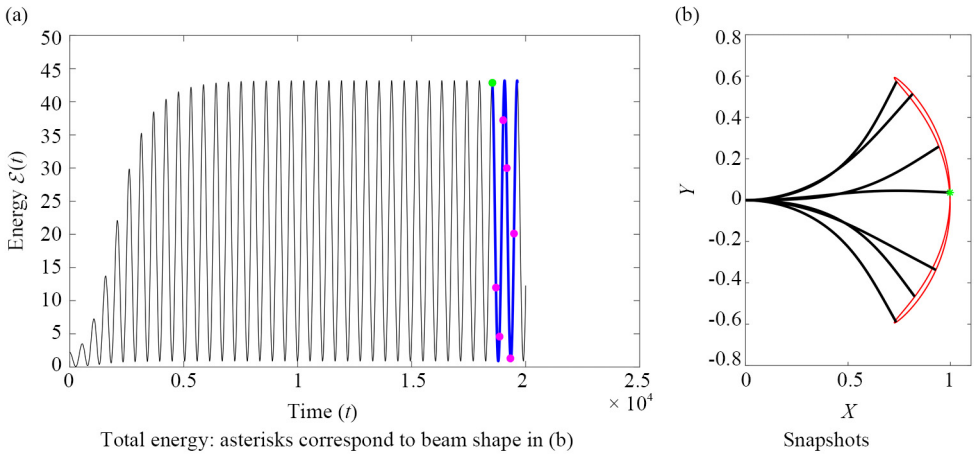

Figure 8. Time evolution of total energy and snapshots of beam motion for flow velocity $U_\infty=400$ and damping $\alpha_0=0.1$

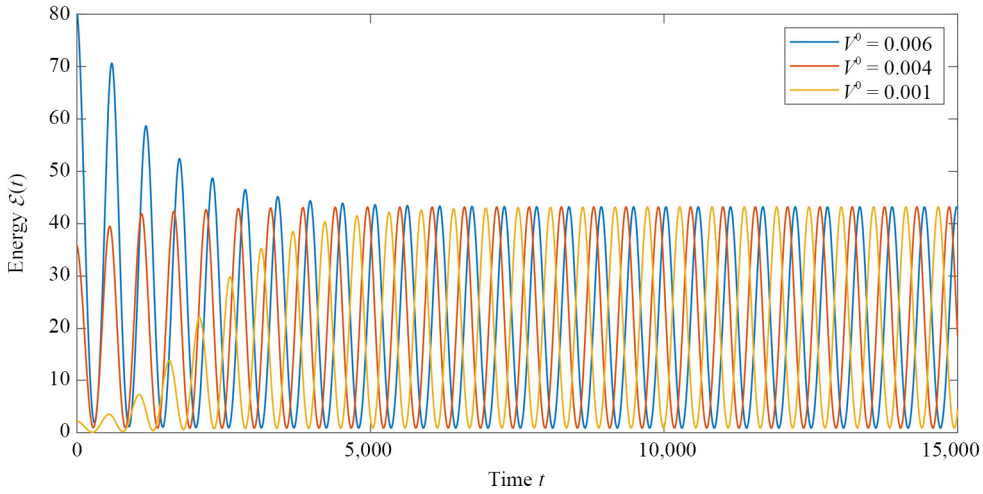

Figure 9. Total energies of beam with varying initial velocities. $U_\infty = 400$, $\alpha_0 = 0.1$

Stability of the System

The stability of the system depends on the flow speed U_∞. According to Deliyianni et al. [8], eigenvalue analysis of a linearized model estimates the critical flow speed below which the system is stable as $U_{\text{crit}} \approx 135.9$ for the case of no internal damping ($\alpha_0 = 0$). Simulations with our nonlinear model confirm this estimate: for $U_\infty \geq 136.0$, the total energy evolves into limit cycle oscillations, while for $U_\infty \leq 135.9$, the energy decays exponentially, as seen in Figure 10.

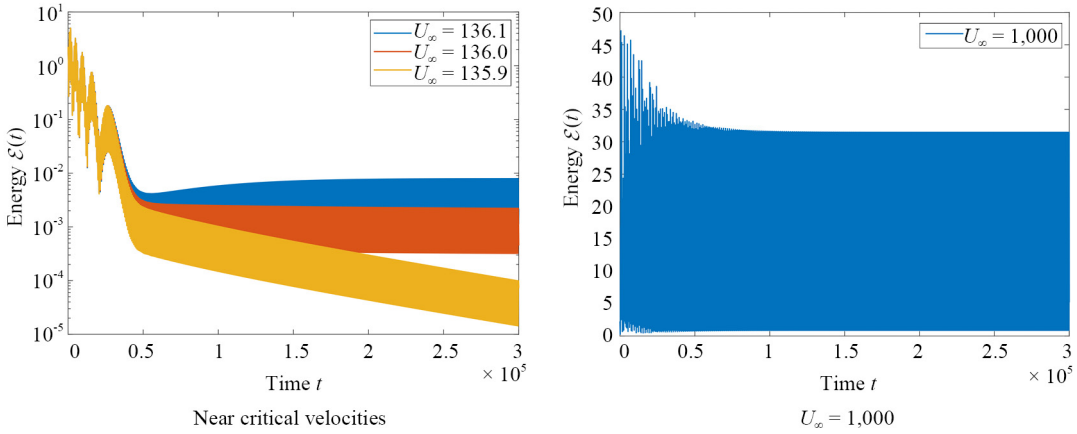

Figure 10. Total energies of undamped ($\alpha_0 = 0$) beam in the flow

As observed in Beck's beam, internal damping in a certain range destabilizes the system. Figure 11a shows that the energy level $\bar{\mathscr{E}}$ of the limit cycle, calculated as the time-averaged $\mathscr{E}(t)$ during the limit cycle phase, increases with the damping coefficient α_0 over a certain range for all U_∞ values.

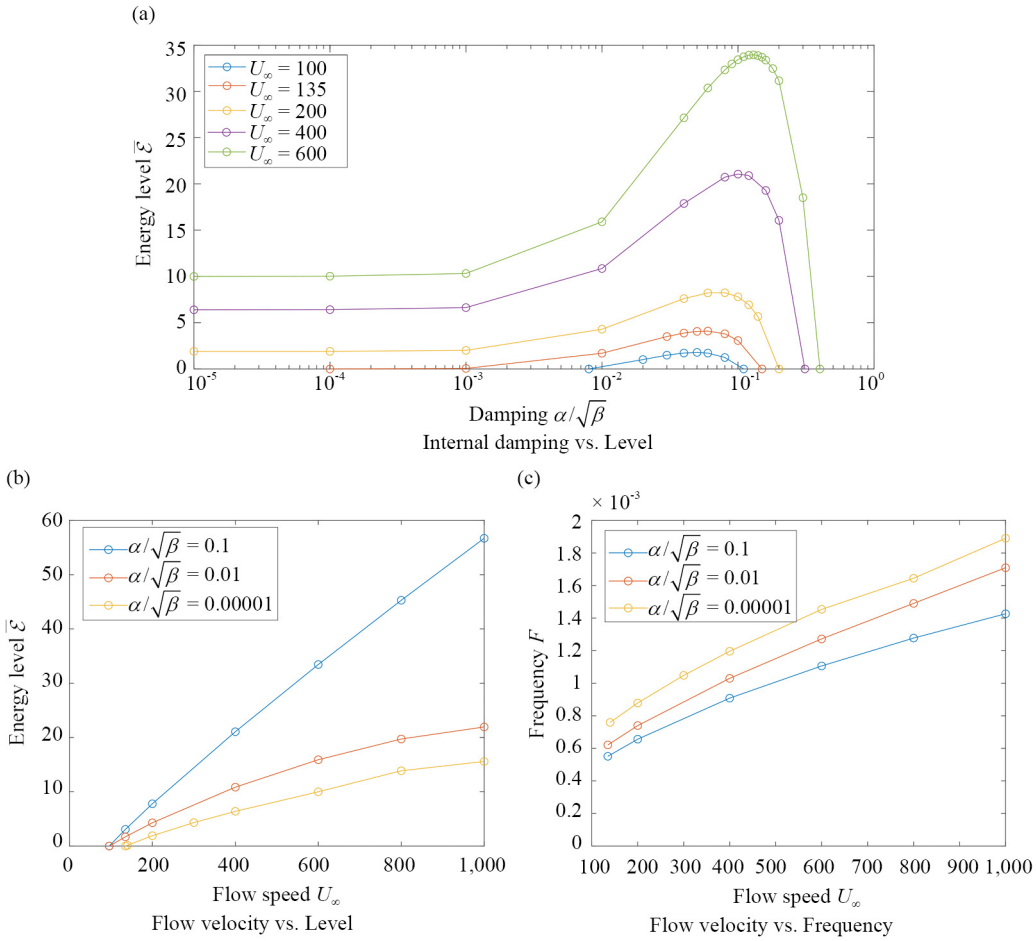

Figure 11. Energy levels and frequencies of limit cycle oscillations caused by axial flow

Additionally, Figure 11b shows that the critical flow speed U_{crit}, estimated as the point where each curve intersects the abscissa, slightly decreases with increasing α_0 (Table 4):

Table 4. Estimates of critical values for flow speed

α_0 $(= \alpha/\sqrt{\beta})$	0	0.00001	0.01	0.1
	$135.9 < U_{crit} < 136.0$	$135 < U_{crit} < 136$	$96 < U_{crit} < 97$	$96 < U_{crit} < 97$

On the other hand, increasing damping reduces the frequency of the limit cycle oscillations for a given U_∞, as seen in Figure 11c.

6.2.3 *Negative flow speed*

Finally, we consider the inverted flag configuration, where the flow moves from the free end to the clamped end, i.e., $U_\infty < 0$. The internal damping is fixed as $\alpha_0 = 0.01$.

There exists a critical flow speed $\hat{U}_{\text{crit}}$ such that for $0 \leq -U_\infty < -\hat{U}_{\text{crit}}$, the beam converges to a trivial equilibrium, whereas for $-\hat{U}_{\text{crit}} < -U_\infty$, the beam converges to a non-trivial equilibrium. Figure 12 illustrates that the critical value

lies within $6.3 < -\hat{U}_{\text{crit}} < 6.4$. Figure 13a shows the equilibrium configurations for various U_∞ values, and Figure 13b does the corresponding convergence of total energy to equilibrium levels.

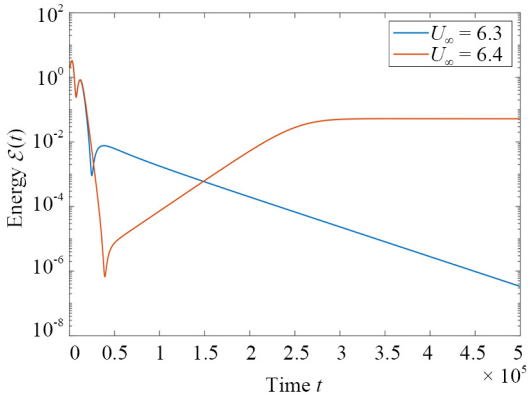

Figure 12. Total energies of inverted flag in the flow of near critical velocities. $\alpha_0 = 0.01$

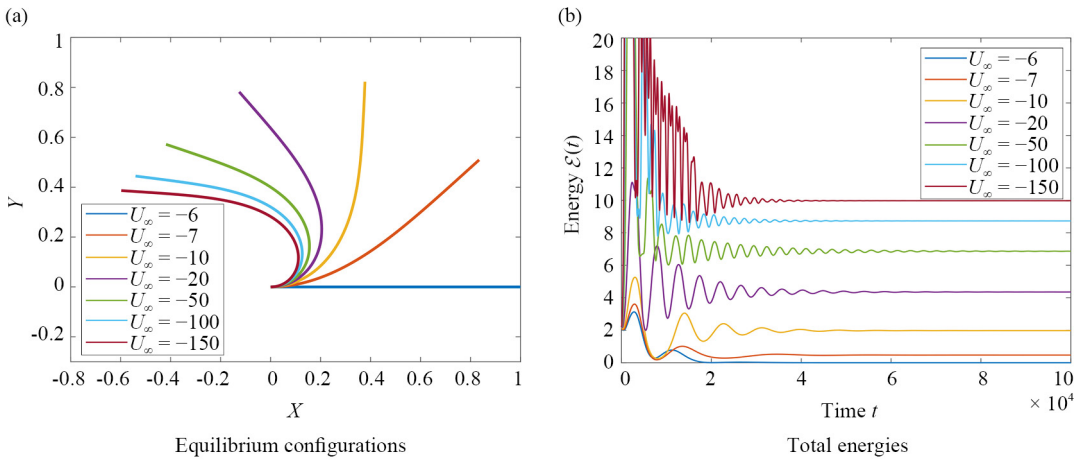

Figure 13. Asymptotic behavior of inverted flag in the flow of varying velocities. $\alpha_0 = 0.01$

7. Conclusions and future scope

This paper presented a numerical study of a nonlinear model, referred to as the ϕ-model, which describes the geometrically large deflection and planar motion of an inextensible cantilevered beam subjected to non-conservative forces. Two types of non-conservative forces were considered: a compressive follower force (Beck's beam) and piston-theoretic pressure resulting from axial airflow.

To approximate solutions of the initial-boundary value problem (IBVP) for this model, a spectral Galerkin scheme was proposed. Its numerical validity was established by proving that if the exact solution of the original problem belongs to the C^m-class for some m, the scheme is convergent of order $O(N^{-m+\varepsilon}) + O(\Delta t^2)$.

Numerical simulations were conducted using the proposed scheme to investigate the critical values of non-conservative forces and the system's post-critical behavior. For Beck's beam, critical follower forces were determined for varying levels of internal damping. The values determined using the ϕ-model were shown to agree with previously published results based on eigenvalue analyses of linearized models. Regarding the post-critical regime, a detailed analysis

of the destabilizing effects of internal damping revealed that, for each magnitude of follower force, there exists a finite range of internal damping where increasing the damping increases the amplitude of limit cycle oscillations (LCOs).

Similar findings were obtained for the piston-theoretic beam. Critical flow velocities were determined for varying levels of internal damping. In particular, in the absence of internal damping, the determined critical flow velocity was shown to agree with previously published results. Furthermore, it was demonstrated that the increasing effect of internal damping on the amplitude of LCOs observed in Beck's beam also applies to the piston-theoretic beam. Additionally, the inverted flag problem, corresponding to negative flow velocities, was discussed.

Building upon the results of this paper, several important directions for future research can be explored:

• In the proposed scheme, the time derivative was discretized using a central difference method, allowing for rigorous error evaluation. However, for improved computational efficiency, the introduction of higher-accuracy discretization methods is desirable. Nevertheless, simply applying conventional discretization methods does not guarantee a scheme with rigorous error evaluation. It is necessary to reconsider the overall structure, including aspects beyond just the time derivative.

• The present study assumes that the beam has a uniform cross-section and homogeneous material properties. An important future direction is to explore whether the proposed scheme can be extended to models with non-uniform cross-sections, nonlinear materials or layered materials. In such cases, a critical challenge remains: ensuring that the mathematical proof of rigorous error evaluation, which is a primary focus of this study, can still be established.

• This paper focused on in-plane beam motion, but beam equations describing three-dimensional motion are also well known [15]. Developing a numerical scheme for such three-dimensional beam equations is another promising research direction.

• For the v-model, some studies compare numerical analyses with experimental results (e.g., [18, 19, 31]). Similar studies are also desirable for the present ϕ-model, which considers geometrically large deformations of beams.

Acknowledgement

The author sincerely thanks the reviewers for their valuable comments and suggestions, which have greatly contributed to improving the manuscript and providing insights for future research.

Conflict of interest

The author declares no competing financial interest.

References

[1] Bolotin VV. *Nonconservative Problem of the Theory of Elastic Stability*. Oxford, UK: Pergamon Press; 1963.

[2] Langthjem MA, Sugiyama Y. Dynamic stability of columns subjected to follower loads: A survey. *Journal of Sound and Vibration*. 2000; 238(5): 809-851. Available from: https://doi.org/10.1006/jsvi.2000.3137.

[3] Luongo A, D'Annibale F. Nonlinear hysteretic damping effects on the post-critical behaviour of the visco-elastic Beck's beam. *Mathematics and Mechanics of Solids*. 2017; 22(6): 1347-1365.

[4] McHugh KA, Dowell EH. Nonlinear response of an inextensible, cantilevered beam subjected to a nonconservative follower force. *Journal of Computational and Nonlinear Dynamics*. 2019; 14(3): 031004. Available from: https://doi.org/10.1115/1.4042324.

[5] Luongo A, D'Annibale F. Double zero bifurcation of non-linear viscoelastic beams under conservative and non-conservative loads. *International Journal of Non-Linear Mechanics*. 2013; 55: 128-139. Available from: https://doi.org/10.1016/j.ijnonlinmec.2013.05.007.

[6] Luongo A, D'Annibale F. On the destabilizing effect of damping on discrete and continuous circulatory systems. *Journal of Sound and Vibration*. 2014; 333(24): 6723-6741. Available from: https://doi.org/10.1016/j.jsv.2014.07.030.

[7] Mei C, Abdel-Motagaly K, Chen RR. Review of nonlinear panel flutter at supersonic and hypersonic speeds. *Applied Mechanics Reviews*. 1999; 52(10): 321-332. Available from: https://doi.org/10.1115/1.3098919.

[8] Deliyianni M, Gudibanda V, Howell J, Webster JT. Large deflections of inextensible cantilevers: Modeling, theory, and simulation. *Mathematical Modelling of Natural Phenomena*. 2020; 15: 44. Available from: https://doi.org/10.1051/mmnp/2020033.

[9] Vedeneev VV. Panel flutter at low supersonic speeds. *Journal of Fluids and Structures*. 2012; 29: 79-96. Available from: https://doi.org/10.1016/j.jfluidstructs.2011.12.011.

[10] McHugh KA, Freydin M, Bastos KK, Beran P, Dowell EH. Flutter and limit cycle oscillations of cantilevered plate in supersonic flow. *Journal of Aircraft*. 2021; 58(2): 266-278. Available from: https://doi.org/10.2514/1.C035992.

[11] Stanton SC, Choi SJ, McHugh KA. On the influence of nonlinear inertial forces on the limit cycle oscillations of an inextensible plate in a supersonic axial flow. *Journal of Vibration and Acoustics*. 2023; 145(3): 031002. Available from: https://doi.org/10.1115/1.4056127.

[12] Howell J, Huneycutt K, Webster JT, Wilder S. A thorough look at the (in)stability of piston-theoretic beams. *Mathematics in Engineering*. 2019; 1(3): 614-647. Available from: https://doi.org/10.3934/mine.2019.3.614.

[13] Coleman BD, Dill EH. Flexure waves in elastic rods. *The Journal of the Acoustical Society of America*. 1992; 91(5): 2663-2673. Available from: https://doi.org/10.1121/1.402974.

[14] Coleman BD, Xu JM. On the interaction of solitary waves of flexure in elastic rods. *Acta Mechanica*. 1995; 110: 173-182. Available from: https://doi.org/10.1007/BF01215423.

[15] Coleman BD, Dill EH, Lembo M, Lu Z, Tobias I. On the dynamics of Rods in the theory of Kirchhoff and Clebsch. *Archive for Rational Mechanics and Analysis*. 1993; 121: 339-359. Available from: https://doi.org/10.1007/BF00375625.

[16] Chen JS, Wu HH. Deformation and stability of an elastica under a point force and constrained by a flat surface. *International Journal of Mechanical Sciences*. 2011; 53(1): 42-50. Available from: https://doi.org/10.1016/j.ijmecsci.2010.10.005.

[17] Crespo da Silva MRM. Harmonic non-linear response of Beck's column to a lateral excitation. *International Journal of Solids and Structures*. 1978; 14(12): 987-997. Available from: https://doi.org/10.1016/0020-7683(78)90080-X.

[18] Tang DM, Yamamoto H, Dowell EH. Flutter and limit cycle oscillations of two-dimensional panels in three-dimensional axial flow. *Journal of Fluids and Structures*. 2003; 17(2): 225-242. Available from: https://doi.org/10.1016/S0889-9746(02)00121-4.

[19] Tang L, Païdoussis MP. On the instability and the post-critical behaviour of two-dimensional cantilevered flexible plates in axial flow. *Journal of Sound and Vibration*. 2007; 305(1-2): 97-115. Available from: https://doi.org/10.1016/j.jsv.2007.03.042.

[20] Tang L, Païdoussis MP, Jiang J. Cantilevered flexible plates in axial flow: Energy transfer and the concept of flutter-mill. *Journal of Sound and Vibration*. 2009; 326(1-2): 263-276. Available from: https://doi.org/10.1016/j.jsv.2009.04.041.

[21] Zhao W, Païdoussis MP, Tang L, Liu M, Jiang J. Theoretical and experimental investigations of the dynamics of cantilevered flexible plates subjected to axial flow. *Journal of Sound and Vibration*. 2012; 331(3): 575-587. Available from: https://doi.org/10.1016/j.jsv.2011.08.014.

[22] Ito K. An error estimate for an energy conserving spectral scheme approximating the dynamic elastica with free ends. *Applied Numerical Mathematics*. 2017; 120: 1-20. Available from: https://doi.org/10.1016/j.apnum.2017.05.001.

[23] Caflisch RE, Maddocks JH. Nonlinear dynamical theory of the elastica. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*. 1984; 99(1-2): 1-23. Available from: https://doi.org/10.1017/S03082105, 00025920.

[24] Ito K. Uniform stabilization of the dynamic elastica by boundary feedback. *SIAM Journal on Control and Optimization*. 1998; 37(1): 319-329. Available from: https://doi.org/10.1137/S0363012997322352.

[25] Dowell E, McHugh K. Equations of motion for an inextensible beam undergoing large deflections. *Journal of Applied Mechanics*. 2016; 83(5): 051007. Available from: https://doi.org/10.1115/1.4032795.

[26] McHugh K, Dowell E. Nonlinear responses of inextensible cantilever and free-free beams undergoing large deflections. *Journal of Applied Mechanics*. 2018; 85(5): 051008. Available from: https://doi.org/10.1115/1.4039478.

[27] Shen J, Tang T, Wang LL. *Spectral Methods, Algorithms, Analysis and Applications*. Berlin, Germany: Springer-Verlag; 2011.

[28] Canuto C, Hussaini MY, Quarteroni A, Zang TA. *Spectral Methods, Fundamentals in Single Domains*. Berlin, Germany: Springer-Verlag; 2006.

[29] Falk RS, Xu JM. Convergence of a second-order scheme for the nonlinear dynamic equations of elastic rods. *SIAM Journal on Numerical Analysis*. 1995; 32(4): 1185-1209.

[30] Ito K. Convergence of a spectral Galerkin scheme for the dynamic elastica with periodical or hinged boundary. *Applied Numerical Mathematics*. 2022; 176: 56-82. Available from: https://doi.org/10.1016/j.apnum.2022.02.010.

[31] Tang D, Zhao M, Dowell EH. Inextensible beam and plate theory: Computational analysis and comparison with experiment. *Journal of Applied Mechanics*. 2014; 81(6): 061009. Available from: https://doi.org/10.1115/1.4026800.