A CLASS OF EFFICIENT SPECTRAL METHODS AND ERROR ANALYSIS FOR NONLINEAR HAMILTONIAN SYSTEMS[*]

JING AN[†], WAIXIANG CAO[‡], AND ZHIMIN ZHANG[§]

Abstract. We investigate efficient numerical methods for nonlinear Hamiltonian systems. Three polynomial spectral methods (including spectral Galerkin, Petrov-Galerkin, and collocation methods) coupled with domain decomposition are presented and analyzed. Our main results include the energy and symplectic structure-preserving properties and error estimates. We prove that the spectral Petrov-Galerkin method preserves the energy exactly while both the spectral Gauss collocation and spectral Galerkin methods are energy conserving up to spectral accuracy. While it is well known that collocation at Gauss points preserves symplectic structure, we prove that, for both the Petrov-Galerkin method and the spectral Galerkin method, the error in symplecticity decays with spectral accuracy. Finally, we show that all three methods converge exponentially with respect to the polynomial degree. Numerical experiments indicate that our algorithms are efficient.

Keywords. Nonlinear Hamiltonian system; spectral methods; error analysis; energy conservation; symplectic structure.

AMS subject classifications. 65M15; 65M70; 65N30.

1. Introduction

Hamiltonian dynamical systems have many applications in classical mechanics, molecular dynamics, hydrodynamics, electrodynamics, plasma physics, relativity, astronomy, and other scientific fields [32, 33]. However, in practice, it is impossible to find exact solutions for most nonlinear Hamiltonian systems, and therefore efficient numerical methods are desired. On the other hand, a Hamiltonian system has many remarkable properties, most important among which are its symplectic structure and energy preservation. A good numerical scheme should be able to mimic as many of these physical properties as possible.

The symplectic geometry algorithm that maintains the symplectic structure for Hamiltonian systems was discussed as early as 1985 by Feng [14]. Since then symplectic algorithms (e.g., Runge-Kutta methods, spectral variational integrators, etc.) have been intensively studied. We refer to [3, 15–21, 25, 26, 30, 33–35, 37, 38] for an incomplete list of references. However, none of the symplectic algorithms are energy-preserving in general. In fact, it was proved in [12, 43] that there exists no constant step size energy-preserving symplectic algorithm that is not the exact solution for general nonlinear Hamiltonian systems. Then we face a dilemma and have to choose between preserving energy and preserving symplectic structure. So far, it is not clear which algorithms should be preferred for a given application. Although symplectic algorithms do not

[†]Beijing Computational Science Research Center, Beijing 100193, China (anjing@csrc.ac.cn); and School of Mathematical Sciences, Guizhou Normal University, Guiyang 550025, China (aj154@163.com).

[‡]Corresponding author. School of Mathematical Sciences, Beijing Normal University, Beijing 100875, China (caowx@bnu.edu.cn).

[§]Beijing Computational Science Research Center, Beijing 100193, China (zmzhang@csrc.ac.cn); and Department of Mathematics, Wayne State University, Detroit, MI 48202, USA (zzhang@math.wayne.edu).

in general conserve the energy of a mechanical system, energy invariance of symplectic algorithms remains in a reasonable range over long time intervals (see, e.g., [30]). Moreover, the properties of symplectic integrators also make them highly suitable for long-time integration of chaotic Hamiltonian systems.

Energy-preserving algorithms, on the other hand, have also been developed in the last decade. In [6,40], Chen et al. used finite element methods to solve ordinary differential equations and proved that the linear element was a second-order pseudo-symplectic scheme and the quadratic element was a third-order pseudo-symplectic scheme, respectively, and both linear and quadratic elements preserved energy. However, to predict long-time orbital evolution, linear and quadratic finite elements require a tremendous amount of computation, and accumulation of round-off errors will be eventually dominant. In [4, 5], the authors studied the energy-preserving Runge-Kutta method for polynomial Hamiltonian system. Other numerical methods have been proposed in the literature to study the long-time near-conservation of the total and oscillatory energies. In [7], Cohen used a modulated Fourier expansion to show long-time near-conservation of the total and oscillatory energies for Hamiltonian systems with highly oscillatory solutions. In a more recent work [29], Hairer and Lubich studied the long-time behavior of the Störmer-Verlet-leapfrog method for highly oscillatory Hamiltonian systems with a slowly varying, solution-dependent high frequency, and proved that the proposed method conserved approximately a modified total energy over a long time interval. We also refer to [8–10,22,23,27,28,30] for some more works in this direction. In the spirit of high-order methods, Kanyamee and Zhang [36], Huang and Zhang [31] proposed algorithms based upon the spectral collocation method and studied energy and symplectic structure preserving properties over long time interval. This approach works well for some problems. The drawback is that the differential matrices are usually full and the condition number of the stiffness matrix increases dramatically with increasing polynomial degrees. Large condition numbers lead to instability of the algorithm in some situations, e.g., the calculation of many-body problems.

The main purpose of the current work is to present and study a class of efficient polynomial spectral methods for nonlinear Hamiltonian systems that preserve the energy and symplectic structure simultaneously, with numerically negligible error in practice. Note that if the error term for the energy and symplectic structure is so small that it reaches the machine epsilon (the computer round-off error), then we say the algorithm is energy and symplectic structure preserving in practice. By "spectral" we mean the convergence is achieved by increasing polynomial degrees N rather than decreasing the step-size in time (as most of ODE and Hamiltonian solvers in the literature). Comparing with the standard h-version method (see, e.g., [2, 6, 40]), whose convergence rate is achieved by decreasing the time step size h (the polynomial degree N is fixed) and the approximation error has polynomial convergence to zero with respect to the variable h, the desirable advantage of the spectral method lies in that the error converges exponentially with respect to the variable N. In other words, to achieve the same error accuracy, the computational cost of the spectral method is less than that of the h-version method. Consequently, when applied to the Hamiltonian system, the approximation errors in energy and symplectic structure of the spectral method can reach the machine epsilon with a reasonable polynomial degree N and computational cost, which makes it possible to study the energy and symplectic structure preserving properties over large time interval efficiently.

In this paper, a rigorous mathematical proof is given to show the high-order accuracy of the proposed spectral methods and numerical experiments are provided to

demonstrate the efficiency of the proposed numerical schemes in simulating the underlying Hamiltonian systems. To be more precise, we present three polynomial spectral methods for nonlinear Hamiltonian systems: spectral Petrov-Galerkin, spectral Gauss collocation, and spectral Galerkin methods, and prove that they share the same property of high-order (spectral) accuracy. Furthermore, we investigate properties of symplectic structure and energy conservation for the three numerical methods and establish results including: (1) The Petrov-Galerkin method preserves the energy exactly, while both Gauss collocation and spectral Galerkin methods are energy conserving up to numerically negligible error terms; (2) both spectral Galerkin and Petrov-Galerkin methods have spectral accuracy in symplectic structure error. Due to the spectral accuracy, the three spectral methods preserve the energy and symplectic structure in practical experiments, with numerically negligible machine epsilon and reasonable computational cost. Moreover, note that the stiff matrices of Petrov-Galerkin and spectral Galerkin methods are both sparse and hence well conditioned, which overcomes the problem of large condition numbers in [36].

To end this introduction, we would like to indicate that Gauss collocation is a well-known symplectic scheme (see, e.g., [30]) and the Gauss spectral collocation method has been used to solve some ordinary differential equations (see, e.g., [24]) and its conservation of energy up to a numerical quadrature error was mentioned in [36]. However, we have not found any rigorous mathematical proof of energy conservation property for spectral Gauss collocation methods when applied to Hamiltonian systems. A preconditioning technique is also proposed for Gauss collocation methods, which yields a sparse and well conditioned stiff matrix. Furthermore, to the best of our knowledge, there isn't any error analysis in symplectic structure and convergence for the spectral Petrov-Galerkin method coupled with time domain decomposition in this context (i.e., the error bound with respect to the polynomial degree), and therefore, our error analysis in symplectic structure for the spectral Petrov-Galerkin method is new for solving Hamiltonian dynamical systems.

The rest of this paper is organized as follows. In the next section, we present three numerical schemes for Hamiltonian systems and study the efficient implementation of them. In §3, we analyze the conservation properties (including energy preserving and symplectic structure) of the three spectral methods. In §4, we establish error estimates for the algorithms and prove that the approximation error converges exponentially. Several numerical experiments are presented in §5 to demonstrate the accuracy and efficiency of the methods. Finally, in §6 we give some concluding remarks.

2. Numerical schemes for Hamiltonian systems

We consider the canonical Hamiltonian system

$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}, \quad \frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \qquad i = 1, 2, \cdots, n, \tag{2.1}$$

$$p_i(0) = p_{i0}, \qquad q_i(0) = q_{i0}, \qquad i = 1, 2, \cdots, n, \tag{2.2}$$

where $H(p_1, \cdots, p_n; q_1, \cdots, q_n)$ is a Hamiltonian function. We shall use the spectral Galerkin method, the spectral Gauss collocation method, and the spectral Petrov-Galerkin method coupled with domain decomposition to solve it.

The basic idea of the polynomial spectral method coupled with domain decomposition is similar to that in [24,36]. To approximate the solution (p, q) at any time $t > 0$, we first choose some positive r (which could be large) and solve the above system on $[0, r]$ by spectral methods, and then we use the obtained value $(p_i(r), q_i(r)), i \leq n$ as an initial

condition to repeat the process on $[r, 2r]$, and so on. In other words, for all $k = 0, 1, 2, \cdots$, we use three spectral methods to solve the following system:

$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}, \quad \frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \quad t \in (kr, (k+1)r], \tag{2.3}$$

$$p_i(kr) = \tilde{p}_{i0}, \quad q_i(kr) = \tilde{q}_{i0}, \quad i = 1, 2, \cdots, n. \tag{2.4}$$

Here $\tilde{p}_{i0}, \tilde{q}_{i0}$ are initial values obtained from the interval $[(k-1)r, kr]$ for all $k \geq 1$ and $\tilde{p}_{i0} = p_{i0}, \tilde{q}_{i0} = q_{i0}$ for $k = 0$.

To simplify our discussion and demonstrate the idea clearly, we rewrite (2.3)-(2.4) into a system with homogenous initial conditions on $[-1, 1]$. Actually, by a scaling from $[-1, 1]$ to $[kr, (k+1)r]$ and the variable substitution (i.e., (p_i, q_i) replaced by $(\hat{p}_i, \hat{q}_i) = (p_i - \tilde{p}_{i0}, q_i - \tilde{q}_{i0}), i = 1, \cdots, n)$, (2.3)-(2.4) is equivalent to the following system:

$$\frac{d\hat{p}_i}{dx} = -\frac{\partial \hat{H}}{\partial \hat{q}_i}, \quad \frac{d\hat{q}_i}{dx} = \frac{\partial \hat{H}}{\partial \hat{p}_i}, \quad x \in (-1, 1], \tag{2.5}$$

$$\hat{p}_i(-1) = 0, \quad \hat{q}_i(-1) = 0, \quad i = 1, 2, \cdots, n, \tag{2.6}$$

where $\hat{H}(\hat{p}_1, \cdots, \hat{p}_n; \hat{q}_1, \cdots, \hat{q}_n) = rH(p_1, \cdots, p_n; q_1, \cdots, q_n)/2$, and $x = 2t/r - 2k - 1 \in [-1, 1]$.

In our practical calculation, we obtain the initial values $\tilde{p}_{i0}, \tilde{q}_{i0}$ at the interval $(r, 2r]$ by the following way: We first obtain $\hat{p}_i, \hat{q}_i$ on $[-1, 1]$ from (2.5)-(2.6) and the initial solution p_{i0}, q_{i0} with p_{i0}, q_{i0} given by (2.2) . Then a simple variable substitution yields the value of the numerical solution at the point $t = r$. That is,

$$(p_i, q_i)(r) = (\hat{p}_i(1) + p_{i0}, \hat{q}_i(1) + q_{i0}).$$

Then the initial solution at the interval $(r, 2r]$ is taken as

$$(\tilde{p}_{i0}, \tilde{q}_{i0}) = (p_i, q_i)(r).$$

Repeating the process, it is easy to get the initial solution $\tilde{p}_{i0}, \tilde{q}_{i0}$ at each interval $(kr, (k+1)r], k \geq 2$.

2.1. Spectral Galerkin methods. Denote by P_N the space of polynomial functions with a degree of no more than N on $[-1, 1]$, and define the following approximation space

$$X_N = \{v : v \in P_N, \ v(-1) = 0\}.$$

The spectral Galerkin method for (2.5)-(2.6) is to find $\hat{p}_{iN}, \hat{q}_{iN} \in X_N$ such that for all $v_{iN}, w_{iN} \in X_N$ and $i = 1, 2, \cdots, n$,

$$\left(\frac{d\hat{p}_{iN}}{dx}, v_{iN}\right) = -\left(\frac{\partial \hat{H}(\hat{p}_{1N}, \cdots, \hat{p}_{nN}; \hat{q}_{1N}, \cdots, \hat{q}_{nN})}{\partial \hat{q}_{iN}}, v_{iN}\right), \tag{2.7}$$

$$\left(\frac{d\hat{q}_{iN}}{dx}, w_{iN}\right) = \left(\frac{\partial \hat{H}(\hat{p}_{1N}, \cdots, \hat{p}_{nN}; \hat{q}_{1N}, \cdots, \hat{q}_{nN})}{\partial \hat{p}_{iN}}, w_{iN}\right), \tag{2.8}$$

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

2.2. Spectral Gauss collocation methods. Let ϕ_n and L_n be the standard Lobatto and Legendre polynomials on $[-1,1]$, respectively. To be more precise,

$$\phi_0 = \frac{1-s}{2}, \quad \phi_1 = \frac{1+s}{2}, \quad \phi_n = \int_{-1}^{s} L_{n-1} ds = \frac{1}{2n-1}(L_n - L_{n-2}) \quad n \geq 2. \tag{2.9}$$

Denote by $-1 < s_1 < s_2 < \cdots < s_N < 1$ the Gauss points of degree N. That is, $s_i, i \leq N$ are N zeros of the Legendre polynomial L_N. Then the Gauss collocation method for (2.5)-(2.6) is: Find $\hat{p}_{iN}, \hat{q}_{iN} \in X_N$ such that for all $s_j, j \leq N$,

$$\frac{d\hat{p}_{iN}(s_j)}{dx} = -\frac{\partial \hat{H}(s_j)}{\partial \hat{q}_{iN}}, \quad \frac{d\hat{q}_{iN}(s_j)}{dx} = \frac{\partial \hat{H}(s_j)}{\partial \hat{p}_{iN}}. \tag{2.10}$$

Here $\hat{H}(s_j) = \hat{H}(\hat{p}_{1N}(s_j), \cdots, \hat{p}_{nN}(s_j); \hat{q}_{1N}(s_j), \cdots, \hat{q}_{nN}(s_j))$.

2.3. Spectral Petrov-Galerkin methods. We choose X_N and P_{N-1} as our trial and test spaces, respectively. Then the Petrov-Galerkin method for (2.5)-(2.6) is to find $\hat{p}_{iN}, \hat{q}_{iN} \in X_N$ such that for all $v_{iN}, w_{iN} \in P_{N-1}$,

$$\left(\frac{d\hat{p}_{iN}}{dx}, v_{iN}\right) = -\left(\frac{\partial \hat{H}(\hat{p}_{1N}, \cdots, \hat{p}_{nN}; \hat{q}_{1N}, \cdots, \hat{q}_{nN})}{\partial \hat{q}_{iN}}, v_{iN}\right), \tag{2.11}$$

$$\left(\frac{d\hat{q}_{iN}}{dx}, w_{iN}\right) = \left(\frac{\partial \hat{H}(\hat{p}_{1N}, \cdots, \hat{p}_{nN}; \hat{q}_{1N}, \cdots, \hat{q}_{nN})}{\partial \hat{p}_{iN}}, w_{iN}\right), \tag{2.12}$$

where $i = 1, 2, \cdots, n$.

We would like to point out that the choice of the test space is not unique, and different choices of the test space may lead to different numerical schemes. For example, we may take

$$\bar{X}_N = \{v : v \in P_N, \ v(1) = 0\}$$

as our test space. Without loss of generality, we only consider the case where the test space is chosen as P_{N-1}.

REMARK 2.1. Note that the initial values in the above three spectral methods are obtained from the previous element, which indicates that the three spectral methods are actually continuous, and thus are different from the discontinuous-time finite element methods proposed in [41,42]. Furthermore, the spectral Galerkin method is also different from the k-degree continuous-time finite element (k-C-TFE) method in [41,42], although both the numerical discretizations are based upon variational formulation. Compared with the k-C-TFE method, the spectral Galerkin method uses a different test function space. As for the spectral Petrov-Galerkin methods, it is equivalent to the k-C-TFE method if the test space is chosen as P_{N-1} with $N = k$. For the case where the test space is taken as $\bar{X}_N$, the spectral Petrov-Galerkin methods are also different from the discontinuous-time finite element methods in [41, 42]. In other words, the k-C-TFE method in [41,42] is a special case of the spectral Petrov-Galerkin method. Actually, by calculating the right-hand sides of (2.7)-(2.8) and (2.11)-(2.12) by numerical quadrature formulae and following the same argument as that in [41,42], we can relate our spectral methods to some special collocation methods or Runge-Kutta methods.

2.4. Efficient implementation of the algorithm.   To compare the difference among the above three spectral methods, we rewrite numerical schemes (2.7)-(2.8), (2.10) and (2.11)-(2.12) into the matrix form. Noticing that

$$X_N = \operatorname{span}\{\phi_1, \cdots, \phi_N\}$$

due to the homogenous initial condition, we have

$$\hat{p}_{iN} = \sum_{m=1}^{N} a_{i,m}\phi_m, \qquad \hat{q}_{iN} = \sum_{m=1}^{N} b_{i,m}\phi_m,$$

where $a_{i,m}, b_{i,m}$ are constants to be determined. We denote

$$\mathbf{p_i} = (a_{i,1}, \ldots, a_{i,N})^T, \quad \mathbf{q_i} = (b_{i,1}, \ldots, b_{i,N})^T, \quad i = 1, \ldots, n.$$

For the spectral Galerkin method, we use $v_{iN} = \phi_j, w_{iN} = \phi_j$ in (2.7)-(2.8) for all $j = 1, \ldots N$ to obtain

$$\begin{pmatrix} \mathcal{D} & \mathbf{0} \\ \mathbf{0} & \mathcal{D} \end{pmatrix} \begin{pmatrix} \mathbf{p_i} \\ \mathbf{q_i} \end{pmatrix} = \begin{pmatrix} \mathbf{f_i} \\ \mathbf{g_i} \end{pmatrix}, \tag{2.13}$$

where for all $1 \leq i \leq n$,

$$\mathbf{f_i} = (f_{i,1}, \ldots, f_{i,N})^T, \ \mathbf{g_i} = (g_{i,1}, \ldots, g_{i,N})^T, \ f_{i,j} = -(\frac{\partial \hat{H}}{\partial \hat{q}_{iN}}, \phi_j), \ g_{i,j} = (\frac{\partial \hat{H}}{\partial \hat{p}_{iN}}, \phi_j),$$

and

$$\mathcal{D} = (d_{j,m})_{N \times N}, \ d_{j,m} = (\phi_m', \phi_j).$$

By the property of the Lobatto polynomial, we have

$$d_{j,m} = (\phi_m', \phi_j) = \begin{cases} \frac{1}{2}, & m = j = 1, \\ -\frac{1}{3}, & m = 1, j = 2, \\ \frac{1}{3}, & m = 2, j = 1, \\ \frac{2}{(2j-1)(2j+1)}, & m = j+1, j \geq 2, \\ \frac{-2}{(2j-1)(2j-3)}, & m = j-1, j \geq 3, \\ 0, & \text{others.} \end{cases}$$

Therefore, the matrix $\mathcal{D}$ is a tridiagonal matrix.

As for the spectral Petrov-Galerkin method, we note that $L_{j-1}, j = 1, \ldots, N$ is the basis function of the test space P_{N-1}. By taking $v_{iN} = L_{j-1}, w_{iN} = L_{j-1}$ in (2.11)-(2.12) for all $j = 1, \ldots N$, we immediately get (2.13) with $f_{i,j}, g_{i,j}$ replaced by

$$f_{i,j} = -(\frac{\partial \hat{H}}{\partial \hat{q}_{iN}}, L_{j-1}), \ g_{i,j} = (\frac{\partial \hat{H}}{\partial \hat{p}_{iN}}, L_{j-1}),$$

and $\mathcal{D}$ the diagonal matrix, i.e.,

$$\mathcal{D} = (d_{j,m})_{N \times N}, \ d_{j,m} = (\phi_m', L_{j-1}) = \begin{cases} \delta_{m,j}, j = 1, \\ \frac{2\delta_{m,j}}{2j-1}, j \geq 2. \end{cases}$$

Similarly, for the Gauss collocation method, we get the same equation as in (2.13) with the matrix

$$\mathcal{D} = (d_{j,m})_{N \times N}, \quad d_{j,m} = \begin{cases} \frac{1}{2}, m = 1, \\ L_{m-1}(s_j), m \geq 2. \end{cases}$$

In this case, note that

$$(\omega \mathcal{D})^T \mathcal{D} = 2 * diag(\frac{1}{4}, \frac{1}{3}, \cdots, \frac{1}{2N-1}),$$

where $\omega = \text{diag}(\omega_1, \cdots, \omega_N)$ with $\{\omega_j\}_{j=1}^N$ the weights of Gauss numerical quadrature. Therefore, if we use $(\omega \mathcal{D})^T$ to multiply both sides of (2.13) for the Gauss collocation method, the left-hand side will be a diagonal matrix, which is exactly the same as that for the spectral Petrov-Galerkin method, and the right-hand side will be N-point Gauss quadrature. In other words, the spectral Gauss collocation method is equivalent to the spectral Petrov-Galerkin method up to a numerical integration error.

To end this section, we describe an iteration procedure to solve the nonlinear system (2.7)-(2.8) or (2.11)-(2.12) in our numerical experiments. We first divide $\frac{\partial \hat{H}}{\partial \hat{q}_{iN}}$ and $\frac{\partial \hat{H}}{\partial \hat{p}_{iN}}$ into linear and nonlinear parts, i.e., $\frac{\partial \hat{H}}{\partial \hat{p}_{iN}} = \mathcal{L}_{\hat{p}_{iN}} + \mathcal{N}_{\hat{p}_{iN}}$, $\frac{\partial \hat{H}}{\partial \hat{q}_{iN}} = \mathcal{L}_{\hat{q}_{iN}} + \mathcal{N}_{\hat{q}_{iN}}$. Then we set up the following iterative scheme:

$$(\frac{d\hat{p}_{iN}^{m+1}}{dx}, v_{iN}) = -(\mathcal{L}_{\hat{q}_{iN}}^{m+1}, v_{iN}) - (\mathcal{N}_{\hat{q}_{iN}}^m, v_{iN}), \quad i = 1, 2, \cdots, n, \qquad (2.14)$$

$$(\frac{d\hat{q}_{iN}^{m+1}}{dx}, w_{iN}) = (\mathcal{L}_{\hat{p}_{iN}}^{m+1}, w_{iN}) + (\mathcal{N}_{\hat{p}_{iN}}^m, w_{iN}), \quad i = 1, 2, \cdots, n, \qquad (2.15)$$

where

$$\mathcal{L}_{\hat{q}_{iN}}^{m+1} = \mathcal{L}_{\hat{q}_{iN}}(\hat{p}_{1N}^{m+1}, \cdots, \hat{p}_{nN}^{m+1}; \hat{q}_{1N}^{m+1}, \cdots, \hat{q}_{nN}^{m+1}),$$
$$\mathcal{N}_{\hat{q}_{iN}}^m = \mathcal{N}_{\hat{q}_{iN}}(\hat{p}_{1N}^m, \cdots, \hat{p}_{nN}^m; \hat{q}_{1N}^m, \cdots, \hat{q}_{nN}^m).$$

By denoting $\mathbf{p} = (\mathbf{p_1}, \ldots, \mathbf{p_n})^{\mathbf{T}}, \mathbf{q} = (\mathbf{q_1}, \ldots, \mathbf{q_n})^{\mathbf{T}}$ and taking the test function v_{iN}, w_{iN} as basis function, we obtain an equation similar to (2.13):

$$\begin{pmatrix} \mathcal{D} & \mathbf{0} \\ \mathbf{0} & \mathcal{D} \end{pmatrix} \begin{pmatrix} \mathbf{p_i^{m+1}} \\ \mathbf{q_i^{m+1}} \end{pmatrix} + \begin{pmatrix} \mathcal{A}_{i,1} & \cdots & \mathcal{A}_{i,2n} \\ \mathcal{B}_{i,1} & \cdots & \mathcal{B}_{i,2n} \end{pmatrix} \begin{pmatrix} \mathbf{p^{m+1}} \\ \mathbf{q^{m+1}} \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{f}}_{\mathbf{i}}(\mathbf{p^m}; \mathbf{q^m}) \\ \bar{\mathbf{g}}_{\mathbf{i}}(\mathbf{p^m}; \mathbf{q^m}) \end{pmatrix},$$

or equivalently,

$$\begin{pmatrix} \mathcal{M} & \mathbf{0} \\ \mathbf{0} & \mathcal{M} \end{pmatrix} \begin{pmatrix} \mathbf{p^{m+1}} \\ \mathbf{q^{m+1}} \end{pmatrix} + \begin{pmatrix} \mathcal{S}_1 & \mathcal{S}_2 \\ \mathcal{S}_3 & \mathcal{S}_4 \end{pmatrix} \begin{pmatrix} \mathbf{p^{m+1}} \\ \mathbf{q^{m+1}} \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{f}}(\mathbf{p^m}; \mathbf{q^m}) \\ \bar{\mathbf{g}}(\mathbf{p^m}; \mathbf{q^m}) \end{pmatrix}. \qquad (2.16)$$

Here $\mathcal{M} = diag(\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_n) \in R^{nN \times nN}$ is a block diagonal matrix with diagonal element $\mathcal{D}_i = \mathcal{D}$, $\bar{\mathbf{f}} = (\bar{\mathbf{f}}_1, \ldots, \bar{\mathbf{f}}_N)^T$, and $\bar{\mathbf{g}} = (\bar{\mathbf{g}}_1, \ldots, \bar{\mathbf{g}}_N)^T$. On the other hand, by the orthogonality of Legendre and Lobatto polynomials, we have

$$(\phi_m, \phi_j) = \begin{cases} c_m, & m = j, j \pm 1, j \pm 2 \\ 0, & \text{otherwise} \end{cases}, \quad (\phi_m, L_j) = \begin{cases} d_m, & m = j, j+1, j+2 \\ 0, & \text{otherwise} \end{cases}.$$

Here c_m, d_m are constants depending only on m and can be pre-calculated exactly from (2.9) and properties of Legendre polynomials. The coefficient matrices $\mathcal{A}_{i,j}, \mathcal{B}_{i,j}, j =$

$1,\ldots,2n$ from the linear parts are five-diagonal matrices for both spectral Galerkin and Petrov-Galerkin methods. As a consequence, matrices $\mathcal{S}_j, j=1,\ldots,4$ in (2.16) are sparse. Note that the Gauss collocation method is equivalent to the Petrov-Galerkin method up to a Gauss numerical quadrature error. Then similar arguments are still valid for the spectral Gauss collocation method. To compute the nonlinear integration terms appearing in (2.13) or (2.16), we use the $2N$-point Gauss numerical quadrature formulae in our numerical experiments.

3. Conservation properties

We first introduce some notations and error estimates for the Gauss and Guass-Lobatto numerical quadratures.

3.1. Preliminaries and notations.   We shall use the following notations in the rest of this paper:

$$
\begin{aligned}
z &= (p_1,\cdots,p_n;q_1,\cdots,q_n), \quad z_N = (p_{1N},\cdots,p_{nN};q_{1N},\cdots,q_{nN}), \\
\hat{z} &= (\hat{p}_1,\cdots,\hat{p}_n;\hat{q}_1,\cdots,\hat{q}_n), \quad \hat{z}_N = (\hat{p}_{1N},\cdots,\hat{p}_{nN};\hat{q}_{1N},\cdots,\hat{q}_{nN}).
\end{aligned}
\tag{3.1}
$$

Denote by $\omega^{\alpha,\beta}(x) = (1-x)^\alpha (1+x)^\beta$ the Jacobi weight function of index (α,β) on $I = [-1,1]$, and $\|\cdot\|_{\omega^{\alpha,\beta},I}$ the weighted L^2 norm corresponding to $\omega^{\alpha,\beta}$, i.e.,

$$
\|v\|^2_{\omega^{\alpha,\beta},I} = (v,v)_{\omega^{\alpha,\beta}} = \int_{-1}^{1} v^2 \omega^{\alpha,\beta} dx.
$$

We introduce the following non-uniformly weighted Sobolev spaces

$$
H^s_{\omega^{\alpha,\beta},*}(I) := \{u : \partial_x^k u \in L^2_{\omega^{\alpha+k,\beta+k}}, 0 \le k \le s\}
\tag{3.2}
$$

equipped with the inner product and norm

$$
(u,v)_{s,\omega^{\alpha,\beta},*} = \sum_{k=0}^{s} (\partial_x^k u, \partial_x^k v)_{\omega^{\alpha+k,\beta+k}}, \quad \|u\|_{s,\omega^{\alpha,\beta},*} = (u,u)^{\frac{1}{2}}_{s,\omega^{\alpha,\beta},*}.
$$

LEMMA 3.1.   *Let $s_j, \omega_j, j=1,\ldots,N$ be N Gauss points and the corresponding weights in $I = [-1,1]$. There holds for any $u \in H^s_{\omega^{0,0},*}(I)$*

$$
\left| \int_{-1}^{1} u(x)dx - \sum_{j=1}^{N} u(s_j)\omega_j \right| \lesssim N^{-s} \|u\|_{s,\omega^{s,s},*}.
\tag{3.3}
$$

Similarly, by denoting $g_j, \bar{\omega}_j, j=1,\ldots,N+1$ the $N+1$ Gauss-Lobatto points and the corresponding weights on I, we have for all $u \in H^s_{\omega^{-1,-1},*}(I)$,

$$
\left| \int_{-1}^{1} u(x)dx - \sum_{j=1}^{N+1} u(g_j)\bar{\omega}_j \right| \lesssim N^{-s} \|u\|_{s,\omega^{-1+s,-1+s},*}.
\tag{3.4}
$$

Here and in the following, $A \lesssim B$ indicates that A can be bounded by B multiplied by a constant independent of N.

Proof.   We denote by $\mathcal{I}_{N-1} u \in P_{N-1}$ the Gauss interpolation of u. That is,

$$
\mathcal{I}_{N-1} u(s_j) = u(s_j), \quad j=1,\ldots,N.
$$

Since N-point Gauss quadrature is exact for polynomials of degree $2N-1$, then

$$\int_{-1}^{1} u(x)dx - \sum_{j=1}^{N} u(s_j)\omega_j = \int_{-1}^{1} (u - \mathcal{I}_{N-1}u)(x)dx.$$

Note that (see, e.g., [39], Theorem 3.41)

$$\|u - \mathcal{I}_{N-1}u\|_{L^2} \lesssim \sqrt{\frac{(N-s)!}{(N-1)!}}(N+s)^{-(s+1)/2}\|u\|_{s,\omega^{s,s},*} \lesssim N^{-s}\|u\|_{s,\omega^{s,s},*}. \tag{3.5}$$

Here in the second step, we have used the Stirling's formula

$$\sqrt{2\pi}p^{p+\frac{1}{2}} < p!e^p < \sqrt{2\pi}p^{p+\frac{1}{2}}(1+\frac{1}{4p}).$$

Then the desired result (3.3) follows. Similarly, (3.4) follows, by using the fact that the $(N+1)$-point Guass-Lobatto quadrature is exact for polynomials of degree $2N-1$ and the inequality (see, e.g., [39])

$$\|u - \pi_N u\|_{L^2} \lesssim \sqrt{\frac{(N-s+1)!}{N!}}(N+s)^{-(s+1)/2}\|u\|_{s,\omega^{-1+s,-1+s},*}.$$

Here $\pi_N u \in P_N$ denotes the Guass-Lobatto interpolation of u. This finishes our proof.
□

REMARK 3.1.   When the function u is sufficiently smooth, the convergence rate of the Gauss numerical quadrature error in (3.3) can be improved to a hypergeometric convergence rate. Actually, it is shown in [11] (see, p. 98, (2.7.12)),

$$\int_{-1}^{1} u(x)dx - \sum_{j=1}^{N} u(s_j)\omega_j = \frac{2^{2N+1}(N!)^4}{(2N+1)[(2N)!]^3}\partial_x^{2N}u(\xi)$$

for some $\xi \in (-1,1)$. In other words, if $u \in C^{2N}(I)$ or u is an analytic funtion, then we have from the Stirling's formula

$$\left| \int_{-1}^{1} u(x)dx - \sum_{j=1}^{N} u(s_j)\omega_j \right| \lesssim \frac{2^{2N+1}(N!)^4}{(2N+1)[(2N)!]^3} \leq N^{-\frac{1}{2}}\left(\frac{e}{4N}\right)^{2N} = N^{-\frac{1}{2}}e^{-\sigma N}$$

with $\sigma = 2\ln(4N-1) > 0$, which indicates that the Gauss numerical quadrature error is convergent exponentially.

3.2. Energy conserving.   One of the important properties of Hamiltonian systems is that the Hamiltonian is conserved along trajectories, i.e.,

$$H(z(t_M)) = H(z(t_0)), \quad \text{or equivalently, } \hat{H}(\hat{z}(1)) = \hat{H}(\hat{z}(-1)).$$

Consequently, it is natural to expect that numerical methods share the same property. In the following, we shall study the energy preserving properties of spectral Galerkin, Petrov-Galerkin and Gauss collocation methods presented in Section 2.

THEOREM 3.1.   *The Petrov-Galerkin method preserves the energy exactly, i.e.,*

$$\hat{H}(\hat{z}_N(1)) = \hat{H}(\hat{z}_N(-1)). \tag{3.6}$$

*While the spectral Galerkin method and the spectral Gauss collocation method is sepa-
rately energy conserving up to a cell-average error and a Gauss numerical quadrature
error. That is,*

$$\hat{H}(\hat{z}_N(1)) = \hat{H}(\hat{z}_N(-1)) + E_0, \tag{3.7}$$

where for the spectral Galerkin method,

$$E_0 = \sum_{i=1}^{n} \left(\frac{d\hat{p}_{iN}(-1)}{dx} \int_{-1}^{1} \left(\frac{d\hat{q}_{iN}}{dx} - \frac{\partial\hat{H}(\hat{z}_N)}{\partial\hat{p}_{iN}} \right) dx + \frac{d\hat{q}_{iN}(-1)}{dx} \int_{-1}^{1} \left(\frac{\partial\hat{H}(\hat{z}_N)}{\partial\hat{q}_{iN}} + \frac{d\hat{p}_{iN}}{dx} \right) dx \right), \tag{3.8}$$

and for the Gauss collocation method,

$$E_0 = \int_{-1}^{1} \frac{d\hat{H}(\hat{z}_N(x))}{dx} dx - \sum_{j=1}^{N} \frac{d\hat{H}(\hat{z}_N(s_j))}{dx} \omega_j. \tag{3.9}$$

Here s_j, ω_j are N Gauss points and weights in the interval $[-1,1]$.

   Proof.   We first consider the Petrov-Galerkin methods. Note that $\frac{d\hat{q}_{iN}}{dx}, \frac{d\hat{p}_{iN}}{dx} \in P_{N-1}$. By subtracting (2.11) from (2.12), and then taking $(v_{iN}, w_{iN}) = (\frac{d\hat{q}_{iN}}{dx}, \frac{d\hat{p}_{iN}}{dx})$, we have

$$\left(\frac{\partial\hat{H}(\hat{z}_N)}{\partial\hat{q}_{iN}}, \frac{d\hat{q}_{iN}}{dx} \right) + \left(\frac{\partial\hat{H}(\hat{z}_N)}{\partial\hat{p}_{iN}}, \frac{d\hat{p}_{iN}}{dx} \right) = 0.$$

Summing up all i from 1 to n yields

$$\int_{-1}^{1} \frac{d\hat{H}(\hat{z}_N(x))}{dx} dx = 0,$$

and thus (3.6) follows.
   As for the spectral Galerkin method, we first denote

$$c_0 = \frac{d\hat{q}_{iN}(-1)}{dx}, \quad c_1 = \frac{d\hat{p}_{iN}(-1)}{dx}.$$

Note that $\frac{d\hat{q}_{iN}}{dx} - c_0, \frac{d\hat{p}_{iN}}{dx} - c_1 \in X_N$. By choosing $(v_{iN}, w_{iN}) = (\frac{d\hat{q}_{iN}}{dx} - c_0, \frac{d\hat{p}_{iN}}{dx} - c_1)$ in (2.7)-(2.8) and following the same argument as we did for the Petrov-Galerkin methods, we get

$$\left(\frac{\partial\hat{H}(\hat{z}_N)}{\partial\hat{q}_{iN}}, \frac{d\hat{q}_{iN}}{dx} \right) + \left(\frac{\partial\hat{H}(\hat{z}_N)}{\partial\hat{p}_{iN}}, \frac{d\hat{p}_{iN}}{dx} \right)$$
$$- \int_{-1}^{1} \left(c_0 \left(\frac{\partial\hat{H}(\hat{z}_N)}{\partial\hat{q}_{iN}} + \frac{d\hat{p}_{iN}}{dx} \right) - c_1 \left(\frac{\partial\hat{H}(\hat{z}_N)}{\partial\hat{p}_{iN}} - \frac{d\hat{q}_{iN}}{dx} \right) \right) dx = 0.$$

Summing up all i, we get (3.7) immediately.
   We now consider the Gauss collocation method. By using $\frac{d\hat{q}_{iN}(s_j)}{dx} \omega_j$ to multiply both sides of the first equation of (2.10) and summing up all j from 1 to N, we get

$$\sum_{j=1}^{N} \frac{d\hat{p}_{iN}(s_j)}{dx} \frac{d\hat{q}_{iN}(s_j)}{dx} \omega_j = - \sum_{j=1}^{N} \frac{\partial\hat{H}(\hat{z}_N(s_j))}{\partial\hat{q}_{iN}} \frac{d\hat{q}_{iN}(s_j)}{dx} \omega_j.$$

Similarly, we use $\frac{d\hat{p}_{iN}(s_j)}{dx}\omega_j$ to multiply both sides of the second equation of (2.10) and then sum up all j to derive

$$\sum_{j=1}^{N} \frac{d\hat{p}_{iN}(s_j)}{dx} \frac{d\hat{q}_{iN}(s_j)}{dx} \omega_j = \sum_{j=1}^{N} \frac{\partial \hat{H}(\hat{z}_N(s_j))}{\partial \hat{p}_{iN}} \frac{d\hat{p}_{iN}(s_j)}{dx} \omega_j.$$

Consequently,

$$\sum_{j=1}^{N} \frac{d\hat{H}(\hat{z}_N(s_j))}{dx} \omega_j = \sum_{j=1}^{N} \left(\frac{\partial \hat{H}(\hat{z}_N(s_j))}{\partial \hat{q}_{iN}} \frac{d\hat{q}_{iN}(s_j)}{dx} \omega_j + \frac{\partial \hat{H}(\hat{z}_N(s_j))}{\partial \hat{p}_{iN}} \frac{d\hat{p}_{iN}(s_j)}{dx} \omega_j \right) = 0,$$

which yields

$$\int_{-1}^{1} \frac{d\hat{H}(\hat{z}_N)}{dx} dx - E_0 = 0,$$

where E_0 denotes the error of Gauss numerical quadrature given in (3.9). Then (3.7) follows. This finishes our proof. □

REMARK 3.2.   We see that, from (3.3), (3.7), and (3.9), if the function $\hat{H}(\hat{z}(x))$ satisfies some regularity condition, e.g., $\frac{d\hat{H}}{dx} \in H^s_{\omega^{0,0},*}(I), 1 \le s \le N$, the Gauss numerical quadrature error E_0 in (3.9) converges with a spectral accuracy, i.e.,

$$E_0 \lesssim N^{-s}.$$

Furthermore, as indicated by Remark 3.1, if $\hat{H}(\hat{z}(x))$ is sufficiently smooth, the error E_0 in (3.9) converges with a hypergeometric rate. Similarly, for the spectral Galerkin method, as we shall prove in our later analysis (see (4.16) in Corollary 4.1), provided $H(z)$ satisfies some regularity assumption, the cell average error E_0 in (3.8) also converges with a spectral accuracy, i.e.,

$$E_0 \lesssim N^{\frac{3}{2}-s}, \quad 2 \le s \le N+1.$$

In other words, there hold for both spectral Galerkin and Gauss collocation methods,

$$\hat{H}(\hat{z}_N(1)) = \hat{H}(\hat{z}_N(-1)) + O(N^{m-s})$$

with $m=0$ for the Gauss collocation method, and $m=\frac{3}{2}$ for the spectral Galerkin method. Due to the spectral accuracy, we are able to control the energy error to the machine epsilon, i.e., 10^{-15}, with a reasonable N. For example, we consider one step and calculate the energy at time $t=r=1$. Assume that the Hamiltonian function is smooth enough. Then the error in energy decays rapidly with respect to N. For spectral Gauss collocation method, a choice of $N=15$ can ensure the machine epsilon 10^{-15} (noticing that $N^{-N-1} \lesssim 10^{-15}$). As we may observe, the needed local degree freedom is $Nr=15$. On the other hand, if we use the standard 4th-order Runge-Kutta method to achieve the same error accuracy at time $t=r$, the time step size should be $\delta t = 10^{-4}$ and thus the local degree freedom is $O(1/\delta t)=10^4$. In other words, the computational cost of the spectral methods is less than that for the low-order Runge-Kutta method and low-order finite element methods, which makes the spectral methods more efficient. The global error estimate in energy and the error accumulation with respect to time t will be discussed in the later sections.

3.3. Symplecticity.   Another important feature or property of the Hamiltonian system is the symplectic structure, the Jacobi matrix of the transformation $\left(\frac{\partial z}{\partial z_0}\right)$ satisfies

$$\left(\frac{\partial z}{\partial z_0}\right)^T J \left(\frac{\partial z}{\partial z_0}\right) = J, \quad J = \begin{pmatrix} \mathbf{0} & -I_n \\ I_n & \mathbf{0} \end{pmatrix}. \tag{3.10}$$

Here I_n denotes the $n \times n$ identity matrix, and $z_0 = (p_{10}, \ldots, p_{n0}; q_{10}, \ldots, q_{n0})$.

As it is well-known that the Gauss collocation method is symplectic (see, e.g., [30]), we next study the symplectic structure of the spectral Galerkin method and the spectral Petrov-Galerkin method.

THEOREM 3.2.   *For both spectral Galerkin and Petrov-Galerkin methods, the numerical error in symplectic structure decays in the following sense: For any $1 \leq s \leq N$,*

1) if $\nabla \hat{H} \in H^s_{\omega^{0,0},*}(I)$, then for the Petrov-Galerkin method

$$\left(\frac{\partial z_N}{\partial z_0}\right)^T J \left(\frac{\partial z_N}{\partial z_0}\right) = J + O(N^{-s}); \tag{3.11}$$

2) if $\nabla \hat{H} \in H^s_{\omega^{-1,-1},*}(I) \cap C^2(I)$, then for the spectral Galerkin method

$$\left(\frac{\partial z}{\partial z_0}\right)^T J \left(\frac{\partial z_N}{\partial z_0}\right) = J + O(N^{1-s}). \tag{3.12}$$

Proof.

1) Denote by $\psi_j \in P_{N-1}$ the Lagrange basis function associated with the Gauss points $s_j, j \leq N$, i.e.,

$$\psi_i(s_j) = \delta_{i,j}.$$

Taking $(v_{iN}, w_{iN}) = (\psi_j, \psi_j)$ in (2.11)-(2.12), we get

$$\frac{d\hat{p}_{iN}(s_j)}{dx} = -\frac{\partial \hat{H}(\hat{z}_N(s_j))}{\partial \hat{q}_{iN}} - \frac{1}{\omega_j}\bar{E}_{ij}, \quad \frac{d\hat{q}_{iN}(s_j)}{dx} = \frac{\partial \hat{H}(\hat{z}_N(s_j))}{\partial \hat{p}_{iN}} + \frac{1}{\omega_j}\tilde{E}_{ij}. \tag{3.13}$$

where both $\bar{E}_{ij}$ and $\tilde{E}_{ij}$ are Gauss numerical quadrature errors. To be more precise,

$$\bar{E}_{ij} = \left(\frac{\partial \hat{H}(\hat{z}_N)}{\partial \hat{q}_{iN}}, \psi_j\right) - \frac{\partial \hat{H}(\hat{z}_N(s_j))}{\partial \hat{q}_{iN}}\omega_j, \quad \tilde{E}_{ij} = \left(\frac{\partial \hat{H}(\hat{z}_N)}{\partial \hat{p}_{iN}}, \psi_j\right) - \frac{\partial \hat{H}(\hat{z}_N(s_j))}{\partial \hat{p}_{iN}}\omega_j. \tag{3.14}$$

Let

$$y = \frac{\partial \hat{z}_N}{\partial \hat{z}_0}, \quad y_1 = \frac{\partial \hat{z}_N(1)}{\partial \hat{z}_0}, \quad y_0 = \frac{\partial \hat{z}_N(-1)}{\partial \hat{z}_0}.$$

Since $\frac{d}{dx}(y^T J y) = \left(\frac{dy}{dx}\right)^T J y + y^T J \frac{dy}{dx}$, it follows from the fact that the Gauss quadrature is exact for all polynomials of degree no more than $2N-1$

$$y_1^T J y_1 - y_0^T J y_0 = \int_{-1}^1 \frac{d}{dx}(y^T J y)dx = \sum_{j=1}^N \left(\left(\frac{dy}{dx}\right)^T J y + y^T J \frac{dy}{dx}\right)(s_j)\omega_j$$

$$= \sum_{j=1}^{N} \left(\left(y^T (\nabla^2 \hat{H})^T J^{-T} \omega_j + E_j^T \right) J y + y^T J \left(E_j + J^{-1} \nabla^2 \hat{H} y \omega_j \right) \right)(s_j)$$

$$= \sum_{j=1}^{N} \left(E_j^T J y + y^T J E_j \right)(s_j).$$

Here $E_j = \frac{\partial}{\partial z_0}(-\bar{E}_{1j}, \ldots, -\bar{E}_{nj}; \tilde{E}_{1j}, \ldots, \tilde{E}_{nj})^T$, $\nabla^2 \hat{H} = \nabla^2 \hat{H}(\hat{z}_N)$ denotes the Hessian matrix of $H(\hat{z}_N)$, and in the second step, we have used (3.13) and the fact that

$$J^{-T} J = -J J^{-1} = -I_{2n}.$$

In light of the error bound for the Gauss numerical quadrature in (3.3), we have

$$|\bar{E}_{ij}| + |\tilde{E}_{ij}| \lesssim N^{-s}, \quad \forall 1 \leq i \leq n,$$

and thus

$$y_1^T J y_1 = y_0^T J y_0 + O(N^{-s}).$$

Then

$$\left(\frac{\partial \hat{z}_N}{\partial \hat{z}_0} \right)^T J \left(\frac{\partial \hat{z}_N}{\partial \hat{z}_0} \right)(1) = \left(\frac{\partial \hat{z}_N}{\partial \hat{z}_0} \right)^T J \left(\frac{\partial \hat{z}_N}{\partial \hat{z}_0} \right)(-1) + O(N^{-s}).$$

Consequently, for all $t > 0$,

$$\left(\frac{\partial z_N}{\partial z_0} \right)^T J \left(\frac{\partial z_N}{\partial z_0} \right)(t) = \left(\frac{\partial z_N}{\partial z_0} \right)^T J \left(\frac{\partial z_N}{\partial z_0} \right)(0) + O(t N^{-s}) = J + O(N^{-s}).$$

Here in the last step, we have used the fact that $\frac{\partial z_N}{\partial z_0}$ is the identity matrix at time $t = 0$. Then (3.11) follows for the Petrov-Galerkin method.

2) We first denote

$$-1 = g_0 < g_1 < \ldots < g_N = 1$$

the $N+1$ Gauss-Lobatto points, i.e., zeros of the Lobatto polynomial ϕ_{N+1}, and $\psi_j \in P_N$ the Lagrange basis function associated with the Gauss-Lobatto points $g_j, j = 0, \ldots, N$. Note that $\psi_j \in X_N$ for all $j = 1, \ldots, N$. We choose $(v_{iN}, w_{iN}) = (\psi_j, \psi_j), j = 1, \ldots, N$ in (2.7)-(2.8) to obtain the same result as in (3.13) with the Gauss points s_j replaced by Gauss-Lobatto points g_j for all j from 1 to N. At the boundary point $x = -1$, we have

$$\frac{d\hat{p}_{iN}(-1)}{dx} = -\frac{\partial \hat{H}(\hat{z}_N(-1))}{\partial \hat{q}_{iN}} - \bar{E}_{i0}, \quad \frac{d\hat{q}_{iN}(-1)}{dx} = \frac{\partial \hat{H}\hat{z}_N(-1)}{\partial \hat{p}_{iN}} + \tilde{E}_{i0}$$

with

$$\bar{E}_{i0} = \frac{d\hat{p}_i(-1)}{dx} - \frac{d\hat{p}_{iN}(-1)}{dx} + \frac{\partial \hat{H}(\hat{z}(-1))}{\partial \hat{q}_i} - \frac{\partial \hat{H}(\hat{z}_N(-1))}{\partial \hat{q}_{iN}}, \tag{3.15}$$

$$\tilde{E}_{i0} = \frac{d\hat{q}_{iN}(-1)}{dx} - \frac{d\hat{q}_i(-1)}{dx} + \frac{\partial \hat{H}(\hat{z}(-1))}{\partial \hat{p}_i} - \frac{\partial \hat{H}(\hat{z}_N(-1))}{\partial \hat{p}_{iN}}. \tag{3.16}$$

Combining (3.13) with (3.15)-(3.16), we have for all $j = 0, \ldots, N$,

$$\frac{d\hat{p}_{iN}(g_j)}{dx} = -\frac{\partial \hat{H}(\hat{z}_N(g_j))}{\partial \hat{q}_{iN}} - \frac{1}{\bar{\omega}_j}\bar{E}_{ij}, \qquad \frac{d\hat{q}_{iN}(g_j)}{dx} = \frac{\partial \hat{H}(\hat{z}_N(g_j))}{\partial \hat{p}_{iN}} + \frac{1}{\bar{\omega}_j}\tilde{E}_{ij}.$$

Here $\{\bar{\omega}_j\}_{j=0}^N$ are weights of the Gauss-Lobatto numerical quadrature on $[-1,1]$, and

$$\bar{E}_{ij} = \left(\frac{\partial \hat{H}(\hat{z}_N)}{\partial \hat{q}_{iN}}, \psi_j\right) - \frac{\partial \hat{H}(\hat{z}_N(g_j))}{\partial \hat{q}_{iN}}\bar{\omega}_j, \quad \tilde{E}_{ij} = \left(\frac{\partial \hat{H}(\hat{z}_N)}{\partial \hat{p}_{iN}}, \psi_j\right) - \frac{\partial \hat{H}(\hat{z}_N(g_j))}{\partial \hat{p}_{iN}}\bar{\omega}_j$$

denote the errors of the Gauss-Lobatto numerical quadrature. Recalling the error for the Gauss-Lobatto numerical quadrature in (3.4), we derive

$$|\bar{E}_{ij}| + |\tilde{E}_{ij}| \lesssim N^{-s}, \quad 1 \leq i \leq n.$$

On the other hand, we have the following estimates (see (4.16) in Corollary 4.1),

$$|\bar{E}_{i0}| + |\tilde{E}_{i0}| \lesssim N^{1-s},$$

then (3.12) follows by applying the same argument as what we have done for the Petrov-Galerkin method. The proof is complete. □

To end this section, we would like to point out that the measurement of error in symplectic structure for spectral Galerkin methods is different from that for the standard h-version method (i.e., Runge-Kutta methods or low-order finite element methods). It is well known that any numerical method of order N can preserve the symplectic structure up to N-th order in the sense of h-version. In other words, the hidden constant in the error bound may depend upon the polynomial N since the polynomial degree N is fixed in the h-version. However, for the spectral method, the error bounds in (3.11)-(3.12) are measured by the polynomial degree N and the hidden constants are independent of N.

4. Error estimates

In this section, we shall provide error estimates for the domain decomposition spectral Galerkin method and spectral Petrov-Galerkin method. As for the spectral Gauss-collocation method, we refer to [24] for more detailed information.

Given any interval $[a,b]$, we denote by $\tilde{w}^{\alpha,\beta}(t)$ the Jacobi weight function of index (α,β) on $[a,b]$. That is,

$$\tilde{w}^{\alpha,\beta}(t) = w^{\alpha,\beta}(x) = w^{\alpha,\beta}\left(\frac{2t-a-b}{b-a}\right), \quad x \in [-1,1].$$

We define the weighted L^2 norm corresponding to $\tilde{\omega}^{\alpha,\beta}$ on $[a,b]$ as follows

$$\|v\|_{\omega^{\alpha,\beta},[a,b]}^2 = \int_a^b v^2 \tilde{w}^{\alpha,\beta} dt = \frac{b-a}{2}\int_{-1}^1 \hat{v}^2 \omega^{\alpha,\beta} dx, \quad \hat{v}(x) = \hat{v}\left(\frac{2t-a-b}{b-a}\right) = v(t).$$

Now we define the weighted Sobolev spaces on any interval $[a,b]$ by

$$H_{\omega^{\alpha,\beta}}^s([a,b]) := \left\{u(t): \|\partial_t^k u\|_{\omega^{\alpha+k,\beta+k},[a,b]} \lesssim 1, \quad 0 \leq k \leq s\right\}$$

equipped with the norm

$$\|u\|_{s,w^{\alpha,\beta},[a,b]} = \sum_{k=0}^s \|\partial_t^k u\|_{\omega^{\alpha+k,\beta+k},[a,b]}.$$

We have the following error estimates.

THEOREM 4.1.   *Given any $t > 0$, let $p_i, q_i \in H^s_{\omega^{-1},-1}([0,t]), 2 \le s \le N+1$ be solutions of (2.1)-(2.2), and (p_{iN}, q_{iN}) be numerical solutions of (p_i, q_i) for all $i \le n$, which are obtained from the spectral Galerkin method (see (2.7)-(2.8)) or the Petrov-Galerkin method (see (2.11)-(2.12)). Suppose $H(z) \in C^3$. Then at any time $t_m = mr$,*

$$\left(\sum_{i=1}^n \left((p_i - p_{iN})^2 + (q_i - q_{iN})^2\right)(t_m)\right)^{\frac{1}{2}} \lesssim N^{l-s} c_{p,q}, \tag{4.1}$$

$$\left(\sum_{i=1}^n \|p_i - p_{iN}\|^2_{L^2(0,t_m)} + \|q_i - q_{iN}\|^2_{L^2(0,t_m)}\right)^{\frac{1}{2}} \lesssim t_m^{\frac{1}{2}} N^{l-s} c_{p,q}. \tag{4.2}$$

Here $l = 1$ for spectral Petrov-Galerkin method and $l = \frac{3}{2}$ for spectral Galerkin method, and

$$c_{p,q} = \left(\sum_{k=1}^m \sum_{i=1}^n (\|\partial_t^s p_i\|^2_{\omega^{-1+s,-1+s},[t_{k-1},t_k]} + \|\partial_t^s q_i\|^2_{\omega^{-1+s,-1+s},[t_{k-1},t_k]})\right)^{\frac{1}{2}}.$$

Proof.   First, let

$$\hat{\xi}_{v_i} = \pi_N \hat{v}_i - \hat{v}_{iN}, \quad \hat{\eta}_{v_i} = \hat{v}_i - \pi_N \hat{v}_i, \quad v = p \text{ or } q.$$

Here $\pi_N v \in X_N$ denotes the Gauss-Lobatto interpolant of v.

Spectral Petrov-Galerkin methods: For all $j = 1, \dots N$, let $\psi_j \in P_{N-1}$ be the Lagrange interpolation function associated with the N Gauss points s_j. Noticing that the exact solution $(\hat{p}_i, \hat{q}_i)$ also satisfies (2.11)-(2.12), then

$$(\frac{d\hat{\xi}_{p_i}}{dx}, \psi_j) + (\frac{d\hat{\eta}_{p_i}}{dx}, \psi_j) = (\frac{\partial \hat{H}(\hat{z}_N)}{\partial \hat{q}_{iN}} - \frac{\partial \hat{H}(\hat{z})}{\partial \hat{q}_i}, \psi_j) = (\tilde{\pi}_{N-1}(\frac{\partial \hat{H}(\hat{z}_N)}{\partial \hat{q}_{iN}} - \frac{\partial \hat{H}(\hat{z})}{\partial \hat{q}_i}), \psi_j).$$

Here $\tilde{\pi}_N$ denotes the L^2 projection onto P_N, and $\hat{z}, \hat{z}_N$ are given in (3.1). Since N points Gauss numerical quadrature is exact for all polynomials of degree not more than $2N - 1$, we have

$$\frac{d\hat{\xi}_{p_i}(s_j)}{dx} \omega_j = \sum_{m=1}^N (\frac{d\hat{\xi}_{p_i}}{dx} \psi_j)(s_m) \omega_m$$

$$= -\tilde{\pi}_{N-1} \frac{d\hat{\eta}_{p_i}}{dx}(s_j) \omega_j + \tilde{\pi}_{N-1}(\frac{\partial \hat{H}(\hat{z}_N)}{\partial \hat{q}_{iN}} - \frac{\partial \hat{H}(\hat{z})}{\partial \hat{q}_i})(s_j) \omega_j.$$

Multiplying both sides of the above equation by $\hat{\xi}_{p_i}(s_j)$ and then summing up all j yields

$$(\frac{d\hat{\xi}_{p_i}}{dx}, \hat{\xi}_{p_i}) = -(\tilde{\pi}_{N-1} \frac{d\hat{\eta}_{p_i}}{dx}, \hat{\xi}_{p_i}) + (\tilde{\pi}_{N-1}(\frac{\partial \hat{H}(\hat{z}_N)}{\partial \hat{q}_{iN}} - \frac{\partial \hat{H}(\hat{z})}{\partial \hat{q}_i}), \hat{\xi}_{p_i}). \tag{4.3}$$

By the Taylor expansion, we have

$$\frac{\partial \hat{H}(\hat{z})}{\partial \hat{q}_i} - \frac{\partial \hat{H}(\hat{z}_N)}{\partial \hat{q}_{iN}} = \sum_{j=1}^n (\frac{\partial^2 \hat{H}}{\partial \hat{p}_j \partial \hat{q}_i})(\zeta)(\hat{p}_j - \hat{p}_{jN}) + \sum_{j=1}^n (\frac{\partial^2 \hat{H}}{\partial \hat{q}_j \partial \hat{q}_i})(\zeta)(\hat{q}_j - \hat{q}_{jN}), \tag{4.4}$$

where $\zeta = \theta\hat{z} + (1-\theta)\hat{z}_N, \theta \in [0,1]$. Therefore, if $H \in C^2$, we have

$$\frac{\partial^2 \hat{H}}{\partial \hat{p}_j \partial q_i} + \frac{\partial^2 \hat{H}}{\partial \hat{q}_j \partial q_i} = \frac{r}{2}\left(\frac{\partial^2 H}{\partial p_j \partial q_i} + \frac{\partial^2 H}{\partial q_j \partial q_i}\right) = O(r).$$

Substituting the Taylor expansion (4.4) into (4.3) and using the Cauchy-Schwarz inequality, we get

$$\hat{\xi}_{p_i}^2(1) - \hat{\xi}_{p_i}^2(-1) = 2\left(\frac{d\hat{\xi}_{p_i}}{dx}, \hat{\xi}_{p_i}\right) \leq \frac{c_0}{r}\|\partial_x \hat{\eta}_{p_i}\|_{L^2}^2 + r\|\hat{\xi}_{p_i}\|_{L^2}^2 + c_1 r\|\hat{z} - \hat{z}_N\|_{L^2}^2.$$

Here both c_0 and c_1 are constants independent of N and r, and

$$\|\hat{z} - \hat{z}_N\|_{L^2}^2 = \sum_{i=1}^{n}\left(\|\hat{p}_i - \hat{p}_{iN}\|_{L^2}^2 + \|\hat{q}_i - \hat{q}_{iN}\|_{L^2}^2\right)$$

$$\leq 2\sum_{i=1}^{n}\left(\|\hat{\xi}_{p_i}\|_{L^2}^2 + \|\hat{\xi}_{q_i}\|_{L^2}^2 + \|\hat{\eta}_{p_i}\|_{L^2}^2 + \|\hat{\eta}_{q_i}\|_{L^2}^2\right).$$

Summing up all i from 1 to n and using the homogeneous boundary condition yields

$$\sum_{i=1}^{n}\hat{\xi}_{p_i}^2(1) \leq (1 + 2c_1 n)r\sum_{i=1}^{n}\left(\|\hat{\xi}_{p_i}\|_{L^2}^2 + \|\hat{\xi}_{q_i}\|_{L^2}^2\right)$$

$$+ \sum_{i=1}^{n}\left(\frac{c_0}{r}\|\partial_x \hat{\eta}_{p_i}\|_{L^2}^2 + 2c_1 nr\|\hat{\eta}_{p_i}\|_{L^2}^2 + 2c_1 nr\|\hat{\eta}_{q_i}\|_{L^2}^2\right).$$

Similarly, we can derive that

$$\sum_{i=1}^{n}\hat{\xi}_{q_i}^2(1) \leq (1 + c_1 n)r\sum_{i=1}^{n}\left(\|\hat{\xi}_{p_i}\|_{L^2}^2 + \|\hat{\xi}_{q_i}\|_{L^2}^2\right)$$

$$+ \sum_{i=1}^{n}\left(\frac{c_0}{r}\|\partial_x \hat{\eta}_{q_i}\|_{L^2}^2 + 2c_1 nr\|\hat{\eta}_{p_i}\|_{L^2}^2 + 2c_1 nr\|\hat{\eta}_{q_i}\|_{L^2}^2\right).$$

Note that for all $u \in H_{\omega^{-1},-1,*}^s, s \leq N+1$, the Gauss-Lobatto interpolation has the following approximation property (see, e.g., [39])

$$\|u - \pi_N u\|_{L^2} + N^{-1}\|\partial_x(u - \pi_N u)\|_{L^2} \lesssim \sqrt{\frac{(N-s+1)!}{N!}}(N+s)^{-(s+1)/2}\|u\|_{s,\omega^{-1+s,-1+s,*}}$$

$$\lesssim N^{-s}\|u\|_{s,\omega^{-1+s,-1+s,*}}.$$

Then

$$\sum_{i=1}^{n}(\hat{\xi}_{p_i}^2 + \hat{\xi}_{q_i}^2)(1) \leq c_2 r\sum_{i=1}^{n}\left(\|\hat{\xi}_{p_i}\|_{L^2}^2 + \|\hat{\xi}_{q_i}\|_{L^2}^2\right) + c_3 N^{2(1-s)}r^{-1}(\hat{c}_{p,q})^2, \qquad (4.5)$$

where c_2 is a constant dependent only on c_0, c_1 and n, and

$$\hat{c}_{p,q} = \sum_{i=1}^{n}(\|\partial_x^s \hat{p}_i\|_{\omega^{-1+s,-1+s,*}}^2 + \|\partial_x^s \hat{q}_i\|_{\omega^{-1+s,-1+s,*}}^2). \qquad (4.6)$$

On the other hand, by choosing $(v_{iN}, w_{iN}) = (\frac{d\hat{\xi}_{p_i}}{dx}, \frac{d\hat{\xi}_{q_i}}{dx})$ in (2.11)- (2.12) and using the orthogonality, we get

$$\|\frac{d\hat{\xi}_{p_i}}{dx}\|_{L^2}^2 = -(\frac{d\hat{\eta}_{p_i}}{dx}, \frac{d\hat{\xi}_{p_i}}{dx}) + (\frac{\partial \hat{H}(\hat{z}_N)}{\partial \hat{q}_{iN}} - \frac{\partial \hat{H}(\hat{z})}{\partial \hat{q}_i}, \frac{d\hat{\xi}_{p_i}}{dx})$$

$$\lesssim \left(N^{1-s}\|\partial_x^s \hat{p}_i\|_{\omega^{-1+s,-1+s},*} + r\|\hat{z} - \hat{z}_N\|_{L^2}\right) \|\frac{d\hat{\xi}_{p_i}}{dx}\|_{L^2}.$$

Following the same argument, we have a similar estimate for $\|\frac{d\hat{\xi}_{q_i}}{dx}\|_{L^2}^2$. Consequently, there exists a constant c such that

$$\sum_{i=1}^n \left(\|\frac{d\hat{\xi}_{q_i}}{dx}\|_{L^2}^2 + \|\frac{d\hat{\xi}_{q_i}}{dx}\|_{L^2}^2\right) \le c\left(N^{2(1-s)}(\hat{c}_{p,q})^2 + r^2\|\hat{z} - \hat{z}_N\|_{L^2}^2\right). \tag{4.7}$$

Then for any $x \in (-1, 1)$,

$$\sum_{i=1}^n (\hat{\xi}_{p_i}^2 + \hat{\xi}_{q_i}^2)(x) = \sum_{i=1}^n (\hat{\xi}_{p_i}^2 + \hat{\xi}_{q_i}^2)(1) - \sum_{i=1}^n \int_x^1 \frac{d}{dx}(\hat{\xi}_{p_i}^2 + \hat{\xi}_{q_i}^2) dx$$

$$\le \sum_{i=1}^n \left(\hat{\xi}_{p_i}^2(1) + \hat{\xi}_{q_i}^2(1) + \frac{1}{4}\|\hat{\xi}_{p_i}\|_{L^2}^2 + \frac{1}{4}\|\hat{\xi}_{q_i}\|_{L^2}^2 + 4\|\partial_x \hat{\xi}_{p_i}\|_{L^2}^2 + 4\|\partial_x \hat{\xi}_{q_i}\|_{L^2}^2\right)$$

$$\le \sum_{i=1}^n \left(\hat{\xi}_{p_i}^2(1) + \hat{\xi}_{q_i}^2(1) + \frac{1}{4}\|\hat{\xi}_{p_i}\|_{L^2}^2 + \frac{1}{4}\|\hat{\xi}_{q_i}\|_{L^2}^2\right)$$

$$+ 4cN^{2(1-s)}(\hat{c}_{p,q})^2 + 4cr^2\|\hat{z} - \hat{z}_N\|_{L^2}^2,$$

where in the second and third steps, we have used the Cauchy-Schwarz inequality and (4.7), respectively. Integrating the above inequality with x on $[-1, 1]$, we get

$$\sum_{i=1}^n \left(\|\hat{\xi}_{p_i}\|_{L^2}^2 + \|\hat{\xi}_{q_i}\|_{L^2}^2\right) \le 4\sum_{i=1}^n (\hat{\xi}_{p_i}^2 + \hat{\xi}_{q_i}^2)(1) + 16cN^{2(1-s)}(\hat{c}_{p,q})^2 + 16cr^2\|\hat{z} - \hat{z}_N\|_{L^2}^2, \tag{4.8}$$

and thus

$$(1 - 16cr^2)\sum_{i=1}^n \left(\|\hat{\xi}_{p_i}\|_{L^2}^2 + \|\hat{\xi}_{q_i}\|_{L^2}^2\right) \le 4\sum_{i=1}^n (\hat{\xi}_{p_i}^2 + \hat{\xi}_{q_i}^2)(1) + c_4 N^{2(1-s)}(\hat{c}_{p,q})^2. \tag{4.9}$$

Therefore, if the time step size r satisfies

$$16cr^2 + 4rc_2 < 1, \tag{4.10}$$

with c, c_2 exactly the same constants given in (4.5) and (4.7), then we have, by substituting the above inequality into (4.5)

$$(1 - \frac{4rc_2}{1 - 16cr^2})\sum_{i=1}^n (\hat{\xi}_{p_i}^2 + \hat{\xi}_{q_i}^2)(1) \lesssim r^{-1} N^{2(1-s)}(\hat{c}_{p,q})^2.$$

By a scaling from $[-1, 1]$ to $\tau_m = [t_{m-1}, t_m]$ with $m \ge 1$, we immediately get

$$\sum_{i=1}^n \left(\xi_{p_i}^2 + \xi_{q_i}^2\right)(t_m) - \sum_{i=1}^n \left(\xi_{p_i}^2 + \xi_{q_i}^2\right)(t_{m-1})$$

$$\lesssim N^{2(1-s)} \left(\frac{r}{2}\right)^{2(s-1)} \sum_{i=1}^{n} (\|\partial_t^s p_i\|_{\omega^{-1+s,-1+s},\tau_m}^2 + \|\partial_t^s q_i\|_{\omega^{-1+s,-1+s},\tau_m}^2).$$

Noticing that ξ_{p_i}, ξ_{q_i} are continuous, then

$$\sum_{i=1}^{n} \left(\xi_{p_i}^2 + \xi_{q_i}^2\right)(t_m) - \sum_{i=1}^{n} \left(\xi_{p_i}^2 + \xi_{q_i}^2\right)(0) \lesssim N^{2(1-s)} \left(\frac{r}{2}\right)^{2(s-1)} (c_{p,q})^2.$$

Then the desired result (4.1) follows from the triangle inequality and the approximation property of the Gauss-Lobatto interpolation function. In light of (4.9), we have

$$\sum_{i=1}^{n} \int_{t_{m-1}}^{t_m} (\xi_{p_i}^2 + \xi_{q_i}^2) dx \lesssim r^{-1} \sum_{i=1}^{n} \left(\xi_{p_i}^2 + \xi_{q_i}^2\right)(t_m) + r^{-1} N^{2(1-s)} (\hat{c}_{p,q})^2,$$

which yields, together with (4.1)

$$\sum_{i=1}^{n} \int_{0}^{t_m} (\xi_{p_i}^2 + \xi_{q_i}^2) dx \lesssim m N^{2(1-s)} (c_{p,q})^2.$$

Then (4.2) follows.

Spectral Galerkin methods: In light of (2.7)-(2.8), we have

$$\left(\frac{d\hat{\xi}_{p_i}}{dx}, v_{iN}\right) + \left(\frac{d\hat{\eta}_{p_i}}{dx}, v_{iN}\right) = \left(\frac{d(\hat{p}_i - \hat{p}_{iN})}{dx}, v_{iN}\right) = \left(\frac{\partial \hat{H}(\hat{z}_N)}{\partial \hat{q}_{iN}} - \frac{\partial \hat{H}(\hat{z})}{\partial \hat{q}_i}, v_{iN}\right), \quad \forall v_{iN} \in X_N. \tag{4.11}$$

By choosing $v_{iN} = \hat{\xi}_{p_i}$ in the above equation and following the same argument as what we did for the spectral Petrov-Galerkin method, we get the same inequality (4.5) for spectral Galerkin methods.

We next estimate $\|\partial_x \hat{\xi}_{p_i}\|_{L^2}$. In light of (4.11), there holds

$$\left(\frac{d\hat{\xi}_{p_i}}{dx}, v_{iN}\right) = -\left(\tilde{\pi}_N \frac{d\hat{\eta}_{p_i}}{dx}, v_{iN}\right) + \left(\tilde{\pi}_N \left(\frac{\partial \hat{H}(\hat{z}_N)}{\partial \hat{q}_{iN}} - \frac{\partial \hat{H}(\hat{z})}{\partial \hat{q}_i}\right), v_{iN}\right), \quad \forall v_{iN} \in X_N. \tag{4.12}$$

Now we suppose that there exist constants a_j, b_j, d_j such that

$$\frac{d\hat{\xi}_{p_i}}{dx} = \sum_{j=0}^{N-1} a_j L_j, \quad \tilde{\pi}_N \frac{d\hat{\eta}_{p_i}}{dx} = \sum_{j=0}^{N} b_j L_j, \quad \tilde{\pi}_N \left(\frac{\partial \hat{H}(\hat{z}_N)}{\partial \hat{q}_{iN}} - \frac{\partial \hat{H}(\hat{z})}{\partial \hat{q}_i}\right) = \sum_{j=0}^{N} d_j L_j. \tag{4.13}$$

Since $(L_j + (-1)^{j+1+N} L_N) \in X_N$ for all $j = 0, \ldots, N-1$, we choose $v_{iN} = L_j + (-1)^{N+j+1} L_N$ in (4.12) and using the orthogonality of the Legendre polynomial to get

$$a_j = d_j - b_j + (-1)^{N+j+1} \frac{(d_N - b_N)(2j+1)}{2N+1}, \quad j = 0, \ldots, N-1. \tag{4.14}$$

and thus,

$$\|\frac{d\hat{\xi}_{p_i}}{dx}\|_{L^2}^2 = \sum_{j=0}^{N-1} \frac{2a_j^2}{2j+1} \lesssim \sum_{j=0}^{N-1} \left(\frac{b_j^2 + d_j^2}{2j+1} + \frac{(b_N^2 + d_N^2)(2j+1)}{(2N+1)^2}\right)$$

$$\lesssim \|\frac{d\hat{\eta}_{p_i}}{dx}\|_{L^2}^2 + \|\frac{\partial \hat{H}(\hat{z}_N)}{\partial \hat{q}_{iN}} - \frac{\partial \hat{H}(\hat{z})}{\partial \hat{q}_i}\|_{L^2}^2 + b_N^2 + d_N^2$$

$$\lesssim N^{2(1-s)}\|\partial_x^s \hat{p}_i\|_{\omega^{-1+s,-1+s,*}}^2 + r^2\|\hat{z} - \hat{z}_N\|_{L^2}^2 + b_N^2 + d_N^2. \qquad (4.15)$$

Note that

$$b_N = \frac{(\pi_N \frac{d\hat{\eta}_{p_i}}{dx}, L_N)}{(L_N, L_N)} = \frac{2N+1}{2}\int_{-1}^1 \frac{d\hat{\eta}_{p_i}}{dx} L_N dx = \frac{2N+1}{2}\int_{-1}^1 \frac{d\hat{p}_i}{dx} L_N dx$$

$$= -\frac{2N+1}{2}\int_{-1}^1 \frac{d^2\hat{p}_i}{dx^2}\phi_{N+1}dx = -\frac{1}{2}\int_{-1}^1 (\frac{d^2\hat{p}_i}{dx^2} - \pi_{N-2}\frac{d^2\hat{p}_i}{dx^2})(L_{N+1} - L_{N-1})dx,$$

where ϕ_N is the Lobatto polynomial of degree N, and in the last three steps, we have used the integration by parts, the orthogonality of Legendre polynomials. Consequently,

$$|b_N|^2 \lesssim N^{-1}\|\frac{d^2\hat{p}_i}{dx^2} - \pi_{N-2}\frac{d^2\hat{p}_i}{dx^2}\|_{L^2}^2 \lesssim N^{3-2s}\|\partial_x^s \hat{p}_i\|_{\omega^{-1+s,-1+s,*}}, \quad 2 \le s \le N+1.$$

Similarly, there holds

$$d_N = \frac{2N+1}{2}(\frac{\partial \hat{H}(\hat{z}_N)}{\partial \hat{q}_{iN}} - \frac{\partial \hat{H}(\hat{z})}{\partial \hat{q}_i}, L_N) = -\frac{1}{2}(\partial_x(\frac{\partial \hat{H}(\hat{z}_N)}{\partial \hat{q}_{iN}} - \frac{\partial \hat{H}(\hat{z})}{\partial \hat{q}_i}), L_{N+1} - L_{N-1}).$$

Therefore, if $H \in C^3$, we have from the Taylor expansion that

$$|d_N|^2 \lesssim N^{-1}r^2(\|\hat{z} - \hat{z}_N\|^2 + \|\partial_x(\hat{z} - \hat{z}_N)\|^2).$$

By substituting the estimates of b_N and d_N into (4.15) and summing up all i, there exists a constant c independent of N, r such that

$$\sum_{i=1}^n \|\frac{d\hat{\xi}_{p_i}}{dx}\|_{L^2}^2 \le c\left(N^{3-2s}(\hat{c}_{p,q})^2 + r^2\|\hat{z} - \hat{z}_N\|_{L^2}^2 + N^{-1}r^2\sum_{i=1}^n (\|\frac{d\hat{\xi}_{q_i}}{dx}\|_{L^2}^2 + \|\frac{d\hat{\xi}_{q_i}}{dx}\|_{L^2}^2)\right).$$

The similar error estimate also holds for $\|\frac{d\hat{\xi}_{q_i}}{dx}\|_{L^2}^2$. Consequently, if $N^{-1}r$ is chosen such that $cN^{-1}r \le \frac{1}{8}$, then

$$\sum_{i=1}^n \left(\|\frac{d\hat{\xi}_{q_i}}{dx}\|_{L^2}^2 + \|\frac{d\hat{\xi}_{q_i}}{dx}\|_{L^2}^2\right) \lesssim N^{3-2s}(\hat{c}_{p,q})^2 + r^2\|\hat{z} - \hat{z}_N\|_{L^2}^2$$

The rest of the proof is similar to that for the spectral Petrov-Galerkin method. This finishes our proof.   □

REMARK 4.1.   If $\frac{d^s p_i}{dt^s}, \frac{d^s q_i}{dt^s} \in L^\infty(0,T)$ for all $1 \le i \le n$, then the constant $c_{p,q}$ in (4.1) can be bounded by $T^{\frac{1}{2}}$. In other words, we have from (4.1)-(4.2),

$$\left(\sum_{i=1}^n ((p_i - p_{iN})^2 + (q_i - q_{iN})^2)(T)\right)^{\frac{1}{2}} = O(T^{\frac{1}{2}}N^{l-s}),$$

$$\left(\sum_{i=1}^n \|p_i - p_{iN}\|_{L^2(0,T)}^2 + \|q_i - q_{iN}\|_{L^2(0,T)}^2\right)^{\frac{1}{2}} = O(TN^{l-s}).$$

Note that the L^2 error is growing linearly with respect to the time t, which is consistent with that of Gauss collocation methods in [24].

REMARK 4.2.   The condition imposed on the step size r in (4.10) is similar to the stable condition established in [1]. The constants c and c_2 in (4.10) are dependent on n and the specific properties of the Hamiltonian function. We refer to [1] for more detailed information about the choice of r and the discussion of stability of the algorithm.

With error estimates in the above theorem, we can prove that all errors given in (3.8), (3.15)-(3.16) have the spectral accuracy.

COROLLARY 4.1.   *Suppose that all the conditions of Theorem 4.1 hold with $\frac{5}{2} \leq s \leq N+1$. Let E_0, $\bar{E}_{i0}$ and $\tilde{E}_{i0}$ be given by (3.8), (3.15)-(3.16), respectively. Then*

$$|E_0| + |\bar{E}_{i0}| + |\tilde{E}_{i0}| \lesssim N^{\frac{3}{2}-s} \sum_{i=1}^{n} (\|\partial_x^s \hat{p}_i\|_{\omega^{-1+s,-1+s},*} + \|\partial_x^s \hat{q}_i\|_{\omega^{-1+s,-1+s},*}). \qquad (4.16)$$

Proof.   We only consider the error E_0 since the same argument can be applied to the other two errors.

In light of (4.13)-(4.14), we have

$$|\partial_x \hat{\xi}_{p_i}(-1)| = \left| \sum_{j=0}^{N-1} (-1)^j a_j \right| = \left| \sum_{j=0}^{N-1} (-1)^j (d_j - b_j) + (-1)^{N+1} \frac{(d_N - b_N)(2j+1)}{2N+1} \right|$$

$$\lesssim \|\hat{\eta}_{p_i}\|_{L^\infty} + \|\frac{\partial \hat{H}(\hat{z}_N)}{\partial \hat{q}_{iN}} - \frac{\partial \hat{H}(\hat{z})}{\partial \hat{q}_i}\|_{L^\infty} + N|b_N| + N|d_N| \lesssim N^{\frac{5}{2}-s},$$

where in the last step, we have used the estimates of b_N, d_N in Theorem 4.1. Consequently, for $s \geq \frac{5}{2}$,

$$|\partial_x \hat{\xi}_{p_i}(-1)| \leq \|\partial_x (\pi_N \hat{p}_i)\|_{L^\infty} + \|\partial_x \hat{\xi}_{p_i}\|_{L^\infty} \lesssim 1.$$

On the other hand, note that

$$\int_{-1}^{1} \left(\frac{\partial \hat{H}(\hat{z}_N)}{\partial \hat{p}_{iN}} - \frac{d\hat{q}_{iN}}{dx} \right) = \int_{-1}^{1} \left(\frac{\partial \hat{H}(\hat{z}_N)}{\partial \hat{p}_{iN}} - \frac{\partial \hat{H}(\hat{z})}{\partial \hat{p}_i} + \frac{d\hat{q}_i}{dx} - \frac{d\hat{q}_{iN}}{dx} \right) dx$$

$$\lesssim \|\hat{z} - \hat{z}_N\|_{L^2} + |(\hat{q}_i - \hat{q}_{iN})(1)|$$

$$\lesssim N^{\frac{3}{2}-s} \sum_{i=1}^{n} (\|\partial_x^s \hat{p}_i\|_{\omega^{-1+s,-1+s},*} + \|\partial_x^s \hat{q}_i\|_{\omega^{-1+s,-1+s},*}).$$

Then

$$\left| \frac{d\hat{p}_{iN}(-1)}{dx} \int_{-1}^{1} \left(\frac{\partial \hat{H}(\hat{z}_N)}{\partial \hat{p}_{iN}} - \frac{d\hat{q}_{iN}}{dx} \right) \right| \lesssim N^{\frac{3}{2}-s} \sum_{i=1}^{n} (\|\partial_x^s \hat{p}_i\|_{\omega^{-1+s,-1+s},*} + \|\partial_x^s \hat{q}_i\|_{\omega^{-1+s,-1+s},*}).$$

Similarly, we have the same error bound for the term $\left| \frac{d\hat{q}_{iN}(-1)}{dx} \int_{-1}^{1} \left(\frac{\partial \hat{H}(\hat{z}_N)}{\partial \hat{q}_{iN}} + \frac{d\hat{p}_{iN}}{dx} \right) \right|$. Consequently,

$$|E_0| \lesssim N^{\frac{3}{2}-s} \sum_{i=1}^{n} (\|\partial_x^s \hat{p}_i\|_{\omega^{-1+s,-1+s},*} + \|\partial_x^s \hat{q}_i\|_{\omega^{-1+s,-1+s},*}).$$

The proof is complete.   □

REMARK 4.3.   With the global error estimate in (4.1)-(4.2), we can obtain the global error in energy. Actually, there holds for $H \in C^2$ and $\frac{d^s p_i}{dt^s}, \frac{d^s q_i}{dt^s} \in L^\infty(0, t_m)$,

$$|H(z_N(t_m)) - H(z(t_m))| \lesssim |(z - z_N)(t_m)| \lesssim t_m^{\frac{1}{2}} N^{l-s}, \quad t_m = mr.$$

Here in the last step, we have used the conclusions in Remark 4. Note that $H(z(t_m)) = H(z(t_0))$ for any time t_m. Consequently,

$$H(z_N(t_m)) = H(z(t_0)) + O(t_m^{\frac{1}{2}} N^{l-s}) = H(z_N(t_0)) + O(t_m^{\frac{1}{2}} N^{l-s}),$$

which indicates that the global error in energy is growing with respect to $\sqrt{t}$.

5. Numerical experiments

We now perform a sequence of numerical tests to study the conservation properties of the three methods in Section 2. We operate our programs in MATLAB 2016b.

Example 1: Consider a system with a Hamiltonian $H(p, q) = p^2 - q^2 + q^4$ [16]. The corresponding canonical system is

$$p'(t) = -\frac{\partial H}{\partial q} = 2q - 4q^3, \tag{5.1}$$

$$q'(t) = \frac{\partial H}{\partial p} = 2p \tag{5.2}$$

with the initial condition $p(0) = p_0$, $q(0) = q_0$. There are three equilibrium points for this system: $z_1 = (\bar{p}, \bar{q}) = (0, 0)$, $z_2 = (0, \frac{1}{\sqrt{2}})$ and $z_3 = (0, -\frac{1}{\sqrt{2}})$. The zero equilibrium point is a saddle point and the other two are centers. As a result, we have to be careful when we choose initial values for our system in order to avoid the neighborhood of $(0,0)$. The iterative method will not converge otherwise. For comparing with the numerical results in [36], we choose the same initial values as those in [36]. That is, we take $p_0 = 0, q_0 = 0.73$. This gives $H(p_0, q_0) = H_0 = -0.24891759$. Numerical results of the energy $H(p(t), q(t))$ at different times and different N are listed in Table 5.1 for the spectral Galerkin (SG) method, the spectral Petrov-Galerkin (SPG) method, and the spectral Gauss collocation (SGC) method.

Methods	N	$t = 100$	$t = 1000$	$t = 10000$
SG	10	-0.248917591415110	-0.248917604174055	-0.248917735530164
	15	-0.248917590000009	-0.248917590000127	-0.248917590000809
	20	-0.248917589999999	-0.248917590000016	-0.248917590000056
	25	-0.248917589999999	-0.248917590000007	-0.248917589999969
SPG	10	-0.248917590000000	-0.248917589999999	-0.248917589999988
	15	-0.248917590000000	-0.248917589999998	-0.248917589999981
	20	-0.248917590000000	-0.248917589999998	-0.248917589999981
	25	-0.248917590000000	-0.248917589999990	-0.248917589999981
SGC	10	-0.248917590000000	-0.248917590000001	-0.248917590000674
	15	-0.248917590000000	-0.248917589999999	-0.248917589999952
	20	-0.248917590000000	-0.248917589999999	-0.248917589999994
	25	-0.248917590000000	-0.248917589999999	-0.248917589999993

TABLE 5.1. *The energy $H(p(t), q(t))$ at different times and different N for the three spectral methods.*

We see clearly from Table 5.1 that all three methods preserve the energy $H(p(t), q(t))$ very accurately: at least fourteen-digit at time $t = 100$ and thirteen-digit at time $t = 1000, 10000$ with $N \geq 20$.

To further demonstrate the energy conserving property, we plot in Figure 5.1 the error curves for the energy $H(p(t), q(t))$ with respect to N at different times for the three methods. As predicted by Theorem 3.1, we observe that the energy error for the Petrov-Gaerkin method reaches the machine precision for all N, which indicates that the Petrov-Gaerkin method preserves the energy exactly. While for spectral Galerkin and Gauss collocation methods, the error of energy converges exponentially and reaches the machine precision for $N \geq 16$. Consequently, with a reasonable N (e.g., $N \geq 16$), both Gauss collocation and spectral Galerkin methods are energy conserving in practice.
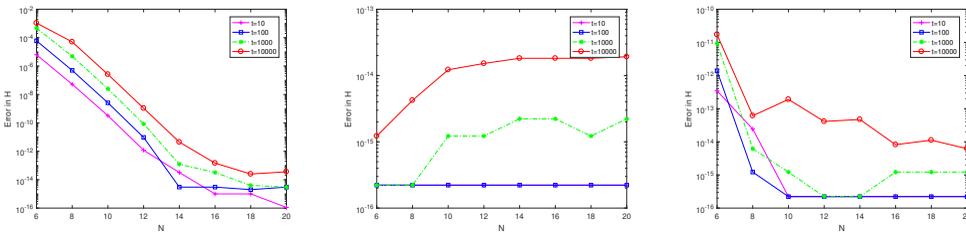
FIG. 5.1. *Error curves of the energy for the spectral Galerkin method (left), the spectral Petrov-Galerkin method (middle), and the Gauss collocation method (right).*

To demonstrate the symplectic structure preserving property, we plot in the left column of Figure 5.2 the phase graph q_N versus p_N for the Gauss collocation method with $N = 30$ on $[0, 10000]$. We observe that the loop is very thin even for large time $t = 10000$, which indicates that the Gauss collocation method can keep trajectory stability for a long time. The phase graphs of the other two methods are almost the same as the Gauss collocation, and hence we only plot their difference. Denote by $e(p_N)$ and $e(q_N)$ the error between the spectral Petrov-Galerkin (or Galerkin) method with the Gauss collocation method for the variable p_N and q_N, respectively, That is,

$$e(p_N) = p_N - p_N^{GC}, \quad e(q_N) = q_N - q_N^{GC},$$

where (p_N^{GC}, q_N^{GC}) are numerical solutions of Gauss collocation methods, and (p_N, q_N) are numerical solutions computed by Petrov-Galerkin or Galerkin methods. We plot in Figure 5.2 (middle and right) the phase plot $e(q_N)$ versus $e(p_N)$ for the spectral Galerkin method (middle) and the Petrov-Galerkin method (right) with the same $N = 30$ and time $t = 10000$. We see that the error between the spectral Galerkin (or Petrov-Galerkin) method and the Gauss collocation method is as small as $10^{-12} - 10^{-11}$ and the error trajectory is stable for a long time. Consequently, like the Gauss collocation method, both Petrov-Galerkin and spectral Galerkin methods preserve symplectic structure in practice.

To show the efficiency of our algorithm, we compare in Table 5.2 the CPU times used for the three spectral methods with the Lobatto spectral collocation method and some symplectic methods proposed in [36]. In our numerical experiments, we choose the time step size $h = 0.001$ for two second-order symplectic schemes 1, 2 and $h = 0.01$ for a sixth-order symplectic method. As we observe, the three spectral methods in this
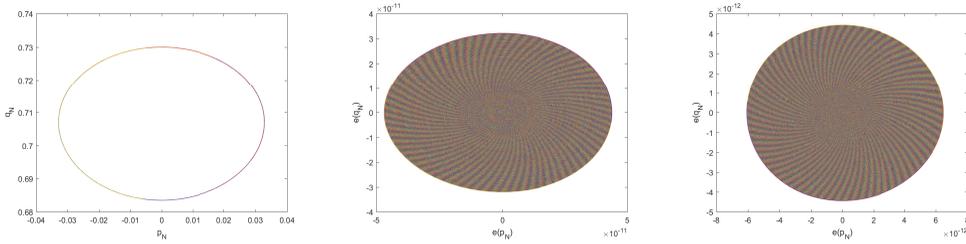
FIG. 5.2. *The phase plot q_N versus p_N for the Gauss collocation method (left), and the phase plot $e(q_N)$ versus $e(p_N)$ between the spectral Galerkin method (middle) or Petrov-Galerkin method (right) with the Gauss collocation method on $[0,10000]$ with $N=30$.*

Methods	time(secs)	Error in Energy
Spectral Galerkin, N=35 on [0,10000]	116.99	1.110×10^{-15}
Spectral Galerkin, N=30 on [0,5000]	42.87	2.054×10^{-15}
Spectral Petrov-Galerkin, N=35 on [0,10000]	118.11	4.552×10^{-15}
Spectral Petrov-Galerkin, N=30 on [0,5000]	42.56	4.940×10^{-15}
Spectral Gauss collocation, N=35 on [0,10000]	119.24	9.215×10^{-15}
Spectral Gauss collocation, N=30 on [0,5000]	43.19	1.388×10^{-15}
Lobatto collocation, N=50 on [0,10000]	2898	5.274×10^{-15}
Lobatto collocation, N=30 on [0,5000]	668	7.661×10^{-15}
Symplectic 1 on [0,450]	2849	9.928×10^{-15}
Symplectic 1 on [0,500]	3585	1.009×10^{-15}
Symplectic 1 on [0,1000]	$>2hrs$	
Symplectic 2 on [0,460]	3010	3.809×10^{-10}
Symplectic 2 on [0,1000]	$>2hrs$	
Symplectic 6th order on [0,4200]	2602	5.496×10^{-15}
Symplectic 6th order on [0,5000]	9833	7.633×10^{-15}
Symplectic 6th order on [0,10000]	$>3hrs$	

TABLE 5.2. *Comparison of CPU times among several numerical methods.*

paper take less CPU time than symplectic methods and the Lobatto spectral collocation method, and thus are more effective.

Listed in Table 5.3 are CPU times on the same interval $[0,1000]$ for the three spectral methods, the Lobatto collocation method in [36], and the sixth-order symplectic method. It seems that the three spectral methods discussed in this paper take much less time than the Lobatto collocation method and the sixth-order symplectic method for achieving the same accuracy in energy.

Example 2: We consider the Henon-Heiles (HH) system [18,19].

The Henon-Heiles (HH) Hamiltonian was introduced in the study of galactic dynamics to describe the motion of stars around the galactic center. The corresponding canonical system is

$$H(p_1,p_2,q_1,q_2) = \frac{1}{2}(p_1^2+p_2^2+q_1^2+q_2^2)+q_1^2q_2-\frac{1}{3}q_2^3,$$

Methods	time(secs)	Error in Energy
Spectral Galerkin, N=18 on [0,1000]	3.53	4.163×10^{-15}
Spectral Petrov-Galerkin, N=18 on [0,1000]	3.41	2.220×10^{-16}
Spectral Gauss collocation, N=18 on [0,1000]	3.37	9.436×10^{-16}
Lobatto collocation, N=18 on [0,1000]	36	5.718×10^{-15}
Symplectic 6th order $h=0.01$ on [0,1000]	80	4.774×10^{-15}

TABLE 5.3.   *Comparison of CPU times among the three spectral methods, the Lobatto collocation method, and the six order symplectic method with the same order of errors.*

and

$$p_1'(t) = -\frac{\partial H}{\partial q_1} = -q_1 - 2q_1 q_2, \quad q_1'(t) = \frac{\partial H}{\partial p_1} = p_1,$$

$$p_2'(t) = -\frac{\partial H}{\partial q_2} = -q_2 - q_1^2 + q_2^2, \quad q_2'(t) = \frac{\partial H}{\partial p_2} = p_2$$

with initial condition $(p_1, p_2, q_1, q_2)(0) = (p_{10}, p_{20}, q_{10}, q_{20})$. Note that the terms q_1^2, q_2^2 in H form a potential well, which is responsible for the oscillations of the particle (the first four terms are related to the kinetic energy). The last two terms $q_1^2 q_2, \frac{1}{3} q_2^3$ are responsible for the existence of the exits from the orbit.

As discussed in [36], there are four equilibrium points for this system which are $z_1 = (\bar{p}_1, \bar{p}_2, \bar{q}_1, \bar{q}_2) = (0,0,0,0)$, a center, $z_2 = (0,0,0,1)$, $z_3 = (0, 0, \frac{\sqrt{3}}{2}, -\frac{1}{2})$ and $z_4 = (0, 0, -\frac{\sqrt{3}}{2}, -\frac{1}{2})$, saddle points. Thus, there are three exits for the energy to escape according to the three saddle points. The total energy $H_z = 0$ for z_1 and $H_z = \frac{1}{6}$ for z_2, z_3, and z_4. If the initial energy is far beyond this H_z, the particles wander inside the region for a certain time in the scattering region until they cross one of the three energy lines and escape to infinity. In other words, when the initial $H < \frac{1}{6}$, the solution is regular; when $H > \frac{1}{6}$, the solution is chaotic. Note that the time they spent in bounded region is named "escape time". The higher the energy, the shorter the escape times are found to be.

To compare with numerical results in [36], we also choose two different sets of initial conditions. The first set represents a regular case with

$$(p_1, p_2, q_1, q_2)(0) = (0.011, 0, 0.013, -0.4), \quad H_0 = 0.101410733333333 < \frac{1}{6}.$$

The second set is a chaotic case with

$$(p_1, p_2, q_1, q_2)(0) = (\sqrt{2 \times 0.15925}, 0.12, 0.12, 0.12), \quad H_0 = 0.1820020 > \frac{1}{6}.$$

We first choose initial conditions from the regular case, and list in Table 5.4 the energy data $H(p_1, p_2; q_1, q_2)$ at different times and different N for the three spectral methods. Similar to Example 1, we see that the energy $H(p_1, p_2; q_1, q_2)$ achieves at least thirteen-digit accuracy at time $t = 100$ and eleven-digit accuracy at time $t = 1000, 10000$ with $N \geq 20$. Consequently, all the three methods preserve the energy for $N \geq 20$.

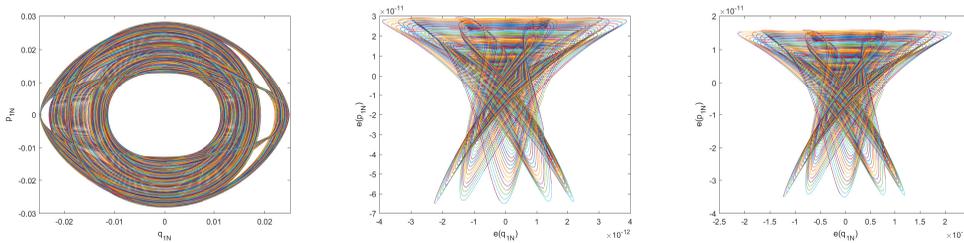Figures 5.3-5.4 represent the phase plot q_{iN} versus p_{iN} for the Gauss collocation method (left), and the error phase plot $e(q_{iN})$ versus $e(p_{iN})$ between the spectral Galerkin method (middle) or the Petrov-Galerkin method (right) with the Guass collocation method on $[0, 2000]$ with $N = 20$ for $i = 1$ and $i = 2$, respectively. Again, just as

Methods	N	$t=100$	$t=1000$	$t=10000$
SG	10	0.101410733274467	0.101410732730912	0.101410727451822
	15	0.101410733333333	0.101410733333333	0.101410733333336
	20	0.101410733333332	0.101410733333312	0.101410733333115
	25	0.101410733333336	0.101410733333356	0.101410733333570
SPG	10	0.101410733333340	0.101410733333398	0.101410733333976
	15	0.101410733333340	0.101410733333397	0.101410733333976
	20	0.101410733333340	0.101410733333396	0.101410733333976
	25	0.101410733333340	0.101410733333398	0.101410733333977
SGC	10	0.101410733333340	0.101410733333398	0.101410733333976
	15	0.101410733333340	0.101410733333397	0.101410733333976
	20	0.101410733333340	0.101410733333396	0.101410733333976
	25	0.101410733333340	0.101410733333398	0.101410733333977

TABLE 5.4. *The energy $H(p(t),q(t))$ at different times and different N for the three spectral methods*

FIG. 5.3. *The phase plot q_{1N} versus p_{1N} for the Gauss collocation method (left), and the error phase plot $e(q_{1N})$ versus $e(p_{1N})$ between the spectral Galerkin method (middle) or the Petrov-Galerkin method (right) with the Guass collocation method on $[0,2000]$ with $N=20$.*

we observe in Example 1, all the three spectral methods can keep trajectory stability for a long time.

In Table 5.5, we compare the CPU times used for the three spectral methods with the Lobatto spectral collocation method and the symplectic methods. We use $h=0.001$ for the symplectic scheme 4 (we refer to [36] for the numerical scheme). Again, we observe that the three spectral methods proposed in this paper take less CPU time and hence are more effective than other numerical methods.

To test the long-time accuracy and discuss how error grows with time, we plot in Figure 5.5 the global L^2 error of the spectral Petrov-Galerkin method on the interval $[0,10000]$ for $N=15$. Since the exact solution is unknown, we take the numerical solution at $N=60$ as our exact solution and test the standard L^2 errors $E_{p_i}, E_{q_i}, i=1,2$ with

$$E_{p_i} = \left(\sum_{l=1}^{m} \int_{(l-1)r}^{lr} (p_i - p_{iN})^2 dt \right)^{\frac{1}{2}}, \quad E_{q_i} = \left(\sum_{l=1}^{m} \int_{(l-1)r}^{lr} (q_i - q_{iN})^2 dt \right)^{\frac{1}{2}}, \quad m=10000/r.$$

To demonstrate how error grows with time, we also plot in Figure 5.5 the curve $y=ct$ with $c=10^{-10}$ in the semi-log scale. From Figure 5.5 we observe that the global error can keep more accurate numerical results for large T. Moreover, it seems that the error grows linearly with respect to the time t.
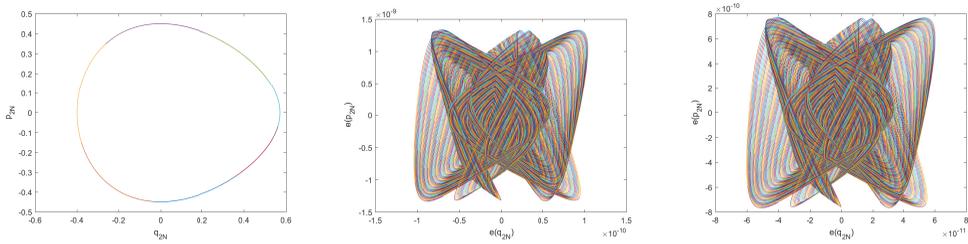
FIG. 5.4. *The phase plot q_{2N} versus p_{2N} for the Gauss collocation method (left), and the phase plot $e(q_{2N})$ versus $e(p_{2N})$ between the spectral Galerkin method (middle) or the Petrov-Galerkin method (right) with the Guass collocation method on $[0,2000]$ with $N = 20$.*

Methods	time(secs)	Error in Energy
Spectral Galerkin, N=20 on [0,10000]	81.71	8.533×10^{-15}
Spectral Galerkin, N=20 on [0,1000]	8.17	8.354×10^{-15}
Spectral Petrov-Galerkin, N=20 on [0,10000]	84.23	1.283×10^{-13}
Spectral Petrov-Galerkin, N=20 on [0,1000]	9.34	1.250×10^{-14}
Spectral Gauss collocation, N=20 on [0,10000]	71.59	4.677×10^{-14}
Spectral Gauss collocation, N=20 on [0,1000]	7.11	4.710×10^{-14}
Collocation, N=20 on [0,10000]	2691	1.900×10^{-12}
Collocation, N=20 on [0,1000]	25	1.907×10^{-13}
Symplectic 4 on [0,65]	21	6.003×10^{-5}
Symplectic 4 on [0,200]	970	6.003×10^{-5}
Symplectic 4 on [0,1000]	$> 2hrs$	

TABLE 5.5. *Comparison of CPU times among several numerical methods.*

Now we choose initial conditions from the chaotic case. Figures 5.6-5.7 represent the chaotic solution $(p_{1N}, p_{2N}, q_{1N}, q_{2N})$ when the particle wanders in the bounded region until it crosses the energy threshold line and escapes, where Figure 5.6 and the left one of Figure 5.7 show the chaotic solutions on $[0,24]$ by spectral Galerkin, Gauss collocation, and Petrov-Galerkin methods with $N = 20$, respectively. For the purpose of observing the energy threshold line clearly, we also plot in Figure 5.7 (the right one) the chaotic solution on $[0,26]$ for $N = 30$ by the Petrov-Galerkin method, from which we observe that the energy threshold line is close to $t = 25$.

In Figures 5.8-5.9, we plot the error curves of the energy for the spectral Gauss collocation method and Galerkin method at different times t with a given $N = 20$. To demonstrate how error grows with time, we also plot in Figures 5.8-5.9 the curve $y = ct^{\frac{1}{2}}$ with $c = 10^{-13}$ in the semi-log scale. Just as demonstrated in Remark 6, the energy error is growing with respect to $t^{\frac{1}{2}}$.

Figure 5.10 shows the intersections of the orbits of Henon-Heiles system with Poincaré section, computed by the spectral Petrov-Galerkin method on $[0,50000]$ with $N = 40$ for different energies (similar numerical results can be obtained by using the other two spectral methods). The points plotted in Figure 5.10 are in the (q_{2N}, p_{2N}) plane, which are obtained by the following way: we first fix the energy H_0 and choose any point $P_0 = (q_{20}, p_{20}), q_{10} = 0$ and obtain a positive root p_{10} from H_0 and P_0; then we follow the solution until it hits again the surface $q_{1N} = 0$ in the positive direction
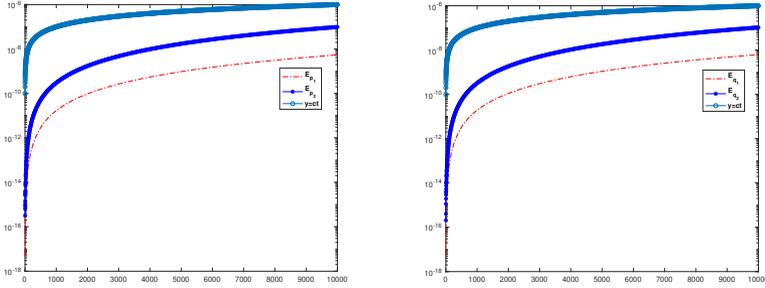
FIG. 5.5.   *Global L^2 errors of the spectral Petrov-Galerkin method for $N = 15$ on the interval* $[0, 10000]$.
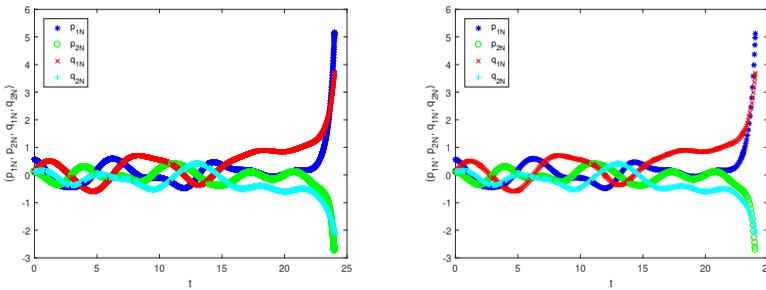

FIG. 5.6.   *Chaotic solutions on $[0, 24]$ by the spectral-Galerkin method (left) and the Gauss collocation method (right) with $N = 20$.*

$p_{1N} > 0$ and obtain a point $P_1 = (q_{2N}, p_{2N})$; in the same way we compute P_2, etc. As we may observe from Figure 5.10, all successive points P_i rotate regularly around the curve and the motion is regular for $H_0 = \frac{1}{12}$; while for $H_0 = \frac{1}{8}$, most of the motions are regular but chaotic motions appear; for higher energy $H_0 = \frac{1}{6}$, it is apparent that this 'ergodic' trajectory covers almost the whole area and almost all motions are chaotic.

Example 3: We consider the problem of Fermi, Pasta & Ulam [13], which is a simple model for simulations in statistical mechanics and reveals highly unexpected dynamical behavior. The problem is described by a Hamiltonian system with total energy

$$H(p,q) = \frac{1}{2} \sum_{i=1}^{m} (p_{2i-1}^2 + p_{2i}^2) + \frac{\omega^2}{4} \sum_{i=1}^{m} (q_{2i} - q_{2i-1})^2 + \sum_{i=0}^{m} (q_{2i+1} - q_{2i})^4,$$

where the variables $q_1, ..., q_{2m} (q_0 = q_{2m+1} = 0)$ stand for the displacements of the mass points, and $p_i = \dot{q}_i$ for their velocities, and ω is assumed to be large. By introducing new variables [30]

$$x_{0,i} = (q_{2i} + q_{2i-1})/\sqrt{2}, \quad x_{1,i} = (q_{2i} - q_{2i-1})/\sqrt{2},$$
$$y_{0,i} = (p_{2i} + p_{2i-1})/\sqrt{2}, \quad y_{1,i} = (p_{2i} - p_{2i-1})/\sqrt{2},$$
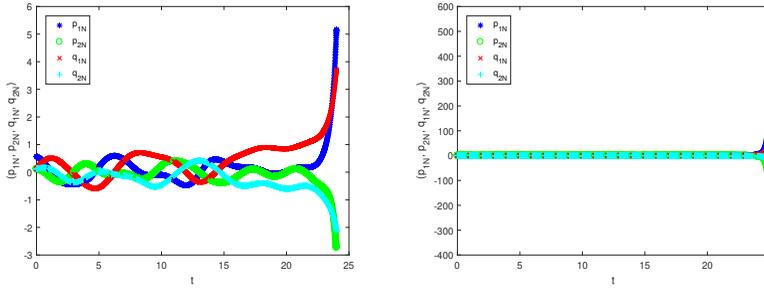
FIG. 5.7. *Chaotic solutions on* $[0,24]$ *for* $N=20$ *(left) and* $[0,26]$ *for* $N=30$ *(right) by the Petrov-Galerkin method.*
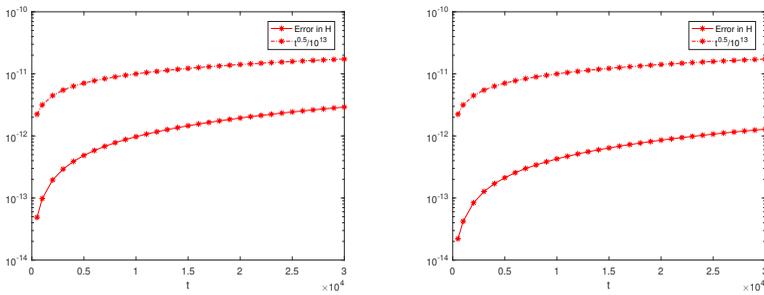

FIG. 5.8. *The curves of energy error versus time by using the spectral Gauss collocation method for Example* 1 *(left) and Example* 2 *(right) with* $N=20$ *on the time interval* $[0,30000]$.

the motion in the new variables is again described by a Hamiltonian system

$$H(y,x) = \frac{1}{2}\sum_{i=1}^{m}(y_{0,i}^2 + y_{1,i}^2) + \frac{\omega^2}{2}\sum_{i=1}^{m}x_{1,i}^2 + \frac{1}{4}\big((x_{0,1}-x_{1,1})^4$$
$$+ \sum_{i=1}^{m-1}(x_{0,i+1}-x_{1,i+1}-x_{0,i}-x_{1,i})^4 + (x_{0,m}+x_{1,m})^4\big).$$

Here $x_{0,i}, x_{1,i}$ represent a scaled displacement and expansion of the i-th stiff spring, respectively, and $y_{0,i}, y_{1,i}$ their velocities.

In our numerical experiment, we take $m=2, \omega=50$ and the following initial values

$$(x_{0,1}, y_{0,1}, x_{1,1}, y_{1,1})(0) = (1,1,\omega^{-1},1), \quad (x_{0,2}, y_{0,2}, x_{1,2}, y_{1,2})(0) = (0,0,0,0).$$

In order to demonstrate long-time accuracy and efficiency of our algorithm, we choose numerical solutions with $N=50$ as reference solutions, and test the approximation error over the time interval $[0,50000]$. We plot in Figure 5.11 the error curves of the approximation solutions p_{1N} and q_{1N} (similarly for other variables $(p_{iN}, q_{iN}), i \leq 4$ or the new variables (y_N, x_N)) with $N = 20, 30, 40$, by using the spectral Petrov-Galerkin method. We observe that with a reasonable N (e.g., $N \geq 30$), the accuracy of approximation error can still reach as high as 10^{-8} even for large $t = 50000$.

Now we consider the conservation properties of our method (i.e., the SPG method). For comparison, we also test spectral variational integrators (SVI) (see, e.g., [25]) and the
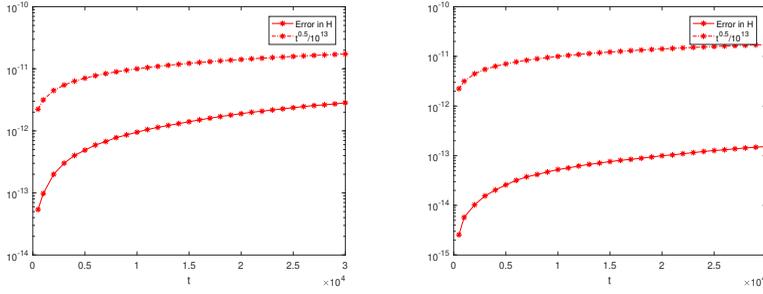
Fig. 5.9.   *The curves of energy error versus time by using the spectral Gakerkin method for Example* 1 *(left) and Example* 2 *(right) with* $N = 20$ *on the time interval* $[0, 30000]$.
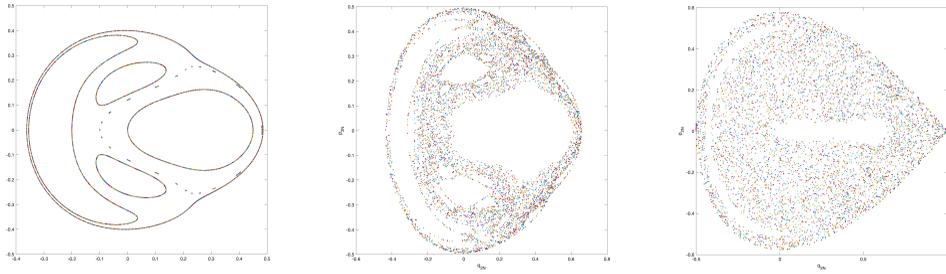

Fig. 5.10.   *Poincaré section for* $q_{1N} = 0, p_{1N} > 0$ *of the spectral Petrov-Galerkin method on* $[0, 50000]$ *with* $N = 40$ *for* $H_0 = \frac{1}{12}$ *(6 orbits, left), and* $H_0 = \frac{1}{8}$ *(1 orbit, middle) and* $H_0 = \frac{1}{6}$ *(1 orbit, right).*
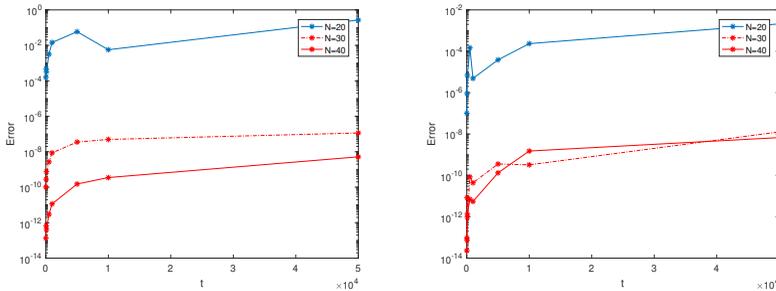

Fig. 5.11. *Global* L^2 *errors of the spectral Petrov-Galerkin method for* p_{1N} *(left) and* q_{1N} *(right) on the time interval* $(0, 50000]$

low-order finite elemenet method (FEM) (see, e.g., [40]) in our numerical experiments. Note that the SVI is symplectic and the FEM preserves energy.

To test the efficiency, we use the same degree of freedom for the above three methods. That is, we take $N = 19, r = 0.5$ for the SVI and SPG methods, and $k = 3, h = 0.1$ for the FEM, where k, h denote the polynomial degree and time step size, respectively. Note
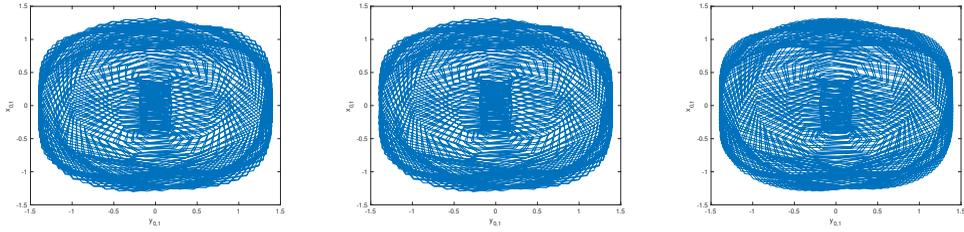
FIG. 5.12. *The phase plot $x_{0,1}$ versus $y_{0,1}$ on the time interval* $[19500, 20000]$ *by using the method of SPG(left), SVI(middle) and FEM (right).*
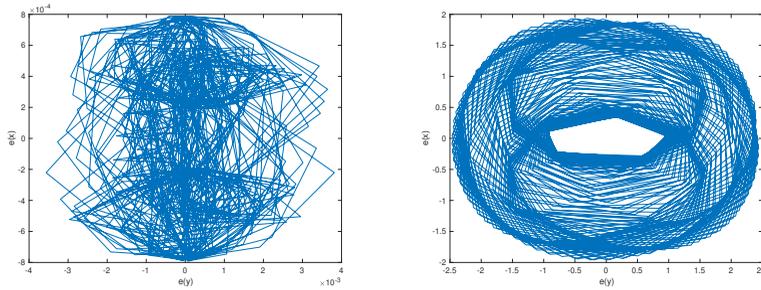

FIG. 5.13. *The error phase plot $e(y)$ versus $e(x)$ between the SPG method (left) or the FEM (right) with the SVI method on* $[19500, 20000]$ *with* $\omega = 50$.
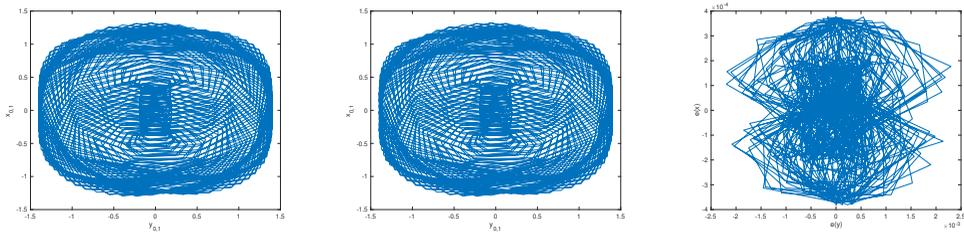

FIG. 5.14. *The phase plot $x_{0,1}$ versus $y_{0,1}$ by using the method of SPG(left), SVI(middle), and the error phase plot $e(y)$ versus $e(x)$ between the SPG method and the SVI method (right) on the time interval* $[99500, 100000]$.

that the total degree of freedom at the time interval $(0, 1]$ is

$$(N+1)/r - 1 = (k+1)/h - 1 = 39.$$

We plot in Figure 5.12 the phase graphs $x_{0,1} = (q_{2N} + q_{1N})/\sqrt{2}$ versus $y_{0,1} = (p_{2N} + p_{1N})/\sqrt{2}$ on the time interval $[19500, 20000]$ by using the above three methods. As we may observe, the trajectory is stable and the phase graphs of the three methods are similar. To show the accuracy of the trajectory, we also compare the difference of the phase graphs for the three methods, e.g., we plot the error phase graphs. Noticing that the SVI method is symplectic, we choose the phase graphs computed by SVI as the reference phase, and then denote by $e(y)$ and $e(x)$ the error phase between the SPG
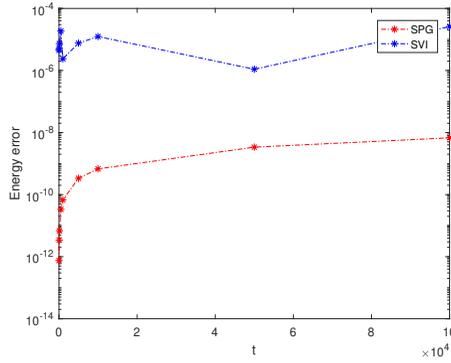
FIG. 5.15. *The curves of the total energy H at the time interval $[0,100000]$ by using the spectral Petrov-Galerkin method and spectral variational integrators with $\omega = 50$.*
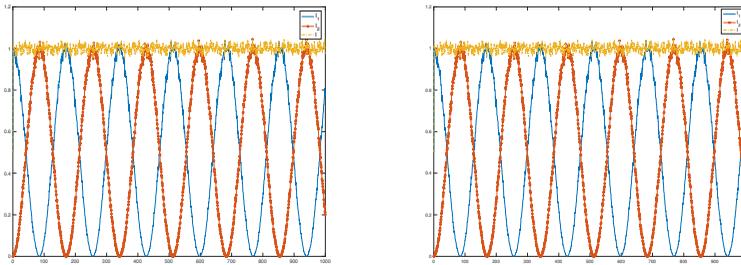

FIG. 5.16. *The total oscillatory energy I and energy exchange of stiff springs calculated by SVI (left) and SPG (right) method with $\omega = 50$.*
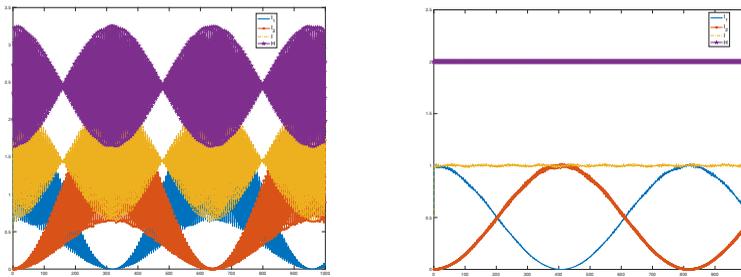

FIG. 5.17. *The total oscillatory energy I and energy of stiff springs calculated by SVI (left) and SPG (right) method with $\omega = 100$.*

method (or FEM) with the SVI method for the variable $y_{0,1}$ and $x_{0,1}$, respectively, That is,

$$e(y) = y_{0,1} - y_{0,1}^{svi}, \quad e(x) = x_{0,1} - x_{0,1}^{svi},$$

where $(y_{0,1}^{svi}, x_{0,1}^{svi})$ are numerical solutions of SVI methods, and $(y_{0,1}, x_{0,1})$ are numerical solutions computed by FEM or SPG methods.

We plot in Figure 5.13 the error phase plot $e(y)$ versus $e(x)$ over the time interval $t = [19500, 20000]$. We see that the error phase between the SPG method and the SVI method is as small as $10^{-4} - 10^{-3}$, while the error phase between the FEM and the SVI method has only $O(1)$ accuracy. In this sense, the SPG method keeps a higher accuracy of the trajectory than the counterpart FEM. To test the accurate trajectories over long time interval for the SPG method, we also plot in Figure 5.14 the phase plot $x_{0,1}$ versus $y_{0,1}$ by using the method of SPG (left), SVI (middle), and the error phase plot $e(y)$ versus $e(x)$ between the SPG method and the SVI method (right) on the time interval $[99500, 100000]$. Again, we observe that the phase graphs for both methods are almost the same and the error between them can still reach as high as $10^{-4} - 10^{-3}$ for large t. Consequently, similar to the SVI method, the SPG method preserves symplectic structure efficiently. Especially, we can take a reasonable N to obtain the desired accuracy in practice.

Now we consider the energy preserving properties of the SPG and SVI method. Let

$$I_j(x_{1,j}, y_{1,j}) = \frac{1}{2}(y_{1,j}^2 + w^2 x_{1,j}^2), \ \ j \le 2.$$

denote the energy of the j-th stiff spring. Note that there is an exchange of the energy between the stiff spring in the exact solution of the Fermi-Pasta-Ulam model, but the total oscillatory energy $I = I_1 + I_2$ remains close to a constant value.

We first plot in Figure 5.15 the curves of the total energy H over the time interval $[0, 100000]$ by using the SPG and SVI methods with $\omega = 50, N = 19$. We observe that the SPG method preserves the total energy more accurately than the counterpart SVI.

In Figures 5.16-5.17, we plot the curves of the total oscillatory energy I and the energy of the j-th stiff spring I_j by using the SPG and SVI methods with $\omega = 50$ and $\omega = 100$, respectively. We see that, for $\omega = 50$, an exchange of energy takes place, going from the first stiff spring with energy I_1 to the second stiff spring and then going back to the first one, and so on; and both SPG and SVI methods preserve the total oscillatory energy I with high accuracy. However, if we choose $\omega = 100$, the total oscillatory energy I and the total energy H are not conserved any more by using the SVI method, and the oscillations appear. The SPG method still preserves the total oscillatory energy I and the total energy H efficiently, and an exchange of energy between the stiff springs is observed again. From the point of view of energy conservation, it seems that the SPG method is superior than the counterpart SVI in this Fermi-Pasta-Ulam model.

6. Conclusions

In this work, we have discussed three efficient numerical methods: spectral Petrov-Galerkin, spectral Gauss collocation, and spectral Galerkin methods for nonlinear Hamiltonian systems. A theoretical investigation has been made for the convergence, symplectic structure preservation and energy conservation. We have proved that the spectral Petrov-Galerkin method is energy conserving and the error in symplecticity decays with spectral accuracy. Furthermore, we show that the global error decays exponentially with respect to the polynomial degree under some regularity assumption. Comparing with existing symplectic methods and low-order Galerkin methods, our theoretical and numerical results have demonstrated that the spectral methods discussed in this paper have some desirable properties and advantages.

(1) They are high-order methods with spectral accuracy and hence require less CPU time than traditional methods to achieve the same accuracy.

(2) They preserve energy and symplectic structure in practice for some reasonable N and computational cost. Especially, the Petrov-Galerkin method preserves the energy exactly.

(3) They predict more accurate trajectories over long time intervals.

The high efficiency of the three spectral methods makes them more appealing for solving partial different equations with high-order derivatives, especially when time-discretization is concerned. Our current and future works include the analysis of the iterative scheme and the application of the algorithms to other time-dependent partial differential equations, e.g., hyperbolic and parabolic problems, which could be more challenging and interesting.

REFERENCES

[1] J. An, W. Cao, and Z. Zhang, *An efficient spectral Petrov-Galerkin method for nonlinear Hamiltonian systems*, Comm. Comput. Phys., 56:1249–1273, 2019. 4.2

[2] A. Aubry and P. Chartier, *Pseudo-symplectic Runge-Kutta methods*, BIT Numer. Math., 38:439–461, 1998. 1

[3] J. Candy and W. Rozmus, *A symplectic integration algorithm for separable Hamiltonian functions*, J. Comput. Phys., 92:230–256, 1991. 1

[4] E. Celledoni, R.I. McLachlan, D.I. McLaren, B. Owren, G.R.W. Quispel, and W.M. Wright, *Energy-preserving Runge-Kutta methods*, M2AN Math. Model. Numer. Anal., 43:645–649, 2009. 1

[5] E. Celledoni, B. Owren, and Y. Sun, *Minimal stage, energy preserving Runge-Kutta method for polynomial Hamiltonian system is the averaged vector field method*, Math. Comput., 83:1689–1700, 2014. 1

[6] C. Chen, Q. Tang, and S. Hu, *Finite element method with superconvergence for nonlinear Hamiltonian systems*, J. Comput. Math., 29:167–184, 2011. 1

[7] D. Cohen, *Conservation properties of numerical integrators for highly oscillatory Hamiltonian systems*, IMA J. Numer. Anal., 26:34–59, 2006. 1

[8] D. Cohen, L. Gauckler, E. Hairer, and C. Lubich, *Long-term analysis of numerical integrators for oscillatory Hamiltonian systems under minimal non-resonance conditions*, BIT Numer. Math., 55:705–732, 2015. 1

[9] D. Cohen, E. Hairer, and C. Lubich, *Numerical energy conservation for multi-frequency oscillatory differential equations*, BIT Numer. Math., 45:287–305, 2005. 1

[10] P. Console, E. Hairer, and C. Lubich, *Symmetric multistep methods for constrained Hamiltonian systems*, Numer. Math., 124:517–539, 2013. 1

[11] P.J. Davis and P. Rabinowitz, *Methods of Numerical Integration*, Courier Corporation, 2007. 3.1

[12] E. Faou, E. Hairer, and T.L. Pham, *Energy conservation with non-symplectic methods: examples and counter-examples*, BIT Numer. Math., 44:699–709, 2004. 1

[13] E. Fermi, J. Pasta, and S. Ulam, *Studies of nonlinear problems*, Document LA-1940, Los Alamos National Laboratory, 1955. 5

[14] K. Feng, *On difference schemes and symplectic geometry*, in K. Feng (ed.), Proceedings of the 5th International Symposium on Differential Geometry and Differential Equations, Science Press, Beijing, 42–58, 1984. 1

[15] K. Feng, *Difference schemes for Hamiltonian formalism and symplectic geometry*, J. Comput. Math., 4:279–289, 1986. 1

[16] K. Feng, *How to compute properly Newton's equation of motion*, in L.A. Ying and B.Y. Guo, Proceedings of 2nd Conference on Numerical Methods for Partial Differential Equations, World Scientific, Singapore, 15–22, 1992. 1, 5

[17] K. Feng, *The Hamiltonian way for computing Hamiltonian dynamics*, in R. Spigler (ed.), Applied and Industrial Mathematics, Springer, Dordrecht, 17–35, 1991. 1

[18] K. Feng and M. Qin, *Symplectic Geometric Algorithms for Hamiltonian Systems*, Springer, Berlin, 2010. 1, 5

[19] K. Feng and M. Qin, *The symplectic methods for the computation of Hamiltonian equations*, In Y.I. Zhu and B. Guo (eds.), Numerical Methods for Partial Differential Equations, Springer, Berlin, Heidelberg, 1–37, 1987. 1, 5

[20] K. Feng and D. Wang, *Variations on a theme by Euler*, J. Comput. Math., 12:97–106, 1998. 1

[21] K. Feng, H. Wu, M. Qin, and D. Wang, *Construction of canonical difference schemes for Hamiltonian formalism via generating functions*, J. Comput. Math., 7:71–96, 1989. 1

[22] B. Garia-Archilla, J.M. Sanz-Serna, and R.D. Skeel, *Long-time-step methods for oscillatory differential equations*, SIAM J. Sci. Comput., 20:930–963, 1998. 1

[23] L. Gauckler, E. Hairer, and C. Lubich, *Energy separation in oscillatory Hamiltonian systems without any non-resonance condition*, Comm. Math. Phys., 321:803–815, 2013. 1

[24] B.-Y. Guo and Q.-Z. Wang, *Legendre-Gauss collocation methods for ordinary differential equations*, Adv. Comput. Math., 30:249–280, 2009. 1, 2, 4, 4.1

[25] J. Hall and M. Leok, *Spectral variational integrators*, Numer. Math., 130:681–740, 2015. 1, 5

[26] J. Hall and M. Leok, *Lie group spectral variational integrators*, Found. Comput. Math., 17:199–257, 2017. 1

[27] E. Hairer and C. Lubich, *Long-time energy conservation of numerical methods for oscillatory differential equations*, SIAM J. Numer. Anal., 38:414–441, 2000. 1

[28] E. Hairer and C. Lubich, *Oscillations over long times in numerical Hamiltonian systems*, in B. Engquist, A. Fokas, E. Hairer and A. Iserles (eds.), Highly Oscillatory Problems, London Mathematical Society Lecture Note Series 366, Cambridge Univ. Press, 2009. 1

[29] E. Hairer and C. Lubich, *Long-term analysis of the Störmer-Verlet method for Hamiltonian systems with a solution-dependent high frequency*, Numer. Math., 134:119–138, 2016. 1

[30] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations*, Second Edition, Springer Science in Computational Mathematics 31, Springer-Verlag, Berlin, 2006. 1, 3.3, 5

[31] C. Huang and Z. Zhang, *Spectral collocation method for differential algebraic equations with arbitrary index: with application to Hamiltonian systems*, J. Sci. Comput., 58:499–516, 2014. 1

[32] J.E. Marsden, G. Misiolek, J.P. Ortega, M. Perlmutter, and T.S. Ratiu, *Hamiltonian Reduction by Stages*, Springer, 2007. 1

[33] J.E. Marsden and M. West, *Discrete mechanics and variational integrators*, Acta Numer., 10:357–514, 2001. 1

[34] F.M. Lasagni, *Canonical Runge-Kutta methods*, Z. Angew. Math. Phys., 39:952–953, 1988. 1

[35] R.I. McLachlan and P. Atela, *The accuracy of symplectic integrators*, Nonlinearity, 5:541, 1992. 1

[36] K. Nanyamee and Z. Zhang, *Comparison of a spectral collocation method and symplectic methods for Hamilton systems*, Int. J. Numer. Anal. Mod., 8:86–104, 2011. 1, 2, 5, 5, 5, 5

[37] J.M. Sanz-Serna, *Runge-Kutta schemes for Hamiltonian systems*, BIT Numer. Math., 28:877–883, 1988. 1

[38] J.M. Sanz-Serna and M.P. Calvo, *Numerical Hamiltonian Problems*, Courier Dover Publications, 2018. 1

[39] J. Shen, T. Tang, and L.L. Wang, *Spectral methods: Algorithms, Analysis and Applications*, Springer Science and Business Media, 2011. 3.1, 3.1, 4

[40] Q. Tang, C. Chen, and L. Liu, *Energy conservation and symplectic properties of continuous finite element methods for Hamiltonian systems*, Appl. Math. Comput., 181:1357–1368, 2006. 1, 5

[41] W. Tang and Y. Sun, *Construction of Runge-Kutta type methods for solving ordinary differential equations*, Appl. Math. Comput., 234:179–191, 2014. 2.1

[42] W. Tang and Y. Sun, *Time finite element methods: A unified framework for numerical discretizations of ODEs*, Appl. Math. Comput., 219:2158–2179, 2012. 2.1

[43] G. Zhong and J.E. Marsden, *Lie-Poisson Hamilton-Jacobi theory and Lie-Poisson integrators*, Phys. Lett. A., 133:134–139, 1988. 1