Explicit symplectic approximation of nonseparable Hamiltonians: Algorithm and long time performance

Molei Tao[*]

Georgia Institute of Technology, Atlanta, Georgia 30332, USA

Explicit symplectic integrators have been important tools for accurate and efficient approximations of mechanical systems with separable Hamiltonians. This article proposes for arbitrary Hamiltonians similar integrators, which are explicit, of any even order, symplectic in an extended phase space, and with pleasant long time properties. They are based on a mechanical restraint that binds two copies of phase space together. Using backward error analysis, Kolmogorov-Arnold-Moser theory, and additional multiscale analysis, an error bound of $O(T\delta^l\omega)$ is established for integrable systems, where T, δ, l, and ω are, respectively, the (long) simulation time, step size, integrator order, and some binding constant. For nonintegrable systems with positive Lyapunov exponents, such an error bound is generally impossible, but satisfactory statistical behaviors were observed in a numerical experiment with a nonlinear Schrödinger equation.

I. INTRODUCTION AND THE ALGORITHM

Symplectic integrators preserve the phase-space volume $dq \wedge dp$ like the exact Hamiltonian flow and thus are the preferred approach for long time simulations of mechanical systems [1–5]. Explicit symplectic integration has been extensively studied for separable Hamiltonians (i.e., $H(q,p) = K(p) + V(q)$; see [1–3,5–11]), but much less so for nonseparable systems. However, nonseparable Hamiltonians model important problems, such as a finite-dimensional representation of the nonlinear Schrödinger equation [12], nearly integrable systems in action-angle coordinates (see [13–16] for astrophysical examples), charged particle dynamics [17–19], mechanical systems in a rotating frame [20], molecular dynamics with thermostats [21,22], time-regularized mechanical systems [23], classical systems with post-Newtonian correction that approximates general relativity effects [24,25], rigid body dynamics [26], pendula dynamics [27], and the scattering of electromagnetic waves by attenuating materials [28].

A common misunderstanding is that a symplectic integration has to be implicit (and hence computationally expensive) when the Hamiltonian is nonseparable. In fact, explicit symplectic integrations have been made possible for several subclasses of nonseparable Hamiltonians [17,19,22,29–33]. Notably, a first step toward explicit approximation of arbitrary H was recently made in [34] by considering a larger system in an extended phase space. Although the method integrator proposed there is only accurate for a short time, the idea of extended phase space is a building block of this research.

We approximate the flow of an arbitrary nonseparable $H(Q,P)$. A generic, high-order, explicit, and symplectic integrator is proposed with a provable pleasant long time performance. This is achieved by considering an augmented Hamiltonian,

$$\bar{H}(q,p,x,y) := H_A + H_B + \omega H_C,$$

in an extended phase space with symplectic 2 form $dq \wedge dp + dx \wedge dy$, where $H_A := H(q,y)$ and $H_B := H(x,p)$ cor-

respond to two copies of the original system with mixed-up positions and momenta, $H_C := \|q - x\|_2^2/2 + \|p - y\|_2^2/2$ is an artificial restraint, and ω is a constant that controls the binding of the two copies.

First of all, note the initial value problems

$$
\begin{aligned}
\dot{Q} &= \partial_P H(Q,P), & Q(0) &= Q_0 \\
\dot{P} &= -\partial_Q H(Q,P), & P(0) &= P_0
\end{aligned}
\quad \text{and}
$$

$$
\begin{aligned}
\dot{q} &= \partial_p \bar{H}(q,p,x,y), & q(0) &= Q_0 \\
\dot{p} &= -\partial_q \bar{H}(q,p,x,y), & p(0) &= P_0 \\
\dot{x} &= \partial_y \bar{H}(q,p,x,y), & x(0) &= Q_0 \\
\dot{y} &= -\partial_x \bar{H}(q,p,x,y), & y(0) &= P_0
\end{aligned}
$$

have the same exact solution in the sense that $q(t) = x(t) = Q(t)$, $p(t) = y(t) = P(t)$. This is because the form of $\bar{H}$ turns the second system into

$$
\begin{aligned}
\dot{q} &= \partial_p H(x,p) + \omega(p - y) \\
\dot{p} &= -\partial_q H(q,y) - \omega(q - x) \\
\dot{x} &= \partial_y H(q,y) + \omega(y - p) \\
\dot{y} &= -\partial_x H(x,p) - \omega(x - q),
\end{aligned}
$$

which upon the substitution $q(t) = x(t) = Q(t)$, $p(t) = y(t) = P(t)$ becomes the first system. The uniqueness of the ordinary differential equation (ODE) solution rules out the possibility of disagreed other solutions.

Second, it is possible to construct high-order symplectic integrators for $\bar{H}$ with explicit updates. Respectively denote by $\phi^\delta_{H_A}, \phi^\delta_{H_B}, \phi^\delta_{\omega H_C}$ the time-δ flow of H_A, H_B, ωH_C. Exact expressions of these flows can be explicitly obtained as

$$
\phi^\delta_{H_A} : \begin{bmatrix} q \\ p \\ x \\ y \end{bmatrix} \mapsto \begin{bmatrix} q \\ p - \delta\partial_q H(q,y) \\ x + \delta\partial_y H(q,y) \\ y \end{bmatrix},
$$

$$
\phi^\delta_{H_B} : \begin{bmatrix} q \\ p \\ x \\ y \end{bmatrix} \mapsto \begin{bmatrix} q + \delta\partial_p H(x,p) \\ p \\ x \\ y - \delta\partial_x H(x,p) \end{bmatrix},
$$

_____

[*]mtao@gatech.edu

$$\phi_{\omega H_C}^{\delta} : \begin{bmatrix} q \\ p \\ x \\ y \end{bmatrix} \mapsto \frac{1}{2} \begin{bmatrix} \begin{pmatrix} q+x \\ p+y \end{pmatrix} + R(\delta)\begin{pmatrix} q-x \\ p-y \end{pmatrix} \\ \begin{pmatrix} q+x \\ p+y \end{pmatrix} - R(\delta)\begin{pmatrix} q-x \\ p-y \end{pmatrix} \end{bmatrix},$$

$$\text{where } R(\delta) := \begin{bmatrix} \cos(2\omega\delta)I & \sin(2\omega\delta)I \\ -\sin(2\omega\delta)I & \cos(2\omega\delta)I \end{bmatrix}. \quad (1)$$

Then we construct a numerical integrator that approximates $\bar{H}$ by composing these maps: It is well known that

$$\phi_2^{\delta} := \phi_{H_A}^{\delta/2} \circ \phi_{H_B}^{\delta/2} \circ \phi_{\omega H_C}^{\delta} \circ \phi_{H_B}^{\delta/2} \circ \phi_{H_A}^{\delta/2}, \quad (2)$$

commonly named as Strang splitting, has a third-order local error (thus a second-order method), and is a symmetric method. An arbitrary high (even)-order integrator can also be obtained. The lth-order version will have an update map ϕ_l^{δ} given by, for instance, "the triple jump" [8,10,35,36]:

$$\phi_l^{\delta} := \phi_{l-2}^{\gamma_l \delta} \circ \phi_{l-2}^{(1-2\gamma_l)\delta} \circ \phi_{l-2}^{\gamma_l \delta}, \quad \text{where } \gamma_l = \frac{1}{2-2^{1/(l+1)}}. \quad (3)$$

Each update ϕ_l^{δ} constitutes one of our proposed integrators, which are symplectic because each flow is symplectic. Each one produces a discrete trajectory

$$\begin{bmatrix} q_N \\ p_N \\ x_N \\ y_N \end{bmatrix} := (\phi_l^{\delta})^N \begin{bmatrix} Q(0) \\ P(0) \\ Q(0) \\ P(0) \end{bmatrix},$$

where q_N, p_N (and x_N, y_N too) approximate the exact solution $Q(N\delta), P(N\delta)$.

This approximation needs justification, because although the exact solutions of H and $\bar{H}$ agree, truncation errors in the numerical solution of $\bar{H}$ may lead to large global error after $O(1)$ time. We will see that this will not be the case if the system is integrable. To discuss the idea, note [34] considered the Hamiltonian $H_A + H_B$ without the binding, i.e., $\bar{H}$ with $\omega = 0$. The resulting integrator did produce (q,p) and (x,y) that well approximate Q,P till $O(1)$ time, but then they quickly diverge. A fix was suggested in [34] based on an extra phase-space mixing substep for inducing a coupling between (q,p) and (x,y); unfortunately, symplecticity is lost due to this substep. We replace this substep using ωH_C. This is because, under reasonable assumptions, the near conservation of $\bar{H}$ by its symplectic integration (established by backward error analysis; see [3]) will imply the boundedness of ωH_C, and thus that $q-x, p-y$ are at most $O(1/\sqrt{\omega})$, which prevents the undesired divergence.

One may worry whether large ω requires small δ, which would undermine the computational efficiency gained by an explicit integrator. Also, will a finite ω introduce another source of error besides truncation error? In addition, one is interested in whether $q-Q, p-P$ are small, but not $q-x$ and $p-y$. Section II will, for integrable $\bar{H}$, bound $\|q-Q\|$ and $\|p-P\|$ until a large time. Section III will show if H is not integrable; although long time accuracy of trajectory should not be expected, it is still possible to numerically capture statistical behaviors of the system, at least in a nonlinear Schrödinger equation example. The difference

between integrabilities of H and $\bar{H}$ is discussed in Sec. IV, which further explains why ω should be larger than a threshold.

II. INTEGRABLE PROBLEMS: LINEAR GROWTH OF LONG TIME APPROXIMATION ERROR

Provided that $\bar{H}$ corresponds to an integrable system (which will be the case if, roughly speaking, H is integrable and $\omega \geqslant \omega_0$ for some constant ω_0; see Sec. IV), we will demonstrate that the proposed lth-order integrator [Eq. (3)] has a numerical error of

$$O(T\delta^l \omega)$$

until at least $T = O(\min(\delta^{-l}\omega^{-1}, \omega^{1/2}))$. Numerical results consistent with this bound for even larger T values will then be shown.

For long time simulation, this linear growth with T is advantageous to nonsymplectic integrators, whose errors can grow exponentially, e.g., $O(e^{CT}\delta^l\omega^{\bar{l}})$ [37]. It also improves the pioneering symplectic integrator in [34], which becomes inaccurate after $T = O(1)$. One can see $\omega \gtrsim \omega_0$ is sufficient for accuracy, and the extra error introduced by a finite ω vanishes as $\delta \to 0$. Interestingly, an ω too large is actually discouraged by this error bound. In addition, $\delta \ll \omega^{-1/l}$ is sufficient, and although there is still a tradeoff between accuracy and efficiency, a larger l allows δ to be much larger than $o(1/\omega)$, i.e., no need to resolve the oscillation induced by ωH_C. Accuracy and efficiency are thus simultaneously improved.

The main idea for establishing this bound is to view the numerical solution as discrete samples of the exact solution of some nearby Hamiltonian $\tilde{H}$ (i.e., backward error analysis) and characterize the distance between $\tilde{H}$ and $\bar{H}$ as a small parameter ϵ. An application of Kolmogorov-Arnold-Moser (KAM) theory [38–43] then bounds the differences between action and angle variables in $\tilde{H}$ and $\bar{H}$. Such bounds will specify how these two systems deviate in q, p coordinates and hence quantify the numerical error, because the flows of $\bar{H}$ and $\tilde{H}$, respectively, correspond to the exact solution of H and the numerical solution. Similar techniques have been established (see Chap X of [3] for a review), and the main novelty of our derivation is a refined estimation of ϵ combined with these techniques. Specifically:

1. Denote by $t := \omega T$ and $h := \omega \delta$ a new time variable and step; in the new time, the Hamiltonian $\bar{H}$ becomes $\frac{1}{\omega}H_A + \frac{1}{\omega}H_B + H_C$. Since the Lie bracket of the Hamiltonian vector fields corresponds to the Poisson bracket of Hamiltonians [44], repeated applications of the Baker-Campbell-Hausdorff formula show that the time-rescaled version of (2),

$$\phi_2^h := \phi_{H_A/\omega}^{h/2} \circ \phi_{H_B/\omega}^{h/2} \circ \phi_{H_C}^h \circ \phi_{H_B/\omega}^{h/2} \circ \phi_{H_A/\omega}^{h/2},$$

corresponds to the symplectic time-h flow of the Hamiltonian $\tilde{H} = \bar{H} + R$, where the perturbative remainder R is defined by

$$R := -\frac{1}{24}h^2\left\{\frac{1}{\omega}H_A, \left\{\frac{1}{\omega}H_A, \frac{1}{\omega}H_B\right\}\right\}$$

$$-\frac{1}{24}h^2\left\{\frac{1}{\omega}H_A, \left\{\frac{1}{\omega}H_A, H_C\right\}\right\}$$

$$+\frac{1}{12}h^2\left\{\frac{1}{\omega}H_B, \left\{\frac{1}{\omega}H_B, \frac{1}{\omega}H_A\right\}\right\}$$

$$-\frac{1}{24}h^2\left\{\frac{1}{\omega}H_B,\left\{\frac{1}{\omega}H_B,H_C\right\}\right\}$$

$$+\frac{1}{12}h^2\left\{H_C,\left\{H_C,\frac{1}{\omega}H_A\right\}\right\}$$

$$+\frac{1}{12}h^2\left\{H_C,\left\{H_C,\frac{1}{\omega}H_B\right\}\right\}$$

$$+\frac{1}{12}h^2\left\{\frac{1}{\omega}H_B,\left\{H_C,\frac{1}{\omega}H_A\right\}\right\}$$

$$+\frac{1}{12}h^2\left\{H_C,\left\{\frac{1}{\omega}H_B,\frac{1}{\omega}H_A\right\}\right\}+O(h^4). \qquad (4)$$

Higher-order methods can be similarly analyzed: it is known that an lth-order integrator based on Hamiltonian splitting samples the exact flow of an $O(h^l)$ perturbation of the exact Hamiltonian, where the perturbation is a sum of terms expressible using at least l nested Poisson brackets of H_A/ω, H_B/ω, and H_C [3]. Therefore, the magnitude of R is at most $O(h^l/\omega)$, because any nonzero nested Poisson bracket has to involve H_A/ω or H_B/ω at least once. (Otherwise, $\{H_C,H_C\}=0$ leads to a zero result.)

2. Assuming q,p,x,y, and R are bounded, conservation of $\tilde{H}$ implies boundedness of H_C, which leads to $\|q-x\|=O(1/\sqrt{\omega})$ and $\|p-y\|=O(1/\sqrt{\omega})$. Under a canonical transformation $\alpha=q-x$, $\beta=p-y$, $\bar{Q}=(q+x)/2$, $\bar{P}=(p+y)/2$,

$$\tilde{H}=\frac{1}{\omega}H\left(\bar{Q}+\frac{\alpha}{2},\bar{P}-\frac{\beta}{2}\right)+\frac{1}{\omega}H\left(\bar{Q}-\frac{\alpha}{2},\bar{P}+\frac{\beta}{2}\right)$$

$$+\frac{1}{2}\|\alpha\|^2+\frac{1}{2}\|\beta\|^2+R$$

for some small $R(\alpha,\beta,\bar{Q},\bar{P})$. Assuming $H(\cdot,\cdot)$ is analytic, then $\tilde{H}$ is analytic in $\bar{Q},\bar{P},\alpha,\beta$. Since $\bar{Q},\bar{P}$ remain bounded, $\alpha(0)=\beta(0)=\mathbf{0}$, and $\alpha(t),\beta(t)=O(1/\sqrt{\omega})$, a multiscale analysis based on normal form shows that $\alpha(t),\beta(t)=O(1/\omega)$ until at least $t=O(\omega^{3/2})$ (see Appendix); this corresponds to $T=O(\sqrt{\omega})$ in the original time, which is still a long time.

3. We now refine the perturbation magnitude estimation. Each time H_A/ω or H_B/ω appears in a nested Poisson bracket

term in (4), that term gets scaled by $1/\omega$. On the other hand, when H_C appears, by the definition of Poisson bracket,

$$\{X,H_C\}=\frac{\partial X}{\partial q}\frac{\partial H_C}{\partial p}-\frac{\partial X}{\partial p}\frac{\partial H_C}{\partial q}+\frac{\partial X}{\partial x}\frac{\partial H_C}{\partial y}-\frac{\partial X}{\partial y}\frac{\partial H_C}{\partial x}.$$

Since derivatives of H_C are $\pm(q-x)$ or $\pm(p-y)$, all of them lead to a scaling by $O(1/\omega)$ too. Therefore, any l-nested Poisson bracket term containing $l+1$ Hamiltonians is actually $O(h^l/\omega^{1+l})$ until at least $t=O(\omega^{3/2})$.

4. Let $\epsilon=h^l/\omega^l$. Then $\tilde{H}-\bar{H}=R=O(\epsilon/\omega)$ until at least $t=O(\omega^{3/2})$, and it is a $O(\epsilon)$ perturbation when compared with $\bar{H}$. A KAM-type estimate (see [3], Chap X) shows that under technical conditions (e.g., diophantine and nondegenerate initial condition) the solution of $\tilde{H}$ differs from that of $\bar{H}$ by $O(t\epsilon)$ for at least $t=O(\min(\epsilon^{-1},\omega^{3/2}))$ (minimum of the two because R estimate is only valid until $t=O(\omega^{3/2})$).

Since $\tilde{H}$ corresponds to the numerical solution and $\bar{H}$ corresponds to the exact solution, the numerical error is thus $O(th^l/\omega^l)$. Converting back to the original time, this corresponds to a numerical error of $O(T\delta^l\omega)$ until at least $T=O(\min(\delta^{-l}\omega^{-1},\omega^{1/2}))$, where δ is the step size used by the proposed integrator (3) in the original time.

A. A numerical demonstration

We now demonstrate the $O(T\delta^l\omega)$ error bound on a 1 degree of freedom system with $H(Q,P)=(Q^2+1)(P^2+1)/2$. This is a system with obtainable exact solution, and thus long time numerical errors can be accurately quantified.

To derive the exact solution, note energy-level sets in this system are closed curves symmetric about $P=0$ and about $Q=0$. A half-period of the dynamics is governed by

$$\dot{Q}=(1+Q^2)\sqrt{\frac{2E}{1+Q^2}-1}, \quad \text{where } P(0)=0,$$

$$E=[1+Q(0)^2]/2, \text{ and } Q(0)<0 \quad \text{is assumed,}$$

and this dynamics is until $\mathcal{T}>0$ such that $Q(\mathcal{T})=-Q(0)$. Its exact solution can be obtained using Jacobi's elliptic function,
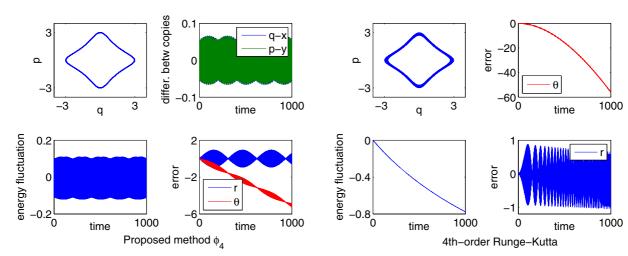

FIG. 1. Comparison between the proposed integrator and a classical nonsymplectic method. Time step $\delta=0.1$ for both methods, $Q(0)=-3$, $P(0)=0$, $\omega=20$, and $T=1000$.

TABLE I. ϕ_4^δ error is proportional to ω. $T = 100$ and $\delta = 0.001$ fixed, $Q(0) = -3, P(0) = 0$.

$\omega =$	20	40	80	160
Max amplitude error $\approx$	6.2×10^{-8}	1.2×10^{-7}	2.5×10^{-7}	5×10^{-7}
Max phase error $\approx$	5.6×10^{-8}	1.1×10^{-7}	2.2×10^{-7}	4.5×10^{-7}

namely,

$$Q(t) = Q(0) \, \text{cn}\left(t\sqrt{1 + Q(0)^2} \left| \frac{Q(0)^2}{1 + Q(0)^2} \right. \right),$$

and thus $\mathcal{T}$ can be obtained using a Gauss hypergeometric function

$$\mathcal{T} = \pi \,_2F_1[0.5, 0.5; 1, -Q(0)^2].$$

(See [45] for more details about these special functions.) The exact solution of H is now available till arbitrary time, because it is $2\mathcal{T}$ periodic, and its other half-period (for time $[(2n + 1)\mathcal{T}, 2n\mathcal{T}], n \in \mathbb{Z}$) is symmetric to the previously obtained half (for time $[2n\mathcal{T}, (2n + 1)\mathcal{T}], n \in \mathbb{Z}$) about the origin.

Figure 1 compares long time simulations by a fourth-order proposed method and the fourth-order Runge-Kutta. Errors of the trajectories are investigated in polar coordinates (column 4 and row 2, column 2). Runge-Kutta has an exponentially increasing phase error and a drifting amplitude error corresponding to undesired numerical viscosity, whereas the proposed method has a linearly growing phase error and a bounded amplitude error, consistent with the error bound but extending much beyond $O(\sqrt{\omega})$ time.

One also sees that δ can indeed be much larger than what is needed for resolving $O(1/\omega)$ time scale oscillations. In addition, ω only needs to be larger than $H(Q(0), P(0))$ in magnitude.

Agreements with other aspects of the error bound are illustrated in Tables I and II. The fact that the error is high order in δ but only first order in ω is another important property, as it leads to simultaneous accuracy and efficiency when l is large ($l = 4$ here).

Figure 2 compares the pioneer method recommended in [34] ($Q\tilde{P}\tilde{Q}P$ with optimized P_1 mixing and projection, second order) with the proposed integrator (Equation (2); here we chose a low-order version so that the focus is on the benefit of symplecticity). Important to recall is that the pioneer method is not symplectic because additional mixing and projection steps were introduced for improved accuracy (without such steps the accuracy time span will be even shorter; results not shown), but such steps breaks symplecticity. It is thus not surprising that the pioneer method remains accurate only for a short time. Note, however, that the pioneer method is symmetric and exhibits no "secular" energy deviation.

B. A second numerical example

Consider the Schwarzschild geodesics problem simulated in [34] (with typos in the Hamiltonian, initial values, and precession rate estimation corrected). The geodesic can be cast as the solution of the 3 degrees of freedom ($Q = [t, r, \phi]$ and $P = [p_t, p_r, p_\phi]$) Hamiltonian system governed by

$$H = \frac{1}{2}\left[\left(1 - \frac{2}{r} \right)^{-1} p_t^2 - \left(1 - \frac{2}{r} \right) p_r^2 - \frac{p_\phi^2}{r^2} \right].$$

Figure 3, left panels, estimates accuracies of the Schwarzschild geodesic computed by second- and fourth-order proposed integrators ϕ_2^δ [Eq. (2)] and ϕ_4^δ [Eq. (3)], as well as that by the pioneer nonsymplectic method recommended in [34] ($Q\tilde{P}\tilde{Q}P$ with P_1 mixing and projection, which was the optimized choice for this problem). Note that accuracy can only be estimated because no exact solution is available for quantifying the numerical error, and we used the adaptive MATLAB ode45 with relative and absolute error tolerances both set to 10^{-20} to generate a benchmark for error estimation. Although the long time fidelity of this benchmark cannot be guaranteed, throughout the simulation it only produced $<7 \times 10^{-15}$ deviation from the conserved energy value. One sees ϕ_4^δ is the most accurate of the three in all aspects, while ϕ_2^δ is less accurate on r than the optimized pioneer method, despite its better energy preservation. Worth mentioning is, for this example, the optimized pioneer method allows larger step sizes than ϕ_l^δ (results not shown).

We then evaluate how each method captures the effect of additional nonconservative forces by adding a simplest-possible dissipation, i.e., consider $\dot{q} = \partial H / \partial p, \dot{p} = \partial H / \partial q - \gamma p$. To add external forces in the proposed and pioneer integrators, we replace the p update in $\phi_{H_A}^\delta$ and the y update in $\phi_{H_B}^\delta$ [Eq. (1)] by

$$p \mapsto p + \delta[-\partial_q H(q, y) - \gamma y],$$
$$y \mapsto y + \delta[-\partial_x H(x, p) - \gamma p].$$

Integration accuracies are again estimated by comparing to an adaptively integrated fine benchmark (Fig. 3, right panels). With dissipation, long time errors of the symplectic versions are much smaller than the pioneer method. The intuition is, as the force breaks time reversibility, the symmetry of the pioneer method is no longer advantageous; symplectic integrators, however, are known to well approximate changes

TABLE II. ϕ_4^δ error is proportional to δ^4. $T = 100$ and $\omega = 20$ fixed, $Q(0) = -3, P(0) = 0$. The $\delta = 0.1$ column anomaly is because $T\delta^l\omega$ is too large to be in the asymptotic regime of the error bound.

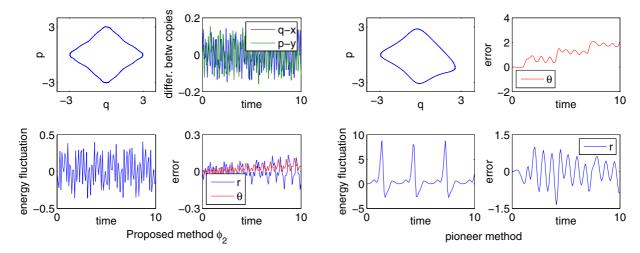| $\delta =$	10^{-1}	$10^{-1.5}$	10^{-2}	$10^{-2.5}$	10^{-3}
Max amplitude error $\approx$	0.76	5.8×10^{-2}	6.1×10^{-4}	6.2×10^{-6}	6.2×10^{-8}
Max phase error $\approx$	0.76	5.2×10^{-2}	5.6×10^{-4}	5.6×10^{-6}	5.6×10^{-8}

FIG. 2. Comparison between a low-order version of the proposed integrator and the pioneer nonsymplectic method recommended in [34]. Time step $\delta = 0.1$ for both methods, $Q(0) = -3, P(0) = 0, \omega = 20$, and $T = 10$.

in phase-space volume [46,47]. In fact, in a longer simulation ($T = 10^5$), the pioneer method became unstable while the proposed methods remained bounded and relatively accurate (results not shown).

III. NONINTEGRABLE SYSTEM: A NUMERICAL DEMONSTRATION ON THE WEAK TURBULENT NONLINEAR SCHRÖDINGER EQUATION

It is known that the nonlinear Schrödinger equation (NLS) is nonintegrable in $\geqslant 2$ spatial dimensions and exhibits weak turbulence [48]. Conditions under which a finite-dimensional Hamiltonian system of Fourier coefficients can approximate the nonlinear Schrödinger equation on 2 torus were rigorously demonstrated in [12], while extensive usages of similar models preceded this rigorous justification [49–53]. This system has a

nonseparable Hamiltonian that can be written as

$$H(q,p) = \frac{1}{4} \sum_{i=1}^{N} \left(q_i^2 + p_i^2 \right)^2 - \sum_{i=2}^{N} \left(p_{i-1}^2 p_i^2 + q_{i-1}^2 q_i^2 \right.$$
$$\left. - q_{i-1}^2 p_i^2 - p_{i-1}^2 q_i^2 + 4 p_{i-1} p_i q_{i-1} q_i \right).$$

Note an explicit symplectic integrator was proposed in [30] for polynomial Hamiltonians, which suit this system. That method requires more computations per time step because it is based on splitting the Hamiltonian into monomials, but it is symplectic in the original phase space, which is certainly an advantage. Our purpose is not to compare that method with the more general method proposed here but only to numerically access whether our method still has good long time performances for a nonintegrable system.
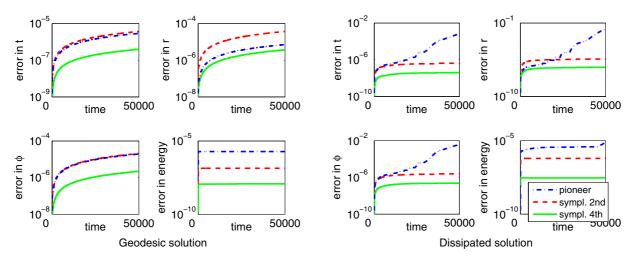

FIG. 3. Comparison between the proposed integrators and the pioneer nonsymplectic method: Schwarzschild geodesic and its dissipated version ($\gamma = 10^{-4}$). Time step $\delta = 0.2$ for both methods, $\omega = 2$, $T = 50\,000$, $M = 10$, $m = 1$, initial condition $Q(0) = [0,20,0], P(0) \approx [0.982,0,-4.472]$ corresponds to initial semimajor axis $a(0) = 20$, eccentricity $e(0) = 0$, $p_r(0) = 0$, $p_\phi(0) = \sqrt{r(0)}$, and $p_t(0)$ solves $H(0) = m^2/2$. Plotted were maxima of relative errors up to given times, respectively scaled by dividing over the Keplerian period $2\pi \sqrt{a(0)^3/M}$ for t, over $a(0)[1 + e(0)]$ for r, over 2π for ϕ, and unscaled for H.
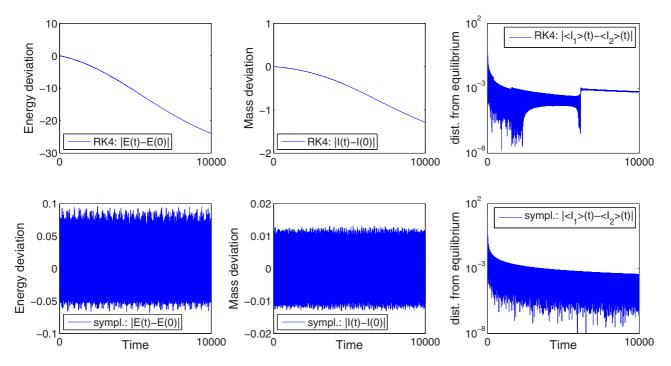
FIG. 4. Long time simulations of NLS with $N = 2$ by fourth-order Runge-Kutta and a proposed integrator ϕ_4^δ. $[q_1,p_1,q_2,p_2](0) = [3,1,0.01,0]$, $\delta = 0.01$, $\omega = 100$. Long time convergence of the trajectory towards the ergodic limit is quantified in the right column based on $\langle I_1 \rangle(t) - \langle I_2 \rangle(t)$ [see Eq. (5)]; this observable was chosen to reduce the interference with the numerical loss of total mass in RK4. Computation done using win32 MATLAB R2010a on x64 Windows 7 i7-4600U CPU.

To do so, we focus on first integrals and the statistical behavior of long time dynamics. If one denotes by $I_i = q_i^2 + p_i^2$ the mass of each mode, the total mass $I := \sum I_i$ can be shown as a second first integral of the system in addition to energy conservation. Also, although mathematically difficult

to prove, it is widely believed the system, due its turbulent nature, is ergodic on first integral foliations. We thus assume an ergodic measure of $\delta[H(q,p) - E]\delta[\sum I_i - I] \, dq \, dp$. It is easy to show that this constrained Liouville measure is at least an invariant measure. Under the ergodicity assumption, long
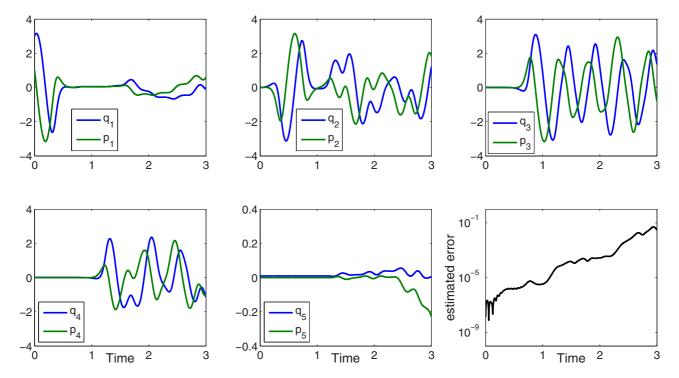

FIG. 5. Short time simulation of NLS with $N = 5$ by ϕ_4. $[q_1,p_1](0) = [3,1]$, $[q_i,p_i](0) = [0.01,0]$ for $i \neq 1$, $\delta = 0.001$ for a closer examination of details, $\omega = 100$. Error is estimated by comparing to fine RK4 simulation with $\delta = 10^{-4}$.
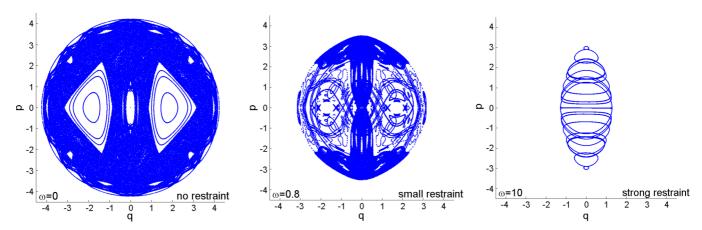
FIG. 6. Poincaré sections of $\bar{H}$ with $H(Q,P) = (Q^2 + 1)(P^2 + 1)/2$ and different ω values at the surface $x = 0$ on constant energy shell $\bar{H} = 10$. Initial conditions are uniformly sampled in the q,p plane. Note curves do intersect because each q,p pair corresponds to two y values.

time averages of phase-space observables converge to their spatial averages with respect to the ergodic measure, and the mode relabeling symmetry of H leads to, if $N = 2$, that

$$\lim_{T \to \infty} \langle I_1 \rangle(T) = \lim_{T \to \infty} \langle I_2 \rangle(T) = \frac{1}{2}I, \quad \text{where } \langle I_i \rangle(T)$$
$$= \frac{1}{T} \int_0^T I_i(t)dt. \tag{5}$$

Figure 4 shows that the proposed method better captures total energy and mass conservations, as well as the assumed convergence toward equilibrium (5). The standard nonsymplectic Runge-Kutta seems to be less accurate in capturing the assumed equilibration, which normally happens at a more consistent rate. In this sense, even though the error analysis in Sec. II does not apply to nonintegrable systems, the proposed integrator still exhibits better long time performance than its nonsymplectic counterpart, at least in this example.

For completeness, Fig. 5 also illustrates an $O(1)$ time simulation of the proposed method for a larger N, where a weak turbulent cascade is clearly observed.

IV. ON INTEGRABILITY OF THE EXTENDED SYSTEM

As discussed above, symplectic integrators have desirable long time performances for integrable systems. (See [3–5] for nonstiff problems, and Sec. II for our specific stiff problem.) However, if the system has a positive Lyapunov exponent, the symplectic integrator can be as bad as a generic integrator, because its error can grow exponentially with time as the local truncation error propagates along.

Unfortunately, the system governed by $H(q,y) + H(x,p)$ in extended phase space may not be integrable even if $H(Q,P)$ is integrable in the original phase space. For instance, consider the 1 degree of freedom problem studied in Sec. II A. The original system is integrable because H is a first integral. Figure 6 illustrates one Poincaré section of the $\omega = 0$ extended system, where large chaotic seas demonstrate nonintegrability. This explains why a symplectic integrator for $\bar{H}$ with $\omega = 0$, as considered in [34], does not have satisfactory performances beyond $O(1)$ time.

However, H_C is integrable. Thus, as ω increases, a larger proportion of the phase space for $\bar{H} = H_A + H_B + \omega H_C$

corresponds to regular behaviors (see [38–43]). Indeed, Poincaré sections in Fig. 6 show smaller chaotic seas when $\omega = 0.8$ and no evidence of chaotic sea but only invariant tori when $\omega = 10$. This suggests that the integrability assumption in Sec. II is reasonable as long as the original system is integrable and ω is larger than a threshold ω_0, with the exception of initial conditions whose measure vanishes.

APPENDIX: REFINED ESTIMATES OF α AND β

This section shows that α, β are not just $O(1/\sqrt{\omega})$ but in fact $O(1/\omega)$ at least for a long time. As this is a self-contained section, we set up the problem in the original time, denote by t the corresponding time variable (as opposed to T in the main text) and by dot the t derivative; the Hamiltonian then is

$$\hat{H} = H\left(\bar{Q} + \frac{\alpha}{2}, \bar{P} - \frac{\beta}{2}\right) + H\left(\bar{Q} - \frac{\alpha}{2}, \bar{P} + \frac{\beta}{2}\right)$$
$$+ \frac{1}{2}\omega \|\alpha\|^2 + \frac{1}{2}\omega \|\beta\|^2 + \mathcal{R}. \tag{A1}$$

We will introduce a near identity transformation to express the governing Hamilton's equations in a nonstandard second-order normal form, which we will utilize to refine the estimates of α, β. To begin, note ωJ is the coefficient matrix associated with the linear dynamical system generated by $H_0 = \frac{1}{2}\omega\alpha^2 + \frac{1}{2}\omega\beta^2$, where $J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$, and $\exp(t\omega J) = O(1)$ for all real t. We use $x \in \mathcal{R}^{2d}$ to represent $[\alpha, \beta]$.

Lemma A1. Given a $2d$-dimensional real symmetric-matrix valued function $S(s)$,

$$\Omega(s) := \frac{1}{2\pi} \int_0^{2\pi} e^{-J\tau} J S(s) e^{J\tau} \, d\tau$$

is a real skew-symmetric-matrix-valued function. Moreover, its associated fundamental matrix $\Phi(s)$, defined as the solution of

$$\Phi'(s) = \Omega(s)\Phi(s), \quad \Phi(0) = I,$$

satisfies $\Phi(s) = O(1)$ and $\Phi(s)^{-1} = O(1)$ for all s.

Proof. Assume S in block form is $S = \begin{bmatrix} A & B \\ B^T & D \end{bmatrix}$ where $A^T = A$, $D^T = D$. Then

$$\Omega = \frac{1}{2\pi} \int_0^{2\pi} \begin{bmatrix} I\cos t & -I\sin t \\ I\sin t & I\cos t \end{bmatrix} \begin{bmatrix} B^T & D \\ -A & -B \end{bmatrix}$$
$$\times \begin{bmatrix} I\cos t & I\sin t \\ -I\sin t & I\cos t \end{bmatrix} dt$$
$$= \frac{1}{2}\begin{bmatrix} B^T - B & A + D \\ -(A + D) & B - B^T \end{bmatrix},$$

and it is real skew-symmetric.

To bound $\Phi(s)$, note $(\Phi^T\Phi)' = (\Omega\Phi)^T\Phi + \Phi^T(\Omega\Phi) = 0$ and thus $\|\Phi(s)\| = \|\Phi(0)\| = 1$. $\Phi(s)^{-1}$ can be similarly bounded since it is easy to verify that $d\Phi^{-1}/ds = -\Phi^{-1}\Omega$. ∎

Theorem A2. Consider in $\mathcal{R}^{2d}$ an ODE

$$\dot{x} = \omega J x + F_0(t) + F_1(t)x + \sum_{i=1}^{2d} x^T F_{2i}(t)x\mathbf{e}_i + O(x^3),$$

where solutions are assumed to exist and x satisfies $x = O(1/\sqrt{\omega})$, $x(0) = \mathbf{0}$, $\mathbf{e}_i$'s are standard basis of $\mathcal{R}^{2d}$, F_0, F_1, F_{2i} have bounded derivatives, and $F_1(t) = JS(t)$ for some symmetric-matrix-valued $S(t)$. Then

$$x(t) = O(1/\omega), \quad \text{until at least } t = O(\sqrt{\omega}).$$

Proof. Let $y(t) = \exp(-\omega Jt)x(t)$, then

$$\dot{y} = e^{-\omega Jt}F_0(t) + e^{-\omega Jt}F_1(t)e^{\omega Jt}y$$
$$+ e^{-\omega Jt}\sum_{i=1}^{2d}y^T e^{-\omega Jt}F_{2i}(t)e^{\omega Jt}y\mathbf{e}_i + O(\omega^{-3/2}).$$

Let $\epsilon = 1/\omega$ introduce a dummy slow variable $s = t$ and a fast variable $\tau = \omega t$ (which corresponds to the angle associated with the rotation in α, β) and use a prime to denote $d/d\tau$, then the governing equation rewrites as a slow–fast system

$$y' = \epsilon e^{-J\tau}F_0(s) + \epsilon e^{-J\tau}F_1(s)e^{J\tau}y$$
$$+ \epsilon e^{-\omega Jt}\sum_{i=1}^{2d}y^T e^{-J\tau}F_{2i}(s)e^{J\tau}y\mathbf{e}_i + O(\epsilon^{5/2})$$

$$s' = \epsilon$$
$$\tau' = 1. \tag{A2}$$

We look for a near-identity transformation in the form of

$$y = z + \epsilon u_1(z,s,\tau) + \epsilon^2 u_2(z,s,\tau) + O(\epsilon^{5/2}),$$

where u_1, u_2 are 2π-periodic in τ, such that

$$z' = \epsilon f_1(z,s) + \epsilon^2 f_2(z,s) + \epsilon S + O(\epsilon^{5/2})$$

$$S := e^{-\omega Jt}\sum_{i=1}^{2d}z^T e^{-J\tau}F_{2i}(s)e^{J\tau}z\mathbf{e}_i$$

for some f_1, f_2 independent of the fast variable τ.

Under this transformation, the right-hand side of (A2) becomes

$$\epsilon e^{-J\tau}F_0(s) + \epsilon e^{-J\tau}F_1(s)e^{J\tau}z + \epsilon^2 e^{-J\tau}F_1(s)e^{J\tau}u_1$$
$$+ \epsilon e^{-\omega Jt}\sum_{i=1}^{2d}z^T e^{-J\tau}F_{2i}(s)e^{J\tau}z\mathbf{e}_i$$
$$+ \epsilon^2 e^{-\omega Jt}\sum_{i=1}^{2d}[z^T e^{-J\tau}F_{2i}(s)e^{J\tau}u_1$$
$$+ u_1^T e^{-J\tau}F_{2i}(s)e^{J\tau}z]\mathbf{e}_i + O(\epsilon^{5/2})$$
$$= \epsilon e^{-J\tau}F_0(s) + \epsilon e^{-J\tau}F_1(s)e^{J\tau}z$$
$$+ \epsilon^2 e^{-J\tau}F_1(s)e^{J\tau}u_1 + \epsilon S + O(\epsilon^{5/2}),$$

where the last equality is due to $z = y + O(\epsilon) = O(1)x + O(\epsilon) = O(\epsilon^{1/2})$.

The left-hand side of (A2), on the other hand, becomes

$$z' + \epsilon\frac{\partial u_1}{\partial z}z' + \epsilon\frac{\partial u_1}{\partial s}s' + \epsilon\frac{\partial u_1}{\partial \tau}\tau' + \epsilon^2\frac{\partial u_2}{\partial z}z' + \epsilon^2\frac{\partial u_2}{\partial s}s'$$
$$+ \epsilon^2\frac{\partial u_2}{\partial \tau}\tau' + O(\epsilon^{5/2})$$
$$= \epsilon f_1 + \epsilon^2 f_2 + \epsilon S + \epsilon\frac{\partial u_1}{\partial z}\epsilon f_1 + \epsilon\frac{\partial u_1}{\partial z}\epsilon S + \epsilon^2\frac{\partial u_1}{\partial s}$$
$$+ \epsilon\frac{\partial u_1}{\partial \tau} + \epsilon^2\frac{\partial u_2}{\partial \tau} + O(\epsilon^{5/2})$$
$$= \epsilon f_1 + \epsilon^2 f_2 + \epsilon S + \epsilon^2\frac{\partial u_1}{\partial z}f_1 + \epsilon^2\frac{\partial u_1}{\partial s} + \epsilon\frac{\partial u_1}{\partial \tau}$$
$$+ \epsilon^2\frac{\partial u_2}{\partial \tau} + O(\epsilon^{5/2}),$$

where the last equality is due to that $S = O(z^2) = O(\epsilon^{-1})$.

By matching $O(\epsilon)$ and $O(\epsilon^2)$ terms, respectively, we obtain the following requirements on u_1, u_2, f_1, f_2:

$$f_1 + \frac{\partial u_1}{\partial \tau} = e^{-J\tau}F_0(s) + e^{-J\tau}F_1(s)e^{J\tau}z, \tag{A3}$$

$$f_2 + \frac{\partial u_1}{\partial z}f_1 + \frac{\partial u_1}{\partial s} + \frac{\partial u_2}{\partial \tau} = e^{-J\tau}F_1(s)e^{J\tau}u_1. \tag{A4}$$

In order for a solution u_1 periodic in τ to exist, f_1 has to satisfy a solvability condition

$$2\pi f_1(z,s) = \int_0^{2\pi}\left[e^{-J\tau}F_0(s) + e^{-J\tau}F_1(s)e^{J\tau}z\right]d\tau,$$

which is obtained after integrating both sides of (A3) over τ. The first term of the integrand vanishes after integration, and Lemma A1 then leads to

$$f_1(z,s) = \Omega(s)z$$

for some real skew-symmetric Ω. Integration then gives

$$u_1 = e^{-J\tau}J F_0(s) + g(\tau,s)z$$

for some $g(\tau,s)$ periodic in τ.

The next order equation (A4) leads to a solvability condition

$$2\pi f_2(z,s) = \int_0^{2\pi}\left(e^{-J\tau}F_1(s)e^{J\tau}u_1 - \frac{\partial u_1}{\partial s} - \frac{\partial u_1}{\partial z}f_1\right)d\tau,$$

and thus $f_2 = -\langle \frac{\partial g}{\partial s} \rangle z - \langle g \rangle \Omega z$, where $\langle \cdot \rangle$ indicates time average with respect to τ.

Consequently, z satisfies

$$z' = \epsilon \Omega(s) z + \epsilon e^{-\omega J t} \sum_{i=1}^{2d} z^T e^{-J\tau} F_{2i}(s) e^{J\tau} z \mathbf{e}_i$$

$$- \epsilon^2 \left\langle \frac{\partial g}{\partial s} \right\rangle z - \epsilon^2 \langle g \rangle \Omega z + O(\epsilon^{5/2}).$$

Rescale back to the original time so that the right-hand side gets divided by ϵ. Let $Z(t) = \Phi(t)^{-1} z(t)$, where Φ is the fundamental matrix associated with Ω. Then

$$\dot{Z} = \Phi(t)^{-1} e^{-\omega J t} \sum_{i=1}^{2d} [\Phi(t) Z(t)]^T e^{-J\tau} F_{2i}(s) e^{J\tau} [\Phi(t) Z(t)] \mathbf{e}_i$$

$$- \epsilon \Phi(t)^{-1} \left[\left\langle \frac{\partial g}{\partial s} \right\rangle + \langle g \rangle \Omega \right] \Phi(t) Z(t) + O(\epsilon^{3/2}).$$

Note Φ and Φ^{-1} are bounded according to Lemma A1. Rewrite this in integral form, and bounding one of the two Z's in the first (quadratic) term by $Z = O(z) = O(\epsilon^{1/2})$ we have

$$\|Z(t)\| \leqslant \int_0^t (\epsilon^{1/2} C_1 \|Z(\hat{t})\| + \epsilon C_2 \|Z(\hat{t})\| + \epsilon^{3/2} C_3) \, d\hat{t}$$

for some $C_1, C_2, C_3 > 0$ independent of ϵ. Then Gronwall's lemma leads to

$$\|Z(t)\| \leqslant \epsilon^{3/2} t C_3 \exp[(\epsilon^{1/2} C_1 + \epsilon C_2) t],$$

which means $Z(t) = O(\epsilon)$ until at least $t = O(\epsilon^{-1/2})$. Since Z, z, y, x are at the same order due to the boundedness of Φ and $\exp(\omega J t)$, $x(t) = O(\epsilon) = O(1/\omega)$ until at least $t = O(\epsilon^{-1/2}) = O(\sqrt{\omega})$.

Theorem A3. Consider the dynamics generated by $\hat{H}$ [Eq. (A1)], where $\mathcal{R}$ is given by a sum of nested Poisson brackets (see step 1 in Sec. II). If $H(\cdot, \cdot)$ is analytic, ω is large enough, $\alpha(0) = \beta(0) = \mathbf{0}$, $\alpha(t), \beta(t) = O(\omega^{-1/2})$ until at least $t = O(\sqrt{\omega})$, and $\bar{Q}(t), \bar{P}(t)$ remain bounded independent of ω until at least the same time, then $\alpha(t), \beta(t) = O(\omega^{-1})$ until at least the same time.

Proof. Let $x = [\alpha, \beta]$, then

$$\dot{x} = J \frac{\partial H}{\partial x}.$$

Since H and $\alpha^2 + \beta^2$ are analytic, $\mathcal{R}$ is also analytic because it is a sum of their nested canonical Poisson brackets. Therefore, $\hat{H}$ can be locally written as

$$\hat{H} = H_0(\bar{Q}, \bar{P}) + H_1(\bar{Q}, \bar{P}) x + \frac{1}{2} x' H_2(\bar{Q}, \bar{P}) x$$

$$+ \sum_{n=3}^{\infty} H_n(\bar{Q}, \bar{P})[x^{\otimes}],$$

where $H_1(\bar{Q}, \bar{P})$ is a vector, $H_2(\bar{Q}, \bar{P})$ is a symmetric matrix, and $H_n(\bar{Q}, \bar{P})[x^{\otimes n}]$ stands for a homogeneous polynomial of degree n in x, with coefficients being analytic functions of $\bar{Q}, \bar{P}$.

Therefore, the α, β dynamics can be locally written as

$$\dot{x} = \omega J x + F_0(\bar{Q}, \bar{P}) + F_1(\bar{Q}, \bar{P}) x + \sum_{i=1}^{2d} x^T F_{2i}(\bar{Q}, \bar{P}) x \mathbf{e}_i$$

$$+ \sum_{n=3}^{\infty} F_n(\bar{Q}, \bar{P})(x^{\otimes n}),$$

for some F_1, F_{2i}, and F_n. In particular, $F_1(\bar{Q}, \bar{P}) = J H_2(\bar{Q}, \bar{P})$ for symmetric H_2.

Since $\bar{Q}, \bar{P}$ are functions of t and bounded (independent of ω) until at least $t = O(\omega^{1/2})$, $\dot{\bar{Q}}, \dot{\bar{P}}$, which are given by Hamilton's equations as functions of $\bar{Q}, \bar{P}, \alpha, \beta$, are also bounded. Consequently, $F_n(\bar{Q}, \bar{P})$ are implicitly functions of time [i.e., $F_n(t)$] with bounded first time derivatives.

Therefore, Theorem A2 applies, and thus $x = O(\omega^{-1})$ until at least $t = O(\omega^{1/2})$. ∎

Note we worked at the level of equations but not Hamiltonians, i.e., we did not express (A1) as a nearly integrable system by transforming α, β into action and angle. This is because we are interested in the special initial condition $\alpha(0) = \beta(0) = 0$, which corresponds a singularity of the transformation, and it will break the analyticity of $\hat{H}$ and render classical approaches like KAM or Nekhoroshev's method difficult to apply.

[1] J. M. Sanz-Serna, Acta Numerica **1**, 243 (1992).

[2] J. E. Marsden and M. West, Acta Numerica **10**, 357 (2001).

[3] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*, 2nd ed. (Springer, Berlin, 2006).

[4] M.-P. Calvo and E. Hairer, Appl. Numer. Math. **18**, 95 (1995).

[5] G. Quispel and C. Dyt, Phys. Lett. A **242**, 25 (1998).

[6] R. D. Ruth, IEEE Trans. Nucl. Sci. **30**, 2669 (1983).

[7] K. Feng, J. Comput. Math. **4**, 279 (1986).

[8] H. Yoshida, Phys. Lett. A **150**, 262 (1990).

[9] J. Wisdom and M. Holman, Astron. J. **102**, 1528 (1991).

[10] R. I. McLachlan and G. R. W. Quispel, Acta Numerica **11**, 341 (2002).

[11] B. Leimkuhler and S. Reich, *Simulating Hamiltonian Dynamics*, Vol. 14 (Cambridge University Press, Cambridge, UK, 2004).

[12] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao, Invent. Math. **181**, 39 (2010).

[13] J. Henrard and A. Lemaitre, Celest. Mech. **30**, 197 (1983).

[14] N. Murray and M. Holman, Astron. J. **114**, 1246 (1997).

[15] G. Li, S. Naoz, B. Kocsis, and A. Loeb, Astrophys. J. **785**, 116 (2014).

[16] G. Li, S. Naoz, M. Holman, and A. Loeb, Astrophys. J. **791**, 86 (2014).

[17] E. Forest, J. Phys. A: Math. Gen. **39**, 5321 (2006).

[18] H. Qin and X. Guan, Phys. Rev. Lett. **100**, 035006 (2008).

[19] M. Tao, J. Comput. Phys. **327**, 245 (2016).

[20] W. S. Koon, M. W. Lo, J. E. Marsden, and S. D. Ross, Chaos **10**, 427 (2000).

[21] S. D. Bond, B. J. Leimkuhler, and B. B. Laird, J. Comput. Phys. **151**, 114 (1999).

[22] J. B. Sturgeon and B. B. Laird, J. Chem. Phys. **112**, 3474 (2000).

[23] J. Worthington, Ph.D. thesis, University of Sydney, 2012.

[24] L. D. Landau, *The Classical Theory of Fields*, Vol. 2 (Elsevier, New York, 2013).

[25] A. Buonanno, Y. Chen, and T. Damour, Phys. Rev. D **74**, 104005 (2006).

[26] A. Kol, B. B. Laird, and B. J. Leimkuhler, J. Chem. Phys. **107**, 2580 (1997).

[27] M. Tao and H. Owhadi, IMA J. Numer. Anal. **36**, 80 (2016).

[28] Z.-X. Huang and X.-L. Wu, Int. J. Quantum Chem. **106**, 839 (2006).

[29] J. Sanz-Serna and M. Calvo, *Numerical Hamiltonian Problems*, 1st ed. (Chapman and Hall/CRC Press, Boca Raton, FL, 1994).

[30] S. Blanes, Phys. Rev. E **65**, 056703 (2002).

[31] R. I. McLachlan and G. R. W. Quispel, BIT Numer. Math. **44**, 515 (2004).

[32] S. A. Chin, Phys. Rev. E **80**, 037701 (2009).

[33] Y. K. Wu, E. Forest, and D. S. Robin, Phys. Rev. E **68**, 046502 (2003).

[34] P. Pihajoki, Celest. Mech. Dyn. Astron. **121**, 211 (2015).

[35] E. Forest, AIP Conf. Proc. **184**, 1106 (1989).

[36] M. Suzuki, Phys. Lett. A **146**, 319 (1990).

[37] E. Hairer, S. P. Nørsett, and G. Wanner, *Solving Ordinary Differential Equations I*, 2nd ed., Springer Series in Computational Mathematics Vol. 8 (Springer-Verlag, Berlin, 1993), pp. xvi+528.

[38] A. Kolmogorov, Dokl. Akad. Nauk SSSR **98**, 527 (1954).

[39] V. I. Arnol'd, Russ. Math. Surv. **18**, 9 (1963).

[40] J. Moser, Nachr. Akad. Wiss. Göttingen, II, 1 (1962).

[41] J. Pöschel, Commun. Pure Appl. Math. **35**, 653 (1982).

[42] L. Eliasson, Math. Phys. Electron. J. **2**, 1 (1996).

[43] A. Giorgilli and U. Locatelli, Z. Angew. Math. Phys. **48**, 220 (1997).

[44] J. E. Marsden and T. S. Ratiu, *Introduction to Mechanics and Symmetry*, 2nd ed. (Springer, New York, 2010), p. 584.

[45] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions, 1965* (Dover Publications, New York, 1964).

[46] N. Bou-Rabee and H. Owhadi, SIAM J. Numer. Anal. **48**, 278 (2010).

[47] S. Ober-Blöbaum, M. Tao, M. Cheng, H. Owhadi, and J. E. Marsden, J. Comput. Phys. **242**, 498 (2013).

[48] S. Dyachenko, A. Newell, A. Pushkarev, and V. Zakharov, Phys. D **57**, 96 (1992).

[49] D. Cai, A. J. Majda, D. W. McLaughlin, and E. G. Tabak, Proc. Natl. Acad. Sci. USA **96**, 14216 (1999).

[50] D. Cai, A. J. Majda, D. W. McLaughlin, and E. G. Tabak, Phys. D **152-153**, 551 (2001).

[51] A. Majda, D. McLaughlin, and E. Tabak, J. Nonlinear Sci. **7**, 9 (1997).

[52] P. A. Milewski, E. G. Tabak, and E. Vanden-Eijnden, Stud. Appl. Math. **108**, 123 (2002).

[53] R. DeVille, P. A. Milewski, R. J. Pignol, E. G. Tabak, and E. Vanden-Eijnden, Commun. Pure Appl. Math. **60**, 439 (2007).