MULTISCALE PROBLEMS IN MECHANICS: SPIN DYNAMICS, STRUCTURE-PRESERVING INTEGRATION, AND DATA-DRIVEN METHODS

A Dissertation
Presented to
The Academic Faculty

By

Renyi Chen

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Mathematics
College of Sciences

Georgia Institute of Technology

August 2021

MULTISCALE PROBLEMS IN MECHANICS: SPIN DYNAMICS, STRUCTURE-PRESERVING INTEGRATION, AND DATA-DRIVEN METHODS

Thesis committee:

Prof. Molei Tao, Advisor
School of Mathematics
Georgia Institute of Technology

Prof. Luca Dieci
School of Mathematics
Georgia Institute of Technology

Prof. Gongjie Li, Co-Advisor
School of Physics
Georgia Institute of Technology

Prof. Sung Ha Kang
School of Mathematics
Georgia Institute of Technology

Prof. Rafael de la Llave
School of Mathematics
Georgia Institute of Technology

Date approved: July 7, 2021

. . . We are no bigger than a gnat in the universe or a grain of corn in the vast ocean. Our life is brief and evanescent, while I envy the eternity of time like the unending flow of this great river. . . . The water passes continually by, and yet it is always here. The moon waxes and wanes but it always remains the same moon. If you look at the changes that take place in the universe, there is nothing in it that lasts more than a fraction of a second. But if you look at the unchanging aspect of things, then you realise that both the things and ourselves are immortal. . . .

-Su Shi, Former Ode on the Red Cliffs, 1082 (translated by Lin Yutang in "The Gay Genius: The Life and Times of Su Tungpo", 1947)

To my family.

ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest gratitude to my advisor Professor Molei Tao, who guided me through the journey of exploring the magic of computational methods with intrinsic mathematical beauties. His constant support and help during my whole graduate study make this thesis possible. Beyond all those inspiring discussions with him (in the office during 2017-2020 and virtually on the screen during 2020-2021), I really appreciate his patience and encouragement in teaching and showing me how to think, write and present academically.

I would also like to thank my co-advisor Professor Gongjie Li, for being so encouraging and helpful all the time. The papers and this thesis would not be possible without her tremendous help. Discussions with her and Prof. Tao together of the astrophysical problems seasoned with a mixture of physical intuitions and mathematical rigors were always interesting and exciting.

Many thanks go to Professor Luca Dieci, who was the very first Professor of Georgia Tech I met in the corridor of Skiles even before I personally met him as my initial advisor. His teaching, guiding and mentoring in the first year of my graduate school inspired me a lot even at that time I was totally new to research and the graduate study. I also appreciate his support as my committee member for the oral exam and the dissertation defense. I am grateful to Professor Rafael de la Llave, for serving my committee member and his valuable suggestions, delightful discussions about my research works. I really appreciate Professor Sung Ha Kang for accepting being my committee member and I enjoyed several interesting casual talks with her about research and some other interesting topics.

Special thanks are due to the friends and fellow graduate students who accompanied me during the unforgettable 5-years-long graduate school journey all the way along. I would like to thank Adrián, Alex, Bhanu, Daniyar, Dantong, Hao, Haodong, He, Hongyu, Jiaqi, Jieun, Jose, Joan, Mengyi, Michail, Ruilin, Shasha, Shu, Tongzhou, Weiwei, Xiao, Xiaofan, Xin, Yian, Yifeng, Yuqing, Zhibo and many more for valuable discussions. And allow me to thank my friends Hongyu, Jiaqi, Jiyuan, Longmei, Mark, Michail, Ruilin, Shijie, Shu, Tongzhou, Weiwei, Xia, Xianda, Xiaofan, Yian, Yifeng, Yuqing, Zili and many more for those enjoyable board game nights and outdoor adventures.

My parents, far away on the other side of the Earth, thank you for the sweet home and your love from the day I was born. For the past several years, deep in my mind, it was always the happy moments of planting and harvesting in the backyard, playing around in the vast country fields and making small objects in the magic engineering attic with you that drove me forward to exploring more of the world.

In the end, for all the great people and furry friends I met in the peach state, I really appreciate and will miss the beautiful days spent here with you, “so long, and thanks for all the fish”.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

This thesis focuses on analyzing the physics and designing multiscale methods for nonlinear dynamics in mechanical systems, such as those in astronomy. The planetary systems (e.g., the Solar System) are of great interest as rich dynamics of different scales contribute to many interesting physics. Outside the Solar System, a bursting number of exoplanets have been discovered in recent years, raising interest in understanding the effects of the spin dynamics to the habitability. In part I of this thesis, we investigate the spin dynamics of circumbinary exoplanets, which are planets that orbit around stellar binaries. We found that habitable zone planets around the stellar binaries in near coplanar orbits may hold higher potential for stable climate compared to their single star analogues. And in terms of methodology, secular theory of the slow dominating dynamics is calculated via averaging. Beyond analyzing the dynamics mathematically, to simulate the spin-orbit dynamics for long term accurately, symplectic Lie-group (multiscale) integrators are designed to simulate systems consisting of gravitationally interacting rigid bodies in part II of the thesis. Schematically, slow and fast scales are tailored to compose efficient algorithms. And the integrators are tested via our package `GRIT`. For the systems that are almost impossible to simulate (e.g., the Solar System with the asteroid belt), how can we understand the dynamics from the observations? In part III, we consider the learning and prediction of nonlinear time series purely from observations of symplectic maps. We represent the symplectic map by a generating function, which we approximate by a neural network (hence the name GFNN). And we will prove, under reasonable assumptions, the global prediction error grows at most linearly with long prediction time as the prediction map is symplectic.

CHAPTER 1

INTRODUCTION

Since “the Big Bang” around 13.8 billion years ago [1], the universe has been evolving never-ending – for days and nights, minutes and seconds. Zooming in the evolution of the Earth, dynamics of different scales have been constantly changing our world – in the millions to billions of years’ scale, the tectonic plates drift and collide, forming the landscape of the Earth’s surface [2]; in the thousands to millions of years’ scale, ice sheets advance and retreat during cycles of glacial periods, dynamically changing the climate of the Earth [3, 4]; every year, the Earth revolves around the Sun once, bringing a variety of weathers for life on the Earth; furthermore, the changing and evolution of our world take place every moment - in tiny scales of seconds, milliseconds etc. Since the mechanics of our universe are naturally evolving with fast and slow scales mixed together, properly modeling and understanding the multiscale dynamics in methodology and in specific scientific questions are important both mathematically and physically.

Over the past several centuries, the Solar System was discovered and understood by humanity step by step through observations, analyses, simulations and explorations. However, in the mechanical view, the Solar System is still far from fully understood because of its rich dynamics of different scales. For example, to understand the orbital dynamics of the planets in the Solar System, beyond the point mass planetary interactions, in the long run, how the asteroid belt and the planetary spins affect the orbital dynamics are still unclear. To understand the spin-orbit dynamics, simulations of systems consisting of rigid bodies with long term accuracies is the holy grail. However, simulation is not always an appropriate approach – the asteroid belt, consisting of millions of asteroids, is almost impossible to simulate. A possible approach to tackle those problems is to understand the dynamics (e.g., the asteroid belt) by fusing machine learning with domain knowledge.

Towards understanding not only our Solar System, but also the exoplanetary systems, three projects are explored in the following chapters in terms of analyzing the spin dynamics, designing accurate symplectic integrators for spin-orbit coupled systems and predicting the (Hamiltonian) dynamics from observations.

In Chapter 2, the spin dynamics of Earth-like planets in circumbinary systems are modeled and analyzed to study the effects of the spin-axis variations to the stability of the planet's climate. In this system, periodic / oscillatory dynamics exist in different scales: from fast spins, orbital evolutions to slow precessions, and five major dynamics are list here – 1. The planet self-rotates periodically. 2. The inner binary stars revolve periodically. 3. The planet orbits around the binary stars every orbital period. 4. The planetary orbit precesses, reflected in the oscillation of orbital inclinations and the orbit's longitude of ascending node. 5. The spin-axis of the planet precesses, evolving the phase of the spin. Due to the rich dynamics described above, resonances need to be carefully studied to reveal the underlying physics. Here, we are interested in the spin-orbit resonance of the last two angles above, and this resonance is the main reason of the planets' large obliquity variations (e.g., the Moonless Earth [5, 6], the Mars [7]). For circumbinary systems, we find that the large quadrupole potential of the stellar binary could speed up the planetary orbital precession, and detune the system out of spin-orbit resonances compared to their single star analogues. This leads to very small obliquity variations for planets that reside near the same plane as the stellar binaries. Mathematically, we carefully average the fast angles [8] and isolate the dominating slow dynamics of the spin axis – our secular theory agrees well with the numerical simulations (see Chapter 3 for our symplectic integrators).

In Chapter 3, we develop symplectic Lie-group (multiscale) integrators to simulate systems consisting of gravitationally interacting rigid bodies to study the spin-orbit dynamics of planetary systems (e.g., the circumbinary systems described in Chapter 2). Three different classes of schemes, namely the $\mathcal{T}$-series, the $\mathcal{M}$-series, and the $\mathcal{K}$-series methods, are designed via different splitting and composition strategies. Specifically, tailored splittings ($\mathcal{M}$-series methods and $\mathcal{K}$-series methods) are utilized in the stage of splittings to divide the slow and fast scales, therefore a better trade-off between efficiency and accuracy of numerical schemes can be achieved. Numerically, the conservation properties and convergences of our integrators are tested via our package `GRIT`[1]. As a demonstration, this package is applied to Trappist-I. It shows that the differences in transit timing variations due to spin-orbit coupling could reach a few min in ten year measurements, and strong planetary perturbations can push Trappist-I f, g and h out of the synchronized states.

In Chapter 4, instead of analyzing dynamics using traditional mathematical tools (Chapter 2) / numerical methods (Chapter 3), the fashion of our age, machine learning, is incorporated to investigate the dynamics purely from data. Specifically, we consider the problem of learning the underlying dynamics from a nonlinear time series generated by a latent symplectic map. A special case is Hamiltonian systems, whose solution flows give such symplectic maps. For this special case, both generic approaches based on learning the vector field of the latent ODE and specialized approaches based on learning the Hamiltonian that generates the vector field exist. Our method, however, is different as it does not rely on the vector field nor assume its existence; instead, it directly learns the symplectic evolution map in discrete time. Moreover, we do so by representing the symplectic map via a generating function, which we approximate by a neural network (hence the name GFNN). This way, our approximation of the evolution map is always exactly symplectic. This additional geometric structure allows the local prediction error at each step to accumulate in a controlled fashion – the global prediction error grows at most linearly with long prediction time for an integrable system. In contrast, the error grows exponentially for non-symplectic predictions. In addition, as a map-based and thus purely data-driven method, GFNN avoids two additional sources of inaccuracies common in vector-field based approaches, namely the error in approximating the vector field by finite difference of the data, and the error in numerical integration of the vector field for making predictions. The effectiveness of our

---

[1] https://github.com/GRIT-RBSim/GRIT

algorithm and the validity of our theory on reduced error are also demonstrated experimentally on various types of dynamical systems.

CHAPTER 2

SPIN DYNAMICS OF CIRCUMBINARY PLANETS

2.1 Introduction

Orientation of planetary spin axis plays an important role in the climate of a planet. For Earth, the spin-orbit misalignment, known as obliquity, varies between 21.4° and 24.4°, and this together with the variations in the orbit's eccentricity (oscillating up to ~ 0.06) are the main causes of the Milankovitch cycles of Earth climate variations, and may lead to glacial cycles [3, 9, 10, 11]. For Mars, the spin axis exhibits wild variation that can reach $\sim 60^\circ$ [12, 5, 7], which has drastic effects: with a high obliquity of $\sim 45^\circ$, Mars' atmosphere could have partially collapsed due to CO_2 condensation; this is consistent with the glacier-like landform discovered in the tropics and mid-latitude of Mars [13, 14].

Beyond the Solar System, the rapidly growing number of detected exoplanets calls for a better understanding of the uniqueness of life and searches for habitable worlds in the universe. Inspired by the discovery that, if Moon did not exist, Earth spin-axis would be chaotically varying with obliquity ranging between $\sim 0 - 50^\circ$ at the billion-year timescale [15, 16], this article joins the contemporary effort of studying Exo-Milankovitch cycles by quantifying obliquity variations in exoplanetary systems. Although it was predicted based on energy balance models that the habitable region can be increased if a planet has a high obliquity [17], terrestrial planets with massive CO_2 atmospheres (typically expected in the outer regions of habitable zones) can be subject to snowball transitions and experience partial atmospheric collapse [18, 19]. Therefore, a large obliquity variation could limit the habitability of an Earth-like planet.

Are the spin-axes dynamics of exoplanets typically stable, or do they also require a massive moon for the stabilization, like Earth? Detailed analysis has been conducted for particular exoplanets residing in the habitable zone and with one host star. For instance, Kepler-186f likely has a stable spin axis and does not need a moon, due to their weak couplings with its companion planets [20]. In addition, the maximum extent of chaotic regions of obliquity has then been estimated for a wide range of exoplanetary systems [21].

Although planets orbiting around single stars, analogous to those in Solar System, have been studied in great details, it was unclear how planetary spin-axes would vary if the planet orbits around stellar binaries. This is actually a meaningful question because roughly half of Solar-type stars are in binaries [22, 23]. Moreover, planet orbiting around both of the binary components are in fact common (c.f. planet orbiting around just one star of the binary, whose obliquity variation has been studied in [24]): over a dozen such planets have been discovered [25, 26], and the occurrence rate of these planets are similar to those around single stars [27, 28]. Thus, it is important to better understand the habitability of circumbinary planets, and we do so from the perspective of spin-axis dynamics.

In general, large obliquity variations manifest due to resonances between gravitational torque from the host star and the orbital perturbations from a planet's companions. Specifically, if there were just one star, the planetary spin axis is torqued by the only star. This would lead to spin-axis precession in the same way as the precession of a top under Earth's gravity. If the spin axis precession frequency could coincide with the orbital oscillation frequency, the spin axis can vary with large amplitude due to the spin-orbit resonance. However, because in the single star case the planetary orbit is almost constant over time with no planetary companions[1], there is no spin-orbit resonances and obliquity will be a constant. Having two stars in the center completely changes the physics, because (i) there are torques from both stars, and (ii) the planetary orbit is no longer stationary. Therefore, the obliquity variation can be very different from the well understood case, and the traditional framework for analyzing single star systems cannot be directly applied. Therefore, this article develops a secular theory for the obliquity variations of circumbinary planets

---

[1] It would be exactly constant as the solution to a two-body problem if the planet and star were point masses

and investigates their detailed evolutions. Results are also confirmed and strengthened by a newly developed numerical tool [29] and its simulations.

2.2 Notations

M_{*_i}	The mass of the i-th Star ($i = 1, 2$)
δ	$\frac{M_{*_2}}{M_{*_1}+M_{*_2}}$
m_p	The mass of the planet
μ_p	$\frac{m_p(M_{*_1}+M_{*_2})}{M_{*_1}+M_{*_2}+m_p}$
μ_*	$\frac{M_{*_1}M_{*_2}}{M_{*_1}+M_{*_2}}$
$\boldsymbol{I}_p = \begin{bmatrix} I_p^{(1)} & 0 & 0 \\ 0 & I_p^{(2)} & 0 \\ 0 & 0 & I_p^{(3)} \end{bmatrix}$	The (standard) moment of inertia tensor of the planet
$\mathcal{R}$	The radius of the equator of the planet
$\mathcal{D}_* = \{l_*, g_*, h_*, L_*, G_*, H_*\}$	The Delaunay variables of the inner orbit
$\mathcal{D} = \{l_d, g_d, h_d, L_d, G_d, H_d\}$	The Delaunay variables of the outer orbit
$\mathcal{A} = \{g_a, h_a, l_a, G_a, H_a, L_a\}$	The Andoyer variables of the planet

Spin-Axis Angle and Obliquity. The spin-axis angle is the angle between the planetary spin-axis (i.e. $\overrightarrow{\boldsymbol{G}_a}$) and the normal to the reference plane (i.e. $\boldsymbol{E}_3$) as shown in Figure 2.2. The obliquity is the angle between the spin-axis and the normal to the planetary orbital plane. Variations in both orbital orientation and the spin-axis affect the obliquity.

2.3 Main Results

Different from planets around a single star, we find that spin-axis variations of circumbinary planets are typically low. This is because the stellar binary in the center produces a much larger quadrupole moment, which leads to faster orbital precession comparing to the cases of a single star. This detunes the system from spin-orbit resonances and leads to low variations in planetary spin-axis. section 2.3.1 uses a heuristic calculation to illustrate this intuition, and then we will obtain more accurate and detailed results using both analytical secular theory and full rigid-body simulations. More precisely, the spin-axis dynamics admits different behaviors over separated timescales: the fast timescale corresponds to rotations along the orbit, and the slower timescales correspond to spin-axis and orbital variations. We average out the faster timescales to obtain the secular theory. We start with the single planet case in section 2.3.2 for simplicity, where section 2.3.2 presents our secular theory, and its results are verified using rigid-body simulations in section 2.3.2. Our results are then generalized to multi-planet systems in section 2.3.3: an analytical approach based on the Lagrange-Laplace method for multiplanetary systems is described in section 2.3.3, followed by systematic numerical investigations in section 2.3.3.

2.3.1 Heuristic Calculation

For a planet orbiting around a stellar binary, its orbit precesses around the orbit of the stellar binary, and the spin-axis precesses around the planetary orbital orientation. When the orbital precession frequency matches that of the spin-axis, spin-orbit resonance occurs which can drive large obliquity variations.

The frequency of the orbital precession can be estimated as [30]:

$$\dot{h_d} = \frac{3}{4\cos(\delta i)} n \left(\frac{a_b}{a_p}\right)^2 \frac{M_{*1}M_{*2}}{(M_{*1}+M_{*2})^2}, \tag{2.1}$$

where δi is the mutual inclination between the orbit of the stellar binary and that of the planet, and n is the orbital frequency, a_b and a_p are the semi-major axes of the stellar binary and the planet, and M_{*1} and M_{*2} are the masses of the stellar binary.

On the other hand, the planet spin-axis precession can be estimated too. In the case of single host star, this precession frequency due to the torque from a central star can be

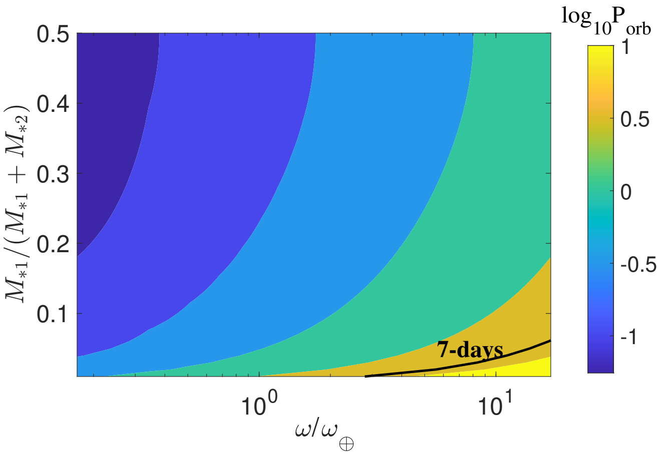


Figure 2.1: Stellar binary period (P_{orb} in days) that leads to planet spin-orbit resonances. Spin-orbit resonances could only occur for closely separated stellar binaries ($P_{orb} \lesssim 1$ day), unless the mass ratio of the stellar binary is below $\lesssim 0.05$ for fast rotating planets with rotating period of $\lesssim 5$ hours.

expressed as $\alpha \cos \epsilon$ [5], where

$$\alpha = \frac{3G}{2\omega}\frac{M_*}{a_p^3}E_d, \tag{2.2}$$

ϵ is the spin-orbit misalignment (obliquity), ω is the spin rate of the planet, M_* is the mass of the central star and E_d is the dynamical ellipticity of the planet. For Earth-like planets (with mass and interior structure similar to Earth), $E_d = E_{d,\oplus}(\omega/\omega_\oplus)^2$, where $E_{d,\oplus} = 3350 \times 10^{-6}$ and $\omega_\oplus$ are the dynamical ellipticity and rotation rate of the Earth separately.

To obtain a rough estimation of the spin-precession frequency in the case around a stellar binary, we substitute the mass of the central star with the total mass of the stellar binary, while assuming the average distance to each of the stars are roughly $\sim a_p$. Then, the spin-axis e.g. the Moonless Earth, Mars precession frequency becomes $3n^2/(2\omega)E_d$.

The separation of the stellar binary components and their masses determines the quadrupole moment and the possibilities of spin-orbit resonances for Earth-like planets. Thus, assuming the planet is Earth-like and located at one Earth flux, we can estimate the period of the stellar binary that could drive to slow enough orbital precession to excite spin-

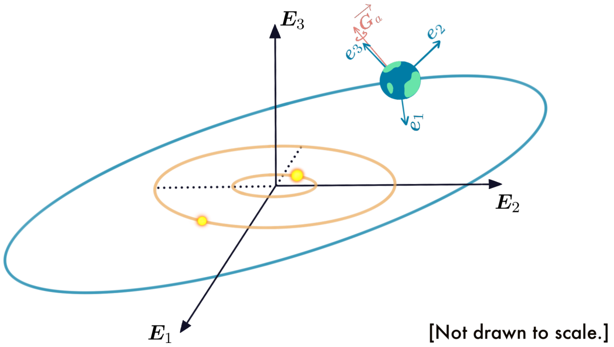


Figure 2.2: The demonstration of a planet orbiting around the stellar binary.

orbit resonances. Setting the total mass of the stars to be one solar mass for illustration, Fig.Figure 2.1 plots the stellar binary period as a function of the planets rotation rate ($\omega/\omega_{\oplus}$) and the stellar binary mass ratio. We assume the luminosity of the stars follows the mass-luminosity relation ($L \propto m^{3.5}$) for simplicity. The qualitative results are the same for different total masses of the stellar binaries. The maximum rotation rate we included in the plot is the breakup angular velocity ($\sqrt{Gm_p/R_p^3}$).

Fig.Figure 2.1 shows that spin-orbit resonances could only occur when stellar binary period is less than 1 day, unless the mass ratio of the stellar binary is very small ($\lesssim 0.05$) and when the planet rotates much faster ($\gtrsim 5\omega_{\oplus}$). However, so far, no circumbinary planets have been detected that orbit stellar binary with a period less than seven days. This is likely due to the formation processes that leads to the orbital decay of the stellar binary to short periods via Kozai-Lidov oscillations [27, 31, 32]. Thus, it is challenging to excite the obliquity of circumbinary planets via spin-orbit resonances.

2.3.2 Single Planet Around Binary Stars

Secular Theory

Existing spin-axis secular theories in the literature focused on single star systems, where the mean anomaly increases linearly with time [33]. This is no longer the case for circumbinary planets, whose mean anomalies change at time-varying rates. To correctly characterize the long term effective behavior, we adopt rigorous normal-form-based treatment for averaging multiple nonlinear angles, which physically corresponds to averaging over time instead of the mean anomaly of a planet in a Hamiltonian setup[2].

In order to accurately characterize the spin dynamics, the planet is modeled as a rigid body instead of a point mass. The problem set up is shown in Fig. 2.2, where the two yellow dots represent the stellar binary and the blue one represents the planet. $E_{1,2,3}$ are the principal axes of the reference frame[3], and $e_{1,2,3}$ are those of the planetary body where e_3 aligns with the symmetry body axis.

The original, unaveraged Hamiltonian in canonical variables (Delaunay variables for the orbit and Andoyer variables for the spin-axis) is denoted by

$$\begin{aligned} & H\left(\mathcal{A}, \mathcal{D}, l_*, L_*\right) \\ =& T^{linear}\left(\mathcal{D}\right) + T^{rot}\left(\mathcal{A}\right) + V\left(\mathcal{A}, \mathcal{D}, l_*, L_*\right) \end{aligned} \tag{2.3}$$

where T^{linear} is the linear kinetic energy of the rigid body, T^{rot} is the rotational kinetic energy of the rigid body Eq. (2.25):

$$T^{rot}(\mathcal{A}) = \frac{H_a^2}{2I_p^{(1)}} + \frac{G_a^2 - H_a^2}{2I_p^{(3)}}. \tag{2.4}$$

and V is the gravitational potential generated by the binary stars on the oblate planet, l_* is the mean anomaly of the inner orbit, and L_* is its conjugate momentum.

Appropriately averaging over the stellar binary's orbital phase, planet's orbital phase and planet spin, we can obtain the secular dynamics of the planetary obliquity, governed by the following ODE system (see sections 2.4.1 and 2.4.2 for derivation):

---

[2] How to do so is well studied [8, 34]. Two distinctions between this work and the rich field of restricted 3-body dynamics [35, 36, 37, 38] to which these tools also apply are: (i) the latter considers point masses only, thus no spin; (ii) there is no/negligible orbital resonance in our considered physical parameter range.

[3] We will follow the convention used for Andoyer variable and call an inertial frame the reference frame.

$$\begin{cases} \dot{X} = \sin(h)\left[C_1 X\sqrt{1-X^2} + 4C_2\cos(h)(1-X^2)\right], \\ \dot{h} = \dfrac{C_1\cos(h)\left(1-2X^2\right)}{\sqrt{1-X^2}} - 2C_2 X\cos(2h) + C_3 + 2C_4 X, \end{cases} \tag{2.5}$$

where $X = H_a/G_a$ is the cosine value of the angle between spin's angular momentum and $\boldsymbol{E}_3$, and $h = h_d - h_a$ is the phase difference between the precessions of the planet's orbit and the planetary spin. Similar to the single star case [15], X, h together give a secular approximation of the planetary spin dynamics. We note that planetary orbital dynamics is not affected by the spin-axis variations in the secular limit. This can be understood intuitively since the angular momentum of the planetary spin is much lower than that of the orbit. However, the spin-axis evolution is sensitive to the orbital oscillations. In particular, in the secular limit for planets around the stellar binary, planetary orbital shape (semi-major axis a and eccentricity e) and inclination are constants over time, and the orbital orientation (argument of pericenter g_d and h_d) are changing slowly due to the stellar binary quadrupole potential. C_{1-4} in our secular equation of motion (eq.Equation 2.5) depend on the constant parameters of the planetary orbit and the masses of the planet and the stars, and the spin-axis variations also depend on the nodal precession of the orbit.

A remark is that the dynamics above (Equation 2.5) corresponds to Hamiltonian

$$\begin{aligned} H(X,h) = & C_1\sqrt{1-X^2}X\cos(h) + C_2\left(1-X^2\right)\cos(2h) \\ & + C_3 X + C_4 X^2. \end{aligned} \tag{2.6}$$

For circumbinary planets in physical parameter range, the value of C_3 is much bigger than those of $C_{1,2,4}$ (see e.g., Fig. 2.3) unless the planet's orbit is nearly orthogonal to the binary's. These values, according to our secular equations of motion (Equation 2.5), lead to X being nearly a constant and h linearly changing with time (note $h \in \mathbb{T}$, not $\mathbb{R}$), which corresponds to fast precession of the spin axis along the orbital normal direction and near

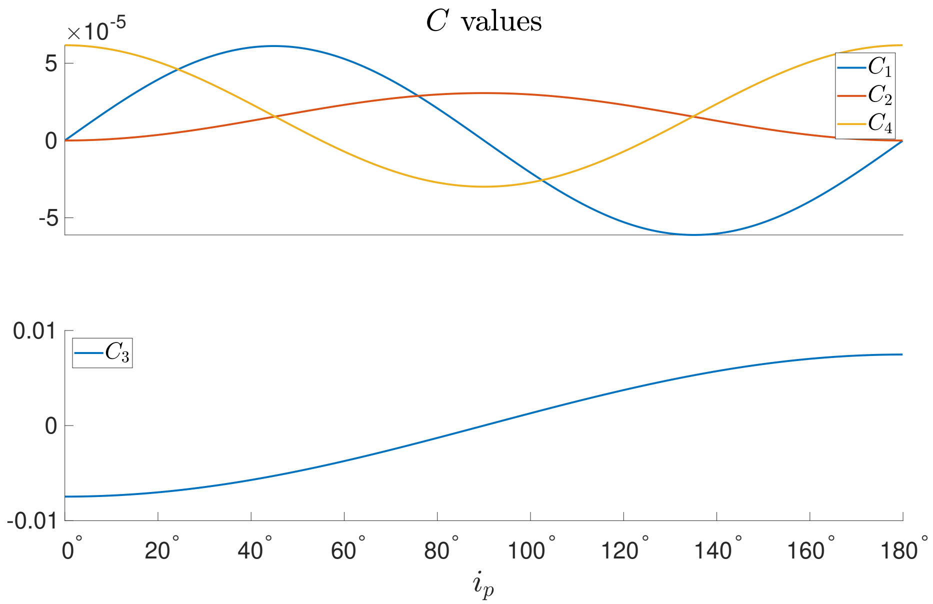


Figure 2.3: C_1,C_2,C_3,C_4 change with respect to mutual inclination i_p of *Kepler-47* system.

constant obliquity. This is how the developed secular theory elucidates the low obliquity variation.

We now illustrate a fuller picture of what the secular theory can say about the effect of stellar binary on the planetary spin dynamics. To be concrete, we use *Kepler-47* system ($a_* = 0.0836AU$, $M_{*_1} = 1.043M_\odot$, $M_{*_2} = 0.362M_\odot$) as an example, but the dynamics and bifurcations are topologically the same for general circumbinary systems. Varying one parameter (i_p) of the system and keeping others fixed, the spin dynamics undergoes a series of bifurcations and switches between topologically different dynamics with $2, 4, 6$ fixed points (the $2, 4, 6$ fixed points cases respectively correspond to $i_p = 0°$ to the 1st bifurcation point near $i_p = 89°$, the 1st to the 2nd bifurcation point near $i_p = 89.99°$, and the 2nd bifurcation point to $i_p = 90°$ in Fig. 2.4).

Comparing to the single star system, the circumbinary system has an additional bifurcation. More precisely, the spin dynamics in the ≤ 1 saddle point parameter regime (from $i_p = 0°$ to the 2nd bifurcation point in Fig. 2.4) are similar to the single star system [21] as both bifurcations are due to the $1 : 1$ spin-orbit resonance [39, 21]; however, the regime with two saddle points is unique to the circumbinary system. In particular, a second saddle

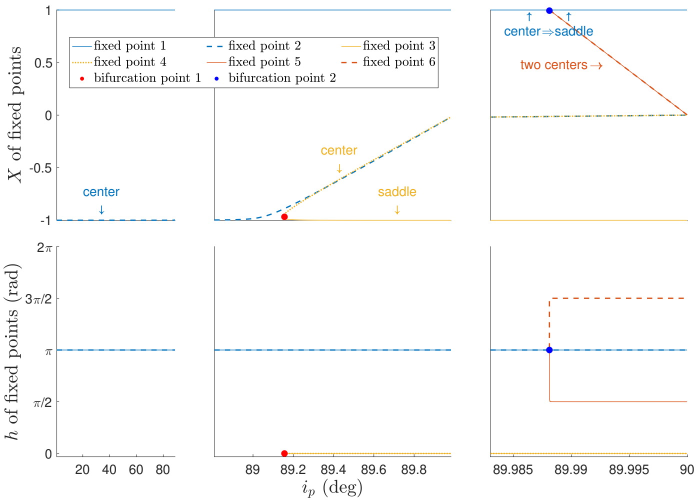


Figure 2.4: Bifurcation diagram with the varying parameter being i_p. Three plots with different scalings of i_p axes are concatenated together. Each family of fixed points is denoted by a color. Dots indicate bifurcation locations. **Bifurcation point 1** is a Hamiltonian saddle-node bifurcation; **Bifurcation point 2** is a Hamiltonian pitch-fork bifurcation.

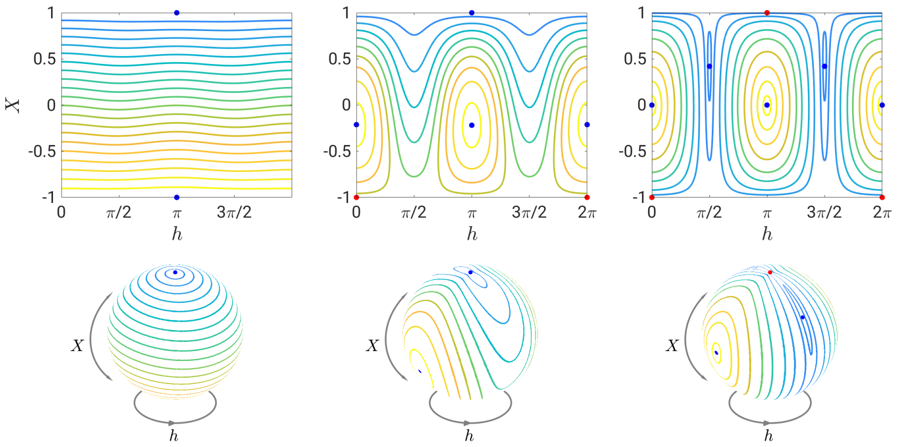


Figure 2.5: Phase diagrams of the (X, h) dynamics in Eq. (2.5) with the varying parameter i_p (in Fig. 2.4) fixed as $i_p = 80°$, $89.8°$, $89.995°$ from the 1st column to the 3rd column respectively. The phase portraits in the second row are the maps of corresponding phase portraits above in spheres with X the latitude coordinate and h the longitude coordinate. Dots indicate fixed pts, blue being nonlinear centers and red being saddles.

point is created out of a previously stable fixed point when the mutual inclination between the inner and the outer orbits reaches nearly 90°, as the consequence of a Hamiltonian pitchfork bifurcation. Two additional centers also appear from the bifurcation, corresponding to configurations for which torques from the two stars cancel out. For illustration, phase portraits of fixing three different i_p values in Fig. 2.4 are plotted in Fig. 2.5 with i_p before the 1st bifurcation, between the 1st and the 2nd bifurcation and after the 2nd bifurcation respectively from the left to the right. To better observe the spin dynamics in 3D space, the phase diagrams in spheres in Fig. 2.5 are wrapped from the 2D phase portraits, demonstrating the trajectories of the spin direction (with respect to the planetary orbit) as the level sets in the sphere. Specifically, in Fig. 2.4, from the 1st column to the 2nd column, one center and one saddle emerge from around $(-1, 0)$ and this bifurcation is due to the $1:1$ spin-orbit resonance (it is common to single star systems); from the 2nd column to the 3rd column, the fixed point around $(\pi, 1)$ changes from a stable center to an unstable saddle with two centers symmetric to $h = \pi$ emerge around it due to the 2nd bifurcation and the two emerged centers take place around $(\pi/2, 0.4)$, $(3\pi/2, 0.4)$ in the phase diagram of the 3rd column.

Mathematically this corresponds to the case where C_3 is no longer dominating. Physically, this corresponds to a previously stable configuration changing into an unstable one, accompanied by two libration-like stable regions emerging nearby, encircled by larger scale transports outside. It happens especially when the planetary orbits are largely tilted with respect to the plane of the stellar binary.

It is unclear yet how physical a near 90° inclination is. However, our main claim, namely the reduced obliquity variation due to the existence of two host stars, is due to the quantitative behavior of the secular dynamics (Equation 2.5) for inclinations away from 90° (not the emergence of a 2nd bifurcation). The new behavior after the 2nd bifurcation (near 90°) is currently just a theoretical prediction and a by-product (our main result is still the modest obliquity variation in the low inclination cases).

Full-body Simulations

An accurate and efficient numerical method for simulating gravitationally interacting rigid bodies (as opposed to point masses) was recently developed based on symplectic Lie group integrators [29]. We conduct such simulations, which will be termed as full-body simulations, for two purposes: (i) to validate our secular theory, and (ii) to demonstrate that stabilization of the long-term obliquity dynamics is a rather general phenomenon.

More specifically, this section will first detail a case study and then sample through the parameter space of circumbinary planetary systems to quantify the ubiquity of stable obliquity dynamics. For the relevance to habitability, we will focus on Earth-like planets in near circular orbits. The planet mass, semi-major axis/rotation period, eccentricity, and oblateness are set to be the same as that of the Earth; however, a wide range of inclinations is considered.

The case study is based on a concrete example: Fig. 2.6 quantifies the spin axis variation of an *Earth-like* planet in a circumbinary system analogous to *Kepler-47* ($a_* = 0.0836AU$, $M_{*_1} = 1.043M_\odot$, $M_{*_2} = 0.362M_\odot$) [40]. Results based on both rigid body dynamics simulations (solid line) and our secular approximation (dashed line; see section 2.4.2 for details of this theory) were both included. Different colors represent different mutual inclinations between the planet and the stellar binary.

The upper panel of Fig. 2.6 presents the results of the *spin-angle*, which measures the tilt of the spin-axis relative to the invariable plane (the plane normal to the total angular momentum of the system). It illustrates the dynamical variation of the spin-axis. The solid lines represent the full body simulation and the dashed lines represent our secular results in figure Fig. 2.6, and it shows that the secular approximation agrees very well with that of the rigid body simulations. Fig. 2.6 shows that the *spin-angle* variations remain very small ($\lesssim 1^\circ$) for all mutual inclinations. This suggests that the effect of a stellar 'perturber' is different from that of planetary perturbers; for example, both Moon-less Earth and Mars are known to have large and chaotic *spin-angle* variations as a result of spin-orbit resonances due to the perturbation of the planetary companions [15, 7]. Similar to the results based on the secular theory, when orbiting around a stellar binary, the spin precession frequencies are much faster than the orbital precession frequency due to the large quadrupole potential of to the stellar binary in the center. Thus, the spin-orbit resonances are absent, and the *spin-angle* only has very small oscillations.

The example in Fig. 2.6 shows that for a near coplanar circumbinary system analogous to *Kepler-47*, both the *spin-angle* and the planetary obliquity variations are small, due to the lack of spin-orbit resonances. Does this feature persist for a single planet orbiting around generic stellar binaries? Typical ranges of the properties of stellar binaries hosting circumbinary planets are not yet well understood, due to the limited sample size of the observed transiting systems (11 so far). However, the understanding so far include that (i) the stellar binaries hosts typically have orbital periods longer than seven days, which are larger than those of eclipsing stellar binaries ($\sim 3days$) [27]; (ii) the mass ratio of the binaries is consistent with that of the stellar binaries in the field (roughly uniform) [41]. In order to make a robust claim, we consider a broad range of binary configurations.

Specifically, we enumerate stellar binary configurations by varying $\frac{M_{*2}}{M_{*1}} \in [0.01, 1]$ and $a_* \in [0.05, 0.3]$. We set the sum of the stellar binaries to be one solar mass for an intuitive comparison with Solar system. The lower limit of the mass ratio is set so that the binary components are both with stellar masses. The minimum semi-major axis corresponds to orbital periods of ~ 4 days, and the maximum semi-major axis corresponds to orbital period of ~ 60days. The maximum semi-major axis is set so that a planet at 1AU remains stable [42]. Then, we calculate the amplitude of the spin-angle variations of an Earth-like planet with 1° inclination (from the orbital plane of the stellar binary) using our secular theory to illustrate the effects quantitatively.

Results are summarized in Fig. 2.7. The spin angles mostly have variations less than 1° for stellar mass ratio larger than 0.01 and stellar separation larger than $\gtrsim 0.05$AU (corresponding to a binary period $\gtrsim 4$ days). This suggests that the latitudinal distribution of

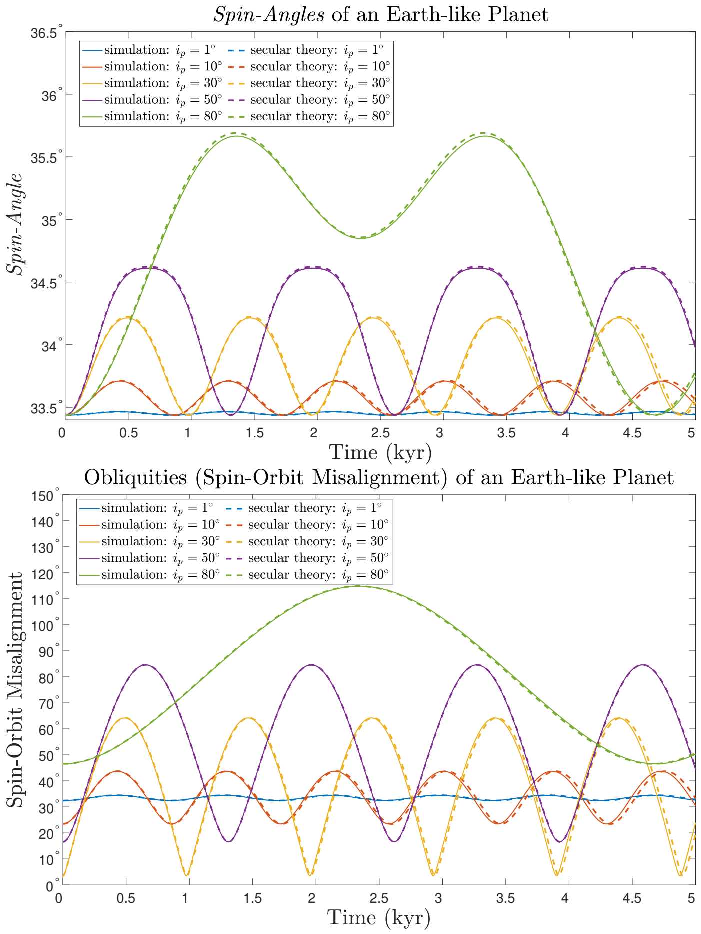


Figure 2.6: *Spin-Angle* (upper panel) and obliquity (lower panel) versus time for an Earth-like planet orbiting around a binary star. The binary stars' orbital parameters were set according to the stellar binary properties of *Kepler-47* system, and the planet was chosen to be *Earth-like*. Multiple colors represent different mutual inclinations between the planet and the stellar binary, and the solid / dashed lines represent results from rigid body simulation / secular theory. It shows that the spin angles have only small variations, and the obliquity variations are mainly due to planetary orbital variations.

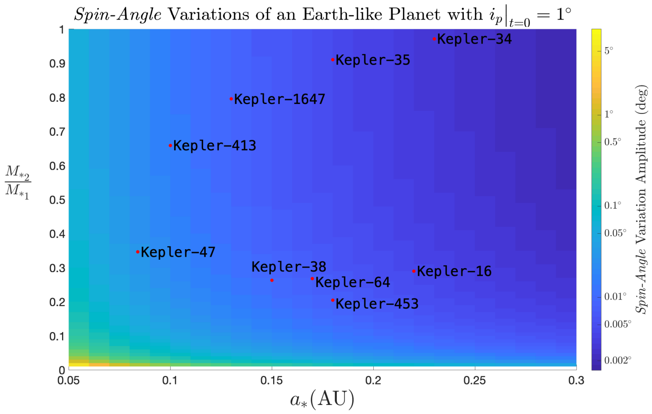


Figure 2.7: *Spin-Angle* variations of circumbinary planets around different types of stellar binaries. The x-axis is the semi-major axis of the orbit of the stellar binary, and the y-axis is the mass ratio of two stars. The color represents the largest *spin-angle* variation for an *Earth-like* planet. Red dots represent currently observed circumbinary planetary systems. It illustrates that the amplitudes of the spin angle variations are very small ($\lesssim$ the variation of the Earth) for a wide range of binary systems, except when the binary separation and mass ratio are both low and the quadrupole moment of the binary is small, corresponding to the lower left corner of the figure.

the stellar radiations have small variations for a wide range of Earth-like planets around stellar binaries in the near co-planar configurations. This phenomenon can be understood intuitively, because higher stellar binary mass ratio and separation both lead to a larger quadrupole momentum at the center of the planetary orbit. This drives fast orbital precession and avoids spin-orbit resonances and variations in the planetary spin-axis. Because circumbinary planets are more likely formed in a coplanar configuration around wide orbit stellar binaries above $\sim$ 7 days [31, 32], it is unlikely to have large spin-axis variations for near co-planar Earth-like planets around stellar binaries.

2.3.3 Circumbinary Systems with Multiple Planets

Similar to single star systems, it is common to have multiple planets in a circumbinary system. An example is the observed Kepler-47 system (at least 3 planets). Interactions between planets can give extra perturbations to their orbital dynamics and increase the likelihood of spin-orbit resonances and large amplitude obliquity variations. For instance,

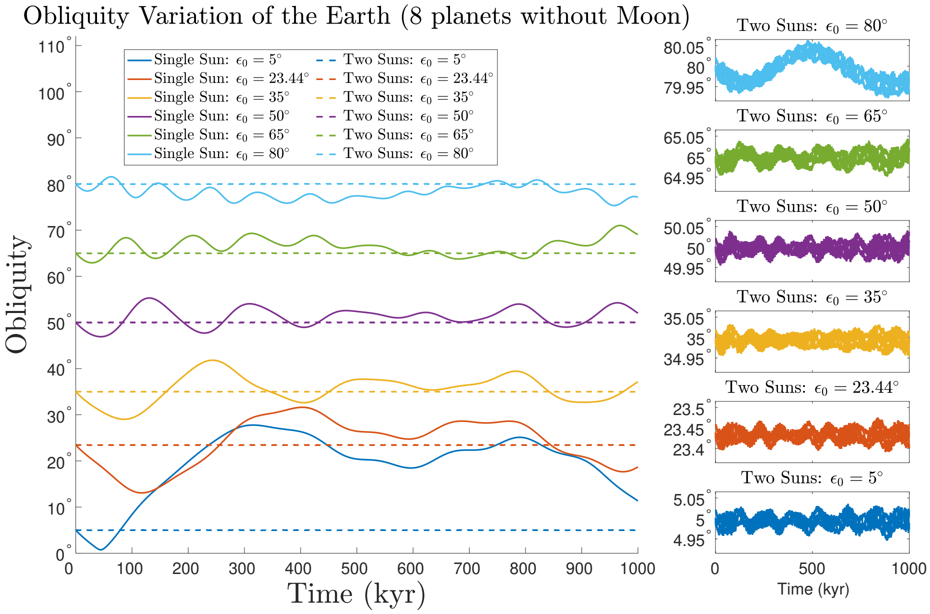


Figure 2.8: Obliquity variations of the Moon-less Earth, orbiting Sun (solid) and orbiting a stellar binary (dash), with companion planets. Similar to the results in [5, 6], the Moon-less Earth has large obliquity variations when obliquity is below 50°. On the other hand, if orbiting a binary, the Earth's obliquity would be stable even without the help of the Moon.

for Earth, perturbations of the terrestrial planets would introduce orbital oscillations that resonate with the precession of Earth's spin-axis and lead to large amplitude obliquity variations, if Moon weren't there to stabilize the spin dynamics [15, 43].

How variable would the spin-axis be, if Earth orbited around a stellar binary with its Solar System planetary companions? To focus on the stabilizing effect of the binary, we consider the moon-less Earth with its seven companion planets in the solar system. The stellar binary are assumed to be in the elliptic plane. We arbitrarily set the stellar binary to be composed of a $0.7M_\odot$ star and a $0.3M_\odot$ star, so the sum of the masses is the same as our own Sun. We set the semi-major axis of the binary to be 0.05AU, so that the effects of the stellar binary is strong comparing with that of the planets but the separation is not wide enough to create instability in millions of years (this was numerically verified).

Fig. 2.8 shows the obliquity variations of the Earth-like planet orbiting around a stellar binary (dashed lines) and around a single star, obtained from our full-body simulations. Different colors correspond to different initial obliquities. The case of the moonless Earth agrees with the secular results and the rigid body simulations in the literature [15, 6], which show large amplitude obliquity oscillations when the obliquity is below $\sim 50°$. Unlike the case of a single-star Moon-less Earth, the obliquity is nearly stable for all initial obliquities around the stellar binary. It shows the Earth's obliquity can be stabilized by the stellar binary even without the Moon.

How general is this stabilization effect? We now investigate it both analytically (section 2.3.3) and numerically (section 2.3.3).

Analytical Results

In single star systems (e.g., the Solar System), the oscillation frequency of a planet's inclination can be estimated using Laplace-Lagrange method [44]. Here, we combine Laplace-Lagrange with our circumbinary secular theory to approximate the orbits of multiple circumbinary planets, so that whether obliquity variations are large can be predicted [5]. Under the common physical approximation that planetary perturbation on the stellar binary is negligible, we assume the stellar binary to have fixed Keplerian orbits. The Hamiltonian of the system is the following:

$$H(\boldsymbol{q},\boldsymbol{p}) = \sum_{i=1}^{n} \frac{\boldsymbol{p}_i^T \boldsymbol{p}_i}{2m_i} - \frac{GM_{*1} m_i}{\|\boldsymbol{q}_i - \boldsymbol{q}_{*1}\|} - \frac{GM_{*2} m_i}{\|\boldsymbol{q}_i - \boldsymbol{q}_{*2}\|} + \mathfrak{R}_i^{planets}, \tag{2.7}$$

where q_i, p_i are respectively planet i's position and momentum, q_{*_1} and q_{*_2} are time-dependent locations of the stars, M_{*_1} and M_{*_2} are stellar masses, the planets [44].

In order to obtain a nearly-integrable form amenable to analysis, we split the gravitational potential between the planets and the stars into two parts and rewrite the Hamiltonian as:

$$H(\boldsymbol{q},\boldsymbol{p}) = \sum_{i=1}^{n} H_i^{kepler} + \left[\mathfrak{R}_i^{*1} + \mathfrak{R}_i^{*2} + \mathfrak{R}_i^{planets}\right]. \tag{2.8}$$

Here H_i^{kepler} is the Hamiltonian of a two-body system, composed of two stars merged at their center of mass and the i^{th} planet, which alone would produce a Keplerian orbit. $\mathfrak{R}_i^{*1}$ and $\mathfrak{R}_i^{*2}$ are disturbing functions modeling (exact) corrections of the gravitation potential generated by the binary due to separations from their center of mass. $\mathfrak{R}_i^{planets}$ are again inter-planet potentials.

We then decompose the perturbative non-integrable part $\sum_{i=1}^{n}\left(\mathfrak{R}_i^{*1} + \mathfrak{R}_i^{*2} + \mathfrak{R}_i^{planets}\right)$ into two groups: the group due to having two stars $\sum_{i=1}^{n}(\mathfrak{R}_i^{*1} + \mathfrak{R}_i^{*2})$, and the group due to having companion planets: $\sum_{i=1}^{n}\mathfrak{R}_i^{planets}$.

Under the influence of the stellar binary (the first group), the planetary orbit would precess with a near constant angular frequency (D_i), in the low eccentricity and low inclination limit (see [45], and SI) , where the orbital inclination I_i is fixed and the longitude of ascending node Ω_i decreases at constant rate D_i. Then, the dynamical evolution under the disturbing potential of $\sum_{i=1}^{n}(\mathfrak{R}_i^{*1} + \mathfrak{R}_i^{*2})$ follows the expression below:

$$\begin{cases} \dfrac{dh_i}{dt} = D_i k_i, \\ \dfrac{dh_j}{dt} = 0, \quad j \neq i. \\ \dfrac{dk_i}{dt} = -D_i h_i, \\ \dfrac{dk_j}{dt} = 0, \quad j \neq i. \end{cases} \tag{2.9}$$

where $h_i = I_i \cos(\Omega_i)\,,\ k_i = I_i \sin(\Omega_i)$.

On the other hand, the approximated dynamics due to planetary companions (the latter group) can be expressed via Laplace-Lagrange theory as described in [44]:

$$\begin{cases} \dfrac{d\boldsymbol{h}}{dt} = \boldsymbol{A} \cdot \boldsymbol{k} \\ \dfrac{d\boldsymbol{k}}{dt} = -\boldsymbol{A} \cdot \boldsymbol{h} \end{cases} \tag{2.10}$$

where $\boldsymbol{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \in \mathbb{R}^{n\times n}$ is a constant matrix that depends on planetary semi-major axes and masses of planets, satisfying $\sum_j a_{ij} = 0$.

For a first-order perturbative approximation, it can be computed that the effective contributions of these two nonresonant fast processes (quadrupole contribution of the binary and planet-planet interaction) are additive, and we have the following approximated dynamics

$$\begin{cases} \dfrac{d\boldsymbol{h}}{dt} = \boldsymbol{B}\cdot\boldsymbol{k} \\ \dfrac{d\boldsymbol{k}}{dt} = -\boldsymbol{B}\cdot\boldsymbol{h} \end{cases} \tag{2.11}$$

where $\boldsymbol{B} = \boldsymbol{A} + \boldsymbol{D}$ and $\boldsymbol{D} = \begin{bmatrix} D_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & D_n \end{bmatrix}$, and this system is a linear system with coefficient matrix $\begin{bmatrix} \boldsymbol{0} & \boldsymbol{B} \\ -\boldsymbol{B} & \boldsymbol{0} \end{bmatrix}$.

For planets in the habitable region close to the stellar binary, the effects of the stellar binary dominates over that of the planetary companions (as illustrated numerically in the next section subsubsection 2.3.3). Therefore, in most of the cases $A \ll D$ and this will thus be assumed (the validity of this assumption will be illustrated in the next section). Under this condition, $\boldsymbol{B}$ is diagonalizable with real eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ satisfying $\lambda_i \approx D_i + A_i$, where A_i is the i^{th} diagonal element of A (proof in section A.1). Denoting the corresponding eigenvectors by $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$, solutions to Eq. (2.11) are in the form

$$\begin{cases} \boldsymbol{h} = \sum\limits_j C_j^{(1)} \sin(\lambda_j t)\, \boldsymbol{v}_j + C_j^{(2)} \cos(\lambda_j t)\, \boldsymbol{v}_j, \\ \boldsymbol{k} = \sum\limits_j -C_j^{(2)} \sin(\lambda_j t)\, \boldsymbol{v}_j + C_j^{(1)} \cos(\lambda_j t)\, \boldsymbol{v}_j. \end{cases} \tag{2.12}$$

Thus, similar to the single planet case, the inclination oscillation frequencies are dom-

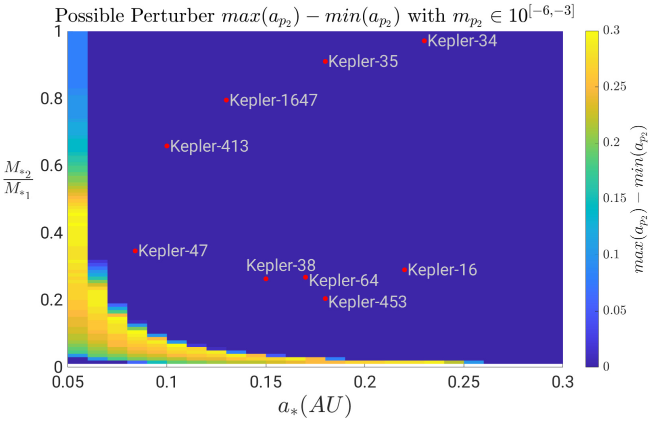


Figure 2.9: Region of the companion planet to allow obliquity variations. The x-axis is the semi-major axis of the orbit of the stellar binary a_*. The y-axis is the mass ratio of two stars $\frac{M_{*2}}{M_{*1}}$. For each choice of these two values as well as p_2's mass log-uniformly sampled from $[10^{-3}, 10^{-6}]\, M_\odot$, we compute the interval of p_2's semi-major values that can potentially place p_1 in spin-orbit resonance, and use color to represent the largest width of this interval (brighter means higher likelihood of larger variation), maximized over all enumerated p_2 mass values. Red dots correspond to observed circumbinary systems.

inated by the stellar potential, and are typically much larger than that of the spin-axis precession frequencies. Therefore, spin-orbit resonances are thus avoided and obliquity variations are typically still low for Earth-like circumbinary planets in the habitable zone like in section 2.3.2, even when planetary companions are present.

Parameter Space Survey Based on the Analytical Theory

To make a robust claim about the obliquity variations of circumbinary multi-planetary systems, we again conduct a systematic study, this time of a binary star system with two planets. We set one of the planets to be an *Earth-like* planet, and vary the mass and location of the other planet. Denote by p_1 the *Earth-like* planet and p_2 the other planet. For dynamically cold systems, we assume that the eccentricity and inclination of the planetary companion are low.

In order to see if the addition of an additional planet can create large obliquity variation, we not only sample the parameter space of the stellar binary over a_* and $\frac{M_{*2}}{M_{*1}}$, but also scan through a wide range of semi-major axes for p_2 (a_{p_2}). For each a_{p_2}, we run over all possible masses of p_2: $m_{p_2} \in [10^{-3}, 10^{-6}]M_{\odot}$. For each setup, we calculate the inclination oscillation modal frequencies based on the secular approach described in section 2.3.3, and record the interval $(\max(a_{p_2}) - \min(a_{p_2}))$ in which the modal frequencies are close to the spin-axis precession frequencies (for an overestimation, we define ‘close’ as $\sim 0.8 - 1.2$ times the spin-axis precession frequency). The interval is considered to be an overestimation of a_{p_2} values that may lead to spin-orbit resonances — for this, we note that the exact proximity of the spin-axis precession frequency that could lead to the spin-orbit resonances depends on the width of the resonances, which is affected by the configuration of the circumbinary system (e.g., obliquity of p_1, the amplitude of the inclination oscillation, as well as the separation of the stellar binary) [12, 20]. In general, the larger the interval length, the higher chance the *Earth-like* planet would experience large *spin-angle* variations.

The results are shown in Fig. 2.9. The likelihood of larger obliquity variation depends on the mass of p_2 and the separation of the stellar binary: when the planets are more massive, they can better compete with the perturbation from the star and allow a higher chance for *spin-angle* variations. However, including maximum planetary mass up to 0.1% that of the star, the planets still need to be very close to each other ($\lesssim 0.1 - 0.3$AU) to allow large *spin-angle* variations.

In short, large *spin-angle* variations may only occur with high probability in the circumstance of a heavy planet ($m_{p_2} > 10^{-3}M_{\odot}$) being very close to the Earth-like planet. We note that separations between planets are typically larger than $\sim 12R_{hill} \sim 0.15 - 1$AU for Earth-like planet with Earth-like to Jupiter-like companions around Sun-like stars, and these separations are larger if the planetary masses are higher. Thus, the yellow region in Fig. 2.9 is unlikely to be physical due to orbital stability of the system.

2.4 Derivation of the Secular Theory

2.4.1 Canonical Variables for Spin-Orbit Dynamics

Building on the canonical Delaunay variables and the Andoyer variables [46, 47, 48, 49], our secular theory of one planet rotating around binary stars not only reduces the dimensions of a system by averaging out the separated fast variables, but also provides insights of the physical behaviors of the system as both sets of variables have clear physical meanings. Specifically, the orbital dynamics are characterized by the Delaunay variables, whereas the spin dynamics are characterized by Andoyer variables. In the following, we will introduce the procedure for properly constructing these variables in a circumbinary system, where two stars are unaffected by the planet and thus modeled as point masses, and the planet is modeled as a moving and rotating rigid body. Firstly, three frames (the *reference frame*, the *body frame* and the *angular momentum frame*) involved in our discussion will be introduced. Secondly, the Delaunay variables will be introduced as a canonical change of coordinates from the spatial positions and momenta of all three bodies' centers of mass. In the end, the orientation and the spin of the rigid planet will be characterized using Andoyer variables.

Three Frames

The *reference frame* is a fixed frame in $\mathbb{R}^3$ with orthogonal basis $(\boldsymbol{E}_1, \boldsymbol{E}_2, \boldsymbol{E}_3)$. Under the reference frame, the spatial position and the translational speed of a body can be expressed by vectors in $\mathbb{R}^3$. In our setup, as the inner orbit's oscillation is relative small compared with the planet's orbit, we make the assumption that the inner orbit is near Keplerian. Thus, we may choose the reference frame (Fig. 2.2) such that $\boldsymbol{E}_1$-axis matches the semi-major axis of the initial inner orbit; $\boldsymbol{E}_2$-axis matches the semi-minor axis of the initial inner orbit; $\boldsymbol{E}_3$-axis matches the normal vector of the initial inner orbit.

On the other hand, the *body frame* (Fig. 2.10) is the moving frame attached to the

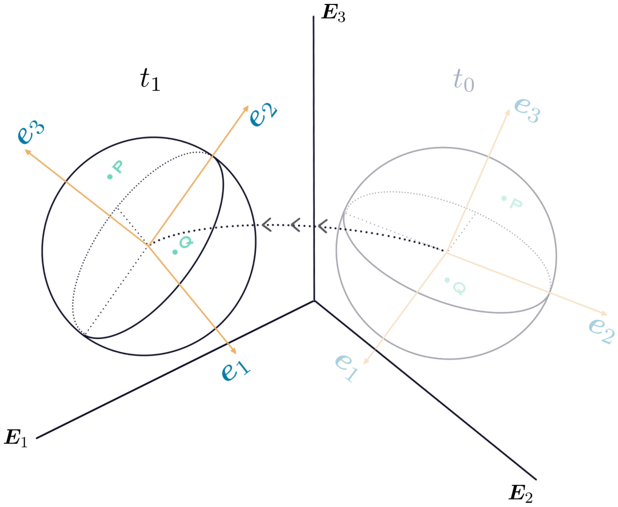


Figure 2.10: The body frame

rotating body (i.e., the planet), giving each particle of the body fixed coordinates. As can be seen from Fig. 2.10, orthogonal bases $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ formed a body frame of this rigid body. As the body moving along the dashed trajectory following arrows as well as self rotating from time t_0 to time t_1, coordinates of points P, Q under the body frame stay the same, despite of the motion of the rigid body. Since we have assumed that each planet is a spheroid, for the body frame (see Fig. 3.1), we fix the body plane (spanned by $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$) the plane of equator and fix the origin the center of the rigid body.

In addition to the *body frame* and the fixed *reference frame* (inertial frame), The *angular momentum frame* is any frame that the rotational angular momentum of the rigid body matches the z-direction.

In Fig. 2.11, three frames of a spinning rigid body are shown together with symbols explained in Table. 2.1. For these three frames, the *reference frame* is fixed; the *body frame* is moving with respect to the orientation of the body; while the *angular momentum frame* is moving with respect to the spinning direction of the body.

Explanations: $\overrightarrow{\boldsymbol{G}_a}$ is the angular momentum vector of the rigid body; α is the angle between the z-axis of the *angular momentum frame* and the z-axis of the *reference frame*;

Table 2.1: x, y, z-axis of three frames

	x-axis	y-axis	z-axis
Reference Frame	$\boldsymbol{E}_1$	$\boldsymbol{E}_2$	$\boldsymbol{E}_3$
Body Frame	$\boldsymbol{e}_1$	$\boldsymbol{e}_2$	$\boldsymbol{e}_3$
Angular Momentum Frame	–	–	$\overrightarrow{\boldsymbol{G}_a}$

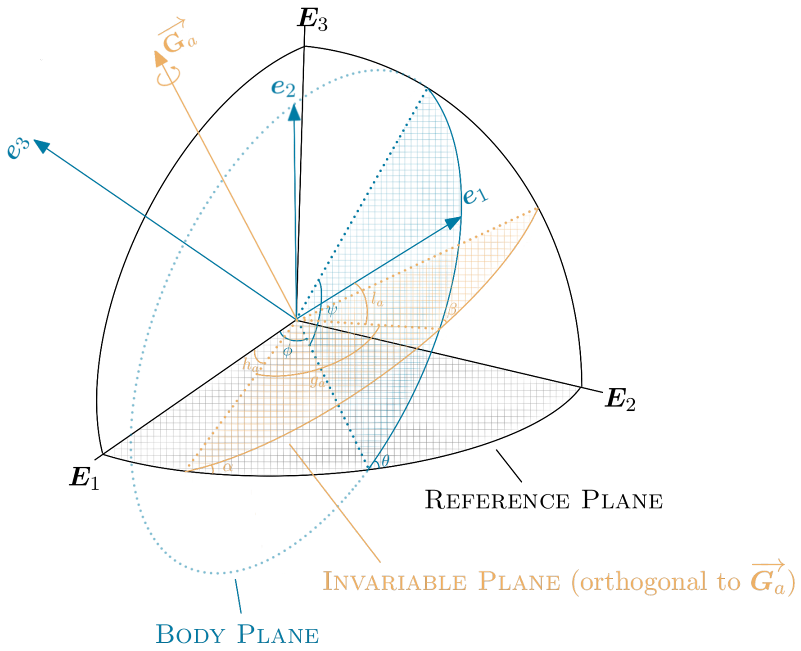


Figure 2.11: Three Frames with Andoyer Variables

β is the angle between the z-axis of the *angular momentum frame* and the z-axis of *body frame*.

Converting Cartesian Spatial Positions and Momenta to the Delaunay Variables

In the system of one planet rotating around binary stars, there are two orbits: the inner orbit of two stars and the outer orbit of the planet rotating around the center of mass of the binary stars. As the inner orbit's oscillation is relative small compared with the planet's orbit, we make the assumption that the inner orbit is near Keplerian, while the outer orbit has its osculating orbital elements being oscillatory (except for the anomaly variable). We denote the orbital elements [50] of the inner orbit and the outer orbit as $(a_*, e_*, i_*, \omega_*, \Omega_*, \nu_*)$, $(a_p, e_p, i_p, \omega_p, \Omega_p, \nu_p)$ respectively. Here, for the inner orbit (relative orbit of the 2nd star around the 1st star), a_* is the length of semi-major axis; e_* is the eccentricity; i_* is the inclination; ω_* is the argument of periapsis; Ω_* is the longitude of ascending node; and we set ν_* to be the true anomaly. For the outer orbit of the planet around the center of mass of two stars, a_p is the length of semi-major axis; e_p is the eccentricity; i_p is the inclination; ω_p is the argument of periapsis; Ω_p is the longitude of ascending node; and ν_p the true anomaly of the planet. Expressing the positions of the three bodies using orbital elements, we have

$$\begin{cases} \boldsymbol{q}_{*_1} = \delta \cdot \dfrac{a_* \left(1 - e_*^2\right)}{1 + e_* \cos\left(\nu_*\right)} \cdot R_x\left(\Omega_*\right) \cdot R_z(i_*) \cdot R_x(\omega_*) \cdot \begin{bmatrix} \cos\nu_* \\ \sin\nu_* \\ 0 \end{bmatrix}, \\ \boldsymbol{q}_{*_2} = -\dfrac{M_{*_1}}{M_{*_2}} \boldsymbol{q}_{*_1}, \\ \boldsymbol{q}_p = \dfrac{a_p \left(1 - e_p^2\right)}{1 + e_p \cos\nu_p} \cdot R_x\left(\Omega_p\right) \cdot R_z(i_p) \cdot R_x(\omega_p) \cdot \begin{bmatrix} \cos\nu_p \\ \sin\nu_p \\ 0 \end{bmatrix}, \end{cases} \tag{2.13}$$

with $\boldsymbol{q}_{*_i}$, $i = 1, 2$ the positions of two stars, $\boldsymbol{q}_p$ the position of the planet and $R_x(\cdot)$, $R_y(\cdot)$, $R_z(\cdot)$ are defined in Eqs. (2.14) to (2.16).

$$R_x(\varphi) := \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{bmatrix}, \tag{2.14}$$

$$R_y(\varphi) := \begin{bmatrix} \cos\varphi & 0 & \sin\varphi \\ 0 & 1 & 0 \\ -\sin\varphi & 0 & \cos\varphi \end{bmatrix}, \tag{2.15}$$

$$R_z(\varphi) := \begin{bmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{2.16}$$

Denote the corresponding Delaunay variables of the inner orbit as

$$\begin{aligned} \mathcal{D}_* &= \{l_*, g_*, h_*, L_*, G_*, H_*\} && (2.17) \\ &= \Big\{M_*, \omega_*, \Omega_*, \mu_*\sqrt{\mathcal{G}(M_{*_1} + M_{*_2})a_*}, L_*\sqrt{1 - e_*^2}, G_* \cos i_*\Big\}, && (2.18) \end{aligned}$$

with $\mu_* = \frac{M_{*_1} M_{*_2}}{M_{*_1} + M_{*_2}}$ and M_* the mean anomaly of the inner orbit. Similarly, the Delaunay variables of the outer orbit is denoted as

$$\begin{aligned} \mathcal{D} &= \{l_d, g_d, h_d, L_d, G_d, H_d\} && (2.19) \\ &= \Big\{M_p, \omega_p, \Omega_p, \mu_p\sqrt{\mathcal{G}\left(M_{*_1} + M_{*_2} + m_p\right)a_p}, L_d\sqrt{1 - e_p^2}, G_d \cos i_p\Big\}, && (2.20) \end{aligned}$$

with $\mu_p = \frac{m_p(M_{*_1}+M_{*_2})}{M_{*_1}+M_{*_2}+m_p}$ and M_p the mean anomaly of the outer orbit. In Eqs. (2.17) and (2.19), $l_*, g_*, h_*, l_p, g_p, h_p$ are angle variables with $L_*, G_*, H_*, L_p, G_p, L_p$ their conjugate momemta. Mapping the orbital elements in Eq. (2.13) to Delaunay variables defined above, the positions can be expressed as functions of Delaunay variables. Noting that different anomalies are used in Eq. (2.13) and Eq. (2.17), Eq. (2.19), one need to solve the Kepler equation to perform the mapping from the orbital elements to the Delaunay variables.

The Orientation and the Rotational Momenta of the Rigid Planet in the Andoyer Variables

As we are interested in the dynamics of the planet's spin, the planet is modeled as a rigid body to account for its finite size, thus the orientation and the rotation of the rigid body are necessary to represent the rigid body. As is well known, each rotation matrix $\boldsymbol{R} \in \mathsf{SO}(3)$ rotates 3-dimensional vectors in Euclidean space in a unique way (we will use the convention of column vectors and left multiplication). Thus we may use a time dependent rotation matrix to represent the transformation from the *body frame* to the *reference frame* in our dynamics (i.e. the orientation of the rigid body). Angles extracted from the Andoyer variables are used to reflect the orientation by a sequence of *standard rotations* (Eqs. (2.14) to (2.16)). Also the rotational angular momentum of the rigid body can be expressed with the momentum variables of the Andoyer variables, conjugating to the three angle variables. Moreover, other than just representing the state of the rigid body, the Andoyer variables also separate the slow and fast scales of the dynamics.

In detail, similar to the Delaunay variables, the Andoyer variables are canonical coordinates. They consist of three angle variables $\{g_a, h_a, l_a\}$, and their conjugate momenta $\{G_a, H_a, L_a\}$. The angle variables are defined by the orientation of three frames (see

Fig. 2.11), while G_a, H_a, L_a are defined as the following [51, 49],

$$\begin{cases} G_a = \left|\overrightarrow{\boldsymbol{G}_a}\right|, \\ H_a = G_a \cos\alpha = \left|\overrightarrow{\boldsymbol{G}_a} \cdot \boldsymbol{E}_3\right|, \\ L_a = G_a \cos\beta = \left|\overrightarrow{\boldsymbol{G}_a} \cdot \boldsymbol{e}_3\right|, \end{cases} \tag{2.21}$$

with α, β are described in Fig. 2.11. In other words, G_a is the magnitude of the angular momentum; H_a is the magnitude of $\overrightarrow{\boldsymbol{G}_a}$'s orthogonal projection to $\boldsymbol{E}_3$; L_a is the magnitude of $\overrightarrow{\boldsymbol{G}_a}$'s orthogonal projection to $\boldsymbol{e}_3$.

Setting the *angular momentum frame* as an intermediate frame, we may represent the rotation matrix $\boldsymbol{R}$ as a product of a sequence of *standard rotations* using angles in Fig. 2.11 [49],

$$\boldsymbol{R} = R_z(h_a) \cdot R_x(\alpha) \cdot R_z(g_a) \cdot R_x(\beta) \cdot R_z(l_a). \tag{2.22}$$

Since $\cos\alpha = \frac{H_a}{G_a}$ and $\cos\beta = \frac{L_a}{G_a}$, replacing α, β in Eq. (2.22), the rotation matrix can be represented purely using the Andoyer variables $\{g_a, h_a, l_a, G_a, H_a, L_a\}$ (name it as $\boldsymbol{R}^{\mathcal{A}}$),

$$\boldsymbol{R}^{\mathcal{A}} = R_z(h_a) \cdot R_x\left(\arccos\left(\frac{H_a}{G_a}\right)\right) \cdot R_z(g_a) \cdot R_x\left(\arccos\left(\frac{L_a}{G_a}\right)\right) \cdot R_z(l_a). \tag{2.23}$$

2.4.2 Dynamics of the Spin for an Earth-like Planet in Circumbinary Systems: two point mass stars + one rigid body planet

Hamiltonian Formulation

In this section, we will express the Hamiltonian using the canonical Delaunay and Andoyer variables $\mathcal{D}, \mathcal{D}_*, \mathcal{A}$ introduced in section 2.4.1. In short, $H = T_{*_1} + T_{*_2} + T_p + V_* + V_1 + V_2$ with T_{*_i} kinetic energies of the ith star, T_p the kinetic energy of the planet, V_* potential energy between two stars and V_i potential energies between ith star and the planet. We assume that the inner orbit is near circular and near Keplerian (as the perturbation of the planet is negligible) such that during the dynamics $\mathcal{O}(i_*) = \mathcal{O}(e_*) = \mathcal{O}\left((a_*/a_p)^3\right)$. Then $T_{*_1} + T_{*_2} + V_* = -\frac{\mathcal{G}^2(M_{*_1}+M_{*_2})^2\mu_*^3}{2L_*^2} + \mathcal{O}\left((a_*/a_p)^3\right)$. We will calculate T_p, V_1 and V_2 terms and conclude the Hamiltonian in action angle variables in the following under the assumption of $\mathcal{O}(e_p) = \mathcal{O}(\beta) = \mathcal{O}\left((a_*/a_p)^3\right)$. Note that all $\mathcal{O}\left((a_*/a_p)^3\right)$ terms are dropped to the remainder term $\tilde{\mathfrak{R}}$ (see Eq. (2.50)) and neglected when analyzing the dynamics, the effective Hamiltonian we consider in fact only relies on variables $\mathcal{A}, \mathcal{D}, L_*, l_*$ and the inner orbit we consider is in fact a Keplerian orbit due to the dropped perturbation. In the end, after dropping $\mathcal{O}\left((a_*/a_p)^3\right)$ terms, an approximated Hamiltonian will be obtained with 6 pairs of conjugate variables corresponding to separated slow and fast scales.

The total kinetic energy of the planet is the sum of its linear kinetic energy $T^{linear}(\mathcal{D})$ and its rotational kinetic energy $T^{rot}(\mathcal{A}, L_*, l_*)$. The former is

$$T^{linear} = \frac{\mathcal{G}^2(M_{*_1} + M_{*_2} + m_p)^2\mu_p^3}{2L_d^2}. \tag{2.24}$$

For the latter, by the definition of the Andoyer variables, we may express the rotational kinetic energy as

$$T^{rot}(\mathcal{A}) = \frac{G_a^2 \sin^2\beta}{2I_p^{(1)}} + \frac{G_a^2 \cos^2\beta}{2I_p^{(3)}} = \frac{H_a^2}{2I_p^{(1)}} + \frac{G_a^2 - H_a^2}{2I_p^{(3)}}. \tag{2.25}$$

with

$$\boldsymbol{I}_p = \begin{bmatrix} I_p^{(1)} & 0 & 0 \\ 0 & I_p^{(2)} & 0 \\ 0 & 0 & I_p^{(3)} \end{bmatrix} \tag{2.26}$$

the (standard) moment of inertia tensor of the planet.

The potential energy of the planet is influenced by both stars,

$$V(\mathcal{A}, \mathcal{D}, l_*, L_*) = V_1 + V_2. \tag{2.27}$$

Without loss of generality, we will only derive V_1 here. Integrating the potential for each mass point of the rigid planet,

$$\begin{aligned}
V_1 =& \int_{\mathcal{B}} -\frac{\mathcal{G}M_{*_1}\rho(\boldsymbol{x})}{\|(\boldsymbol{q}_p + \boldsymbol{R}_p\boldsymbol{x}) - \boldsymbol{q}_{*_1}\|}\, d\boldsymbol{x} \\
=& \mathcal{G}M_{*_1}m_p\left\{ -\frac{1}{\|\boldsymbol{q}_p - \boldsymbol{q}_{*_1}\|} - \frac{tr[\boldsymbol{I}_p]}{m_p\,\|\boldsymbol{q}_p - \boldsymbol{q}_{*_1}\|^3} + \frac{3(\boldsymbol{q}_p - \boldsymbol{q}_{*_1})^T \boldsymbol{R}_p \boldsymbol{I}_p {\boldsymbol{R}_p}^T (\boldsymbol{q}_p - \boldsymbol{q}_{*_1})}{2m_p\,\|\boldsymbol{q}_p - \boldsymbol{q}_{*_1}\|^5} \right\} \\
&+ \mathcal{O}\left(\left(\frac{\mathcal{R}}{\|\boldsymbol{q}_p - \boldsymbol{q}_{*_1}\|}\right)^3\right),
\end{aligned} \tag{2.28}$$

with $\boldsymbol{q}_p, \boldsymbol{q}_{*_i}$ in the Delaunay variables $\mathcal{D}, \mathcal{D}_*$ and $\boldsymbol{R}_p$ the rotation matrix in the Andoyer variables $\mathcal{A}$ (Eq. (2.23)). Term by term, we will show how we express Eq. (2.28) in the Delaunay variables and the Andoyer variables with higher order terms separated.

In detail, to analyze the term in Eq. (2.28) with $\boldsymbol{R}_p$, we assume the angle β (the angle between the spin axis and the $\boldsymbol{e}_3$-axis of the planet) small as physically, the oblate shape of the planet is flattened by the spinning of the planet. Thus we may rewrite $\boldsymbol{R}_p\boldsymbol{I}_p\boldsymbol{R}_p^T$ in

Eq. (2.28) by separating the higher order terms of $\mathcal{O}\left(\beta\right)$,

$$\begin{aligned}\boldsymbol{R}_p\boldsymbol{I}_p\boldsymbol{R}_p^T =& R_z(h_a)\cdot R_x(\alpha)\cdot R_z(g_a)\cdot R_x(\beta)\cdot R_z(l_a)\cdot \boldsymbol{I}_p \\ &\cdot R_z^T(l_a)\cdot R_x^T(\beta)\cdot R_z^T(g_a)\cdot R_x^T(\alpha)\cdot R_z^T(h_a) \\ =& R_z(h_a)\cdot R_x(\alpha)\cdot R_z(g_a)\cdot R_x(\beta)\cdot \boldsymbol{I}_p\cdot R_x^T(\beta)\cdot R_z^T(g_a)\cdot R_x^T(\alpha)\cdot R_z^T(h_a) \\ =& R_z(h_a)\cdot R_x(\alpha)\cdot \boldsymbol{I}_p\cdot R_x^T(\alpha)\cdot R_z^T(h_a)+\mathcal{O}(\beta) \\ =& I_p^{(1)}\boldsymbol{I}_{3\times 3} \\ &+(I_p^{(3)}-I_p^{(1)})\cdot R_z(h_a)\cdot R_x(\alpha)\cdot \begin{bmatrix}0&0&0\\0&0&0\\0&0&1\end{bmatrix}\cdot R_x^T(\alpha)\cdot R_z^T(h_a) \\ &+\mathcal{O}(\beta).\end{aligned} \tag{2.29}$$

For $\|\boldsymbol{q}_p-\boldsymbol{q}_{*_1}\|$, using Eq. (2.13), we have

$$\|\boldsymbol{q}_{*_1}-\boldsymbol{q}_p\|^2=\|\boldsymbol{q}_p\|^2\cdot\left(1+2\cdot\mathfrak{P}(\nu_*,\nu_p,i_p,\omega_p,\Omega_p)\cdot\frac{\|\boldsymbol{q}_{*_1}\|}{\|\boldsymbol{q}_p\|}+\left(\frac{\|\boldsymbol{q}_{*_1}\|}{\|\boldsymbol{q}_p\|}\right)^2\right), \tag{2.30}$$

with

$$\begin{aligned}&\mathfrak{P}(\nu_*,\nu_p,i_p,\omega_p,\Omega_p) \\ =&\sin\left(\nu_p\right)\cos\left(\nu_*\right)\big[\cos\left(i_p\right)\cos\left(\omega_p\right)\sin\left(\Omega_p\right)+\sin\left(\omega_p\right)\cos\left(\Omega_p\right)\big] \\ &-\sin\left(\nu_*\right)\big[\cos\left(i_p\right)\cos\left(\Omega_p\right)\cdot\sin\left(\nu_p+\omega_p\right)+\sin\left(\Omega_p\right)\cos\left(\nu_p+\omega_p\right)\big] \\ &+\cos\left(\nu_p\right)\cos\left(\nu_*\right)\big[\cos\left(i_p\right)\sin\left(\omega_p\right)\sin\left(\Omega_p\right)-\cos\left(\omega_p\right)\cos\left(\Omega_p\right)\big],\end{aligned} \tag{2.31}$$

and

$$\frac{\|\boldsymbol{q}_{*_1}\|}{\|\boldsymbol{q}_p\|}=\delta\frac{a_*(1+\mathcal{O}(e_*))}{a_p(1+\mathcal{O}(e_p))}=\delta\frac{a_*}{a_p}+\mathcal{O}(e_*)+\mathcal{O}(e_p). \tag{2.32}$$

Applying Legendre polynomial for the expansion of $\frac{1}{\|\boldsymbol{q}_{*_1}-\boldsymbol{q}_p\|}$ with $\|\boldsymbol{q}_{*_1}-\boldsymbol{q}_p\|$ calculated in Eq. (2.30) , we have

$$\begin{aligned}\frac{1}{\|\boldsymbol{q}_{*_1}-\boldsymbol{q}_p\|} =&\frac{1}{\|\boldsymbol{q}_p\|}\left(1-\frac{\|\boldsymbol{q}_{*_1}\|}{\|\boldsymbol{q}_p\|}\mathfrak{P}+\left(\frac{\|\boldsymbol{q}_{*_1}\|}{\|\boldsymbol{q}_p\|}\right)^2\left(\frac{3}{2}\mathfrak{P}^2-\frac{1}{2}\right)+\mathcal{O}\left(\left(\frac{\|\boldsymbol{q}_{*_1}\|}{\|\boldsymbol{q}_p\|}\right)^3\right)\right)\\ =&\frac{1}{a_p}\left(1-\delta\left(\frac{a_*}{a_p}\right)\mathfrak{P}+\delta^2\left(\frac{a_*}{a_p}\right)^2\left(\frac{3}{2}\mathfrak{P}^2-\frac{1}{2}\right)+\mathcal{O}\left(\left(\frac{a_*}{a_p}\right)^3\right)\right)\\ &+\mathcal{O}(e_*)+\mathcal{O}(e_p).\end{aligned} \tag{2.33}$$

Denote

$$\mathfrak{D}_1=\frac{1}{a_p}\left[-\delta\mathfrak{P}+\delta^2\left(\frac{a_*}{a_p}\right)\left(\frac{3}{2}\mathfrak{P}^2-\frac{1}{2}\right)\right], \tag{2.34}$$

We can rewrite Eq. (2.33) as

$$\frac{1}{\|\boldsymbol{q}_{*_1}-\boldsymbol{q}_p\|}=\frac{1}{a_p}+\left(\frac{a_*}{a_p}\right)\mathfrak{D}_1+\mathcal{O}\left(\left(\frac{a_*}{a_p}\right)^3\right)+\mathcal{O}(e_*)+\mathcal{O}(e_p). \tag{2.35}$$

Similarly, we have

$$\frac{1}{\|\boldsymbol{q}_{*_1}-\boldsymbol{q}_p\|^3}=\frac{1}{a_p^3}+\left(\frac{a_*}{a_p}\right)\mathfrak{D}_3+\mathcal{O}\left(\left(\frac{a_*}{a_p}\right)^3\right)+\mathcal{O}(e_*)+\mathcal{O}(e_p). \tag{2.36}$$

and

$$\frac{1}{\|\boldsymbol{q}_{*_1}-\boldsymbol{q}_p\|^5}=\frac{1}{a_p^5}+\left(\frac{a_*}{a_p}\right)\mathfrak{D}_5+\mathcal{O}\left(\left(\frac{a_*}{a_p}\right)^3\right)+\mathcal{O}(e_*)+\mathcal{O}(e_p). \tag{2.37}$$

Let

$$\mathfrak{Q}_{*_1} = \delta \cdot a_* \cdot \begin{bmatrix} \cos \nu_* \\ \sin \nu_* \\ 0 \end{bmatrix}, \tag{2.38}$$

and

$$\mathfrak{Q}_p = a_p \cdot R_x\left(\Omega_p\right) \cdot R_z(i_p) \cdot R_x(\omega_p) \cdot \begin{bmatrix} \cos \nu_p \\ \sin \nu_p \\ 0 \end{bmatrix}. \tag{2.39}$$

According to Eq. (2.13), we have

$$\boldsymbol{q}_{*_1} = \mathfrak{Q}_{*_1} + \mathcal{O}(e_*), \qquad \boldsymbol{q}_p = \mathfrak{Q}_p + \mathcal{O}(e_p). \tag{2.40}$$

Plugging Eqs. (2.29), (2.35) to (2.37) and (2.40), into Eq. (2.28) and absorbing $\mathcal{O}\left(\left(\mathcal{R}/a_p\right)^3\right)$ into the $\mathcal{O}\left(\left(a_*/a_p\right)^3\right)$ term as $\mathcal{R} \ll a_*$, we have

$$\begin{aligned}
V_1 =& \mathcal{G} M_{*_1} m_p \Bigg\{ -\frac{1}{a_p} - \frac{a_*}{a_p}\mathfrak{D}_1 - \frac{tr[\boldsymbol{I}_p] - \frac{3}{2} I_p^{(1)}}{m_p} \left(\frac{1}{a_p^3} + \frac{a_*}{a_p}\mathfrak{D}_3 \right) + \frac{I_p^{(3)} - I_p^{(1)}}{m_p} \cdot \left(\frac{1}{a_p^5} + \frac{a_*}{a_p}\mathfrak{D}_5 \right) \\
& \cdot \left(\mathfrak{Q}_p - \mathfrak{Q}_{*_1}\right)^T \cdot R_z(h_a) \cdot R_x(\alpha) \cdot \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot R_x^T(\alpha) \cdot R_z^T(h_a) \left(\mathfrak{Q}_p - \mathfrak{Q}_{*_1}\right) \Bigg\} \\
& + \mathcal{O}(\beta) + \mathcal{O}(e_*) + \mathcal{O}(e_p) + \mathcal{O}\left(\left(\frac{a_*}{a_p} \right)^3 \right) \\
=& -\frac{\mathcal{G} M_{*_1} m_p}{a_p} + \frac{a_*}{a_p}\mathfrak{R}_1 + \mathcal{O}(\beta) + \mathcal{O}(e_*) + \mathcal{O}(e_p) + \mathcal{O}\left(\left(\frac{a_*}{a_p} \right)^3 \right) \\
=& -\frac{\mathcal{G} M_{*_1} m_p}{a_p} + \frac{a_*}{a_p}\mathfrak{R}_1 + \mathcal{O}\left(\left(\frac{a_*}{a_p} \right)^3 \right)
\end{aligned} \tag{2.41}$$

where

$$\begin{aligned}\mathfrak{R}_1 =& \mathcal{G} M_{*_1} m_p \Bigg\{ -\mathfrak{D}_1 - \frac{tr[\boldsymbol{I}_p] - \frac{3}{2} I_p^{(1)}}{m_p} \left(\frac{1}{a_p^2 a_*} + \mathfrak{D}_3 \right) \\ &+ \frac{I_p^{(3)} - I_p^{(1)}}{m_p} \cdot \left(\frac{1}{a_p^4 a_*} + \mathfrak{D}_5 \right) \cdot \left(\mathfrak{Q}_p - \mathfrak{Q}_{*_1} \right)^T \\ &\cdot R_z(h_a) \cdot R_x(\alpha) \cdot \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot R_z^T(h_a) \cdot \left(\mathfrak{Q}_p - \mathfrak{Q}_{*_1} \right) \Bigg\}.\end{aligned} \tag{2.42}$$

Similarly, $V_2 = -\frac{\mathcal{G} M_{*_2} m_p}{a_p} + \frac{a_*}{a_p} \mathfrak{R}_2 + \mathcal{O}\left(\left(\frac{a_*}{a_p} \right)^3 \right)$.

Changing coordinates from $(a_p, e_p, i_p, \omega_p, \Omega_p, \nu_p)$ to $\mathcal{D}$, we have $V = V_1 + V_2$ as a function of $\mathcal{A}, \mathcal{D}$, l_* and L_*. Note that ν_p is the true anomaly and M_p (in $\mathcal{D}$) is the mean anomaly, and we expressed M_p as $\nu_p + \mathcal{O}(e_p)$ instead of solving the Kepler equation exactly.

Summing the kinetic energy and the potential energy, the full Hamiltonian is:

$$\begin{aligned} H\left(\mathcal{A}, \mathcal{D}, l_*, L_*\right) &= T^{linear}\left(\mathcal{D}\right) + T^{rot}\left(\mathcal{A}\right) + V\left(\mathcal{A}, \mathcal{D}, l_*, L_*\right) \\ &= -\frac{\mathcal{G}^2 (M_{*_1} + M_{*_2})^2 \mu_*^3}{2 L_*^2} + \frac{\mathcal{G}^2 (M_{*_1} + M_{*_2} + m_p)^2 \mu_p^3}{2 L_d^2} + T^{rot}\left(\mathcal{A}\right) \\ &+ V\left(\mathcal{A}, \mathcal{D}, l_*, L_*\right). \end{aligned} \tag{2.43}$$

Plugging in the potential V_1 and V_2 to Eq. (2.43), one sees that the Hamiltonian only depends on $\mathcal{S} = \{l_d, h_d, g_d, L_d, H_d, G_d, l_d, g_a, h_a, G_a, H_a, L_a, l_*, L_*\}$,

$$\begin{aligned} H\left(\mathcal{S}\right) =& -\frac{\mathcal{G}^2(M_{*_1}+M_{*_2})^2\mu_*^3}{2L_*^2} + \frac{\mathcal{G}^2(M_{*_1}+M_{*_2}+m_p)^2\mu_p^3}{2L_d^2} + T^{rot}(\mathcal{S}) + V_1 + V_2, \\ =& -\frac{\mathcal{G}^2(M_{*_1}+M_{*_2})^2\mu_*^3}{2L_*^2} + \frac{\mathcal{G}^2(M_{*_1}+M_{*_2}+m_p)^2\mu_p^3}{2L_d^2} + \frac{H_a^2}{2I_p^{(1)}} + \frac{G_a^2 - H_a^2}{2I_p^{(3)}} \\ & - \frac{\mathcal{G}^2(M_{*_1}+M_{*_2}+m_p)^2\mu_p^3}{L_d^2} + \frac{a_*(L_*)}{a_p(L_d)}\left(\mathfrak{R}_1 + \mathfrak{R}_2\right) + \widetilde{\mathfrak{R}}, \\ =& -\frac{\mathcal{G}^2(M_{*_1}+M_{*_2})^2\mu_*^3}{2L_*^2} - \frac{\mathcal{G}^2(M_{*_1}+M_{*_2}+m_p)^2\mu_p^3}{2L_d^2} + \frac{H_a^2}{2I_p^{(1)}} + \frac{G_a^2 - H_a^2}{2I_p^{(3)}} \\ & + \frac{a_*(L_*)}{a_p(L_d)}\left(\mathfrak{R}_1 + \mathfrak{R}_2\right) + \widetilde{\mathfrak{R}}, \end{aligned} \tag{2.44}$$

with $\widetilde{\mathfrak{R}} = \mathcal{O}\left(\left(\frac{a_*}{a_p}\right)^3\right)$.

Using Averaging Theory to Construct an Approximated Dynamics

The dynamics of the above Hamiltonian system (Eq. (2.44)) consists of the slow components $\mathcal{S}_{slow} = \{h_d, g_d, L_d, H_d, G_d, h_a, G_a, H_a, L_*\}$, as well as the fast components $\mathcal{S}_{fast} = \{l_*, l_d, g_a\}$, which correspond to binary's orbital phase, planet's orbital phase, and planet spin. The fast and slow components are separated by scaling η_1 and η_2, where $\eta_1 = \frac{a_*}{a_p}$ is a small parameter corresponding to the stellar binary are closer to each other than the planet, and $\eta_2 = \frac{I_p^{(1)} - I_p^{(3)}}{I_p^{(3)}}$ is another small parameter modeling the oblateness of the planet. In the observed circumbinary systems, $0.084 \leq \eta_1 \leq 0.23$ [30]; and $\eta_2 \approx 0.00334$ for Earth-like planets. Let us now compare the relative timescales of the variables, in order to determine and justify the order of a sequence of averaging approximations. Specifically, the slow variables change with rate either $\mathcal{O}(\eta_1)$ or $\mathcal{O}(\eta_2)$. On the other hand, $\dot{l}_* = \mathcal{O}\left(\eta_1^{-3/2}\right)$, $\dot{l}_d = \mathcal{O}(1)$ and $\dot{g}_a$ equals the spin rate of the planet around the direction of the angular momentum. Here, we assume that the frequencies of l_*, l_d and g_a are not commensurable. Thus, we may average over l_* and l_d in sequence, assuming no resonance occurs between

l_*, l_d, g_a, then we saw that g_a is decoupled from the system.

Average over binary's orbital phase l_*

A closer look at $H\left(\mathcal{S}\right)$ reveals that all terms dependent on l_* are of order $\mathcal{O}(\eta_1^2)$. Thus, $H\left(\mathcal{S}\right)$ can be written as

$$\begin{aligned} & H\left(\mathcal{S}, l_*, L_*\right) \\ =& H_0\left(L_*\right)+\eta_1 H_1\left(\mathcal{S} \backslash \left\{l_*, L_*\right\}\right)+\eta_1^2 H_2\left(\mathcal{S}\right)+\eta_1^3 H_3\left(\mathcal{S}\right) \\ & +\mathcal{O}\left(\eta_1^4\right). \end{aligned} \tag{2.45}$$

As $H(\mathcal{S})$ is nearly-integrable, canonical averaging theory (e.g., [50]) guarantees that averaging $H(\mathcal{S}, l_*, L_*)$ over l_* is equivalent to first-order averaging of the Hamiltonian dynamics. After l_* is averaged, L_* becomes a conserved quantity. Therefore, excluding $H_0(L_*)$, we obtain the averaged Hamiltonian

$$\begin{aligned} \overline{H}\left(\mathcal{S}\right) &= H_1\left(\mathcal{S} \backslash \left\{l_*, L_*\right\}\right)+\frac{1}{2\pi}\int_0^{2\pi} \frac{\eta_1 H_2\left(\mathcal{S}\right)+\eta_1^2 H_3\left(\mathcal{S}\right)}{\dot{l}_*}\, dl_* \\ &= H_1\left(\mathcal{S} \backslash \left\{l_*, L_*\right\}\right)+\eta_1 \cdot 0+\eta_1^2 \overline{H}_3\left(\mathcal{S} \backslash \left\{l_*, L_*\right\}\right) \\ &= H_1\left(\mathcal{S} \backslash \left\{l_*, L_*\right\}\right)+\eta_1^2 \overline{H}_3\left(\mathcal{S} \backslash \left\{l_*, L_*\right\}\right) \end{aligned}. \tag{2.46}$$

which generates an approximated dynamics (see Eq. (2.47)).

$$
\begin{cases}
\dot{l}_d = \dfrac{\mathcal{G}^2(M_{*_1}+M_{*_2}+m_p)^2\mu_p^3}{L_d^3} \\
\qquad - \eta_1^2 \dfrac{(\delta-1)\delta\mathcal{G}^2(M_{*_1}+M_{*_2}+m_p)^2\mu_p^3\left(G_d^2\left(3\cos\left(2\left(g_d+l_d\right)\right)-1\right)+6H_d^2\sin^2\left(g_d+l_d\right)\right)}{4G_d^2L_d^3}, \\
\dot{g}_d = -\dfrac{3(\delta-1)\delta H_d^2\mathcal{G}^2(M_{*_1}+M_{*_2}+m_p)^2\mu_p^3\eta_1^2\sin^2\left(g_d+l_d\right)}{2G_d^3L_d^2}, \\
\dot{h}_d = \dfrac{3(\delta-1)\delta H_d\mathcal{G}^2(M_{*_1}+M_{*_2}+m_p)^2\mu_p^3\eta_1^2\sin^2\left(g_d+l_d\right)}{2G_d^2L_d^2}, \\
\dot{L}_d = \dfrac{3(\delta-1)\delta\mathcal{G}^2(M_{*_1}+M_{*_2}+m_p)^2\mu_p^3\eta_1^2\left(G_d^2-H_d^2\right)\sin\left(2\left(g_d+l_d\right)\right)}{4G_d^2L_d^2}, \\
\dot{G}_d = \dfrac{3(\delta-1)\delta\mathcal{G}^2(M_{*_1}+M_{*_2}+m_p)^2\mu_p^3\eta_1^2\left(G_d^2-H_d^2\right)\sin\left(2\left(g_d+l_d\right)\right)}{4G_d^2L_d^2}, \\
\dot{H}_d = 0, \\
\dot{g}_a = \dfrac{G_a}{I_p^{(3)}} + \eta_2\left(f_0^{(1)}\left(\mathcal{S}\right)+\eta_1^2\cdot f_1^{(1)}\left(\mathcal{S}\right)\right), \\
\dot{h}_a = \eta_2\cdot\left(f_0^{(2)}\left(\mathcal{S}\right)+\eta_1^2\cdot f_1^{(2)}\left(\mathcal{S}\right)\right), \\
\dot{G}_a = 0, \\
\dot{H}_a = \eta_2\cdot\left(f_0^{(3)}\left(\mathcal{S}\right)+\eta_1^2\cdot f_1^{(3)}\left(\mathcal{S}\right)\right).
\end{cases}
\tag{2.47}
$$

From the above averaged Hamiltonian Eq. (2.46), we may derived the corresponding dynamics Eq. (2.47). Then, we average the vector field over the fast angle l_d. Important to note is, $\dot{l}_d$ actually depends on l_d, and therefore one cannot just average $\dot{L}_d$ (for example) uniformly over l_d from 0 to 2π. A proper ergodic averaging can be done either via integration against time, or, essentially equivalently, a weighted average that reflects the non-uniform ergodic measure of l_d on the torus (which can be rigorously derived using normal form).

For example, for L_d, the averaged dynamics is

$$\dot{\bar{L}}_d = \frac{\frac{1}{2\pi}\int_0^{2\pi} \frac{\dot{L}_d}{\dot{l}_d}\, dt}{\frac{1}{2\pi}\int_0^{2\pi} \frac{1}{\dot{l}_d}\, dt}. \tag{2.48}$$

Performing the same operation for $G_d, H_d, g_d, h_d, G_a, H_a, g_a, h_a$, we obtain the averaged dynamics Eq. (2.49).

Resulting Approximated Dynamics

After averaging over l_* and l_d, another fast angle g_a is decoupled from the system (see Eq. (2.49)).

$$
\begin{cases}
\dot{\overline{L}}_d = \dot{\overline{G}}_d = \dot{\overline{H}}_d = \dot{\overline{G}}_a = 0, \\
\dot{\overline{g}}_d = \dfrac{3(\delta-1)\delta H_d^2 \mathcal{G}^2 (M_{*_1}+M_{*_2}+m_p)^2 \mu_p^3 \eta_1^2}{L_d^2 \left(G_d^3\left((\delta-1)\delta\eta_1^2-4\right) - 3(\delta-1)\delta G_d H_d^2 \eta_1^2\right)}, \\
\dot{\overline{h}}_d = -\dfrac{3(\delta-1)\delta H_d \mathcal{G}^2 (M_{*_1}+M_{*_2}+m_p)^2 \mu_p^3 \eta_1^2}{L_d^2 \left(G_d^2\left((\delta-1)\delta\eta_1^2-4\right) - 3(\delta-1)\delta H_d^2 \eta_1^2\right)}, \\
\dot{\overline{h}}_a = \dfrac{3\eta_2 \mathcal{G}^4 I_p^{(3)} \mu_p^7 (m_p+M_{*_1}+M_{*_2})^4}{8 m_p G_d^2 L_d^6 L_a^3 \sqrt{1-\frac{H_a^2}{L_a^2}} \left(G_d^2\left((\delta-1)\delta\eta_1^2-4\right) - 3(\delta-1)\delta H_d^2 \eta_1^2\right)} \\
\quad \left(G_d^4 H_a L_a \sqrt{1-\dfrac{H_a^2}{L_a^2}} \left(-85(\delta-1)\delta\eta_1^2 + \left(223(\delta-1)\delta\eta_1^2-16\right)\cos\left(2\left(h_d-h_a\right)\right) - 16\right) \right. \\
\quad + 2G_d^3 H_d \left(125(\delta-1)\delta\eta_1^2-16\right)\sqrt{1-\dfrac{H_d^2}{G_d^2}}\cos\left(h_d-h_a\right)\left(2H_a^2-L_a^2\right) \\
\quad + 2G_d^2 H_d^2 H_a L_a \sqrt{1-\dfrac{H_a^2}{L_a^2}} \left(-95(\delta-1)\delta\eta_1^2 + \left(8-68(\delta-1)\delta\eta_1^2\right)\cos\left(2\left(h_d-h_a\right)\right) + 24\right) \\
\quad + 174(\delta-1)\delta G_d H_d^3 \eta_1^2 \sqrt{1-\dfrac{H_d^2}{G_d^2}}\cos\left(h_d-h_a\right)\left(2H_a^2-L_a^2\right) \\
\quad \left. - 87(\delta-1)\delta H_d^4 H_a L_a \eta_1^2 \left(\cos\left(2\left(h_d-h_a\right)\right)+3\right)\sqrt{1-\dfrac{H_a^2}{L_a^2}} \right) \\
\dot{\overline{H}}_a = -\dfrac{3\eta_2 \mathcal{G}^4 I_p^{(3)} \mu_p^7 \sin\left(h_d-h_a\right) (m_p+M_{*_1}+M_{*_2})^4}{4 m_p G_d^2 L_d^6 L_a^2 \left(G_d^2\left((\delta-1)\delta\eta_1^2-4\right) - 3(\delta-1)\delta H_d^2 \eta_1^2\right)} \\
\quad \left(G_d^4 \left(223(\delta-1)\delta\eta_1^2-16\right)\cos\left(h_d-h_a\right)\left(H_a^2-L_a^2\right) \right. \\
\quad - 8G_d^2 H_d^2 \left(17(\delta-1)\delta\eta_1^2-2\right)\cos\left(h_d-h_a\right)\left(H_a^2-L_a^2\right) \\
\quad + G_d^3 H_d H_a L_a \left(16-125(\delta-1)\delta\eta_1^2\right)\sqrt{1-\dfrac{H_d^2}{G_d^2}}\sqrt{1-\dfrac{H_a^2}{L_a^2}} \\
\quad - 87(\delta-1)\delta G_d H_d^3 H_a L_a \eta_1^2 \sqrt{1-\dfrac{H_d^2}{G_d^2}}\sqrt{1-\dfrac{H_a^2}{L_a^2}} \\
\quad \left. + 87(\delta-1)\delta H_d^4 \eta_1^2 \cos\left(h_d-h_a\right)\left(L_a^2-H_a^2\right) \right)
\end{cases}
\tag{2.49}
$$

and $\overline{g}_a$ decoupled from the system. Changing coordinates to $X = H_a/G_a$ and $h = h_d - h_a$, from Eq. (2.49), we have

$$\begin{cases} \dot{X} = \sin(h)\left[C_1 X\sqrt{1-X^2} + 4C_2\cos(h)(1-X^2)\right], \\ \dot{h} = \dfrac{C_1\cos(h)\left(1-2X^2\right)}{\sqrt{1-X^2}} - 2C_2 X\cos(2h) + C_3 + 2C_4 X, \end{cases} \tag{2.50}$$

with

$$\begin{cases} C_1 = & \dfrac{3\eta_2\mathcal{G}^4 I_p^{(3)}\mu_p^7\sqrt{1-\frac{H_d^2}{G_d^2}}(m_p+M_{*_1}+M_{*_2})^4}{4m_pG_dL_d^6L_a\left(G_d^2\left((\delta-1)\delta\eta_1^2-4\right)-3(\delta-1)\delta H_d^2\eta_1^2\right)} \\ & \Big(G_d^2H_d\,\big(125(\delta-1)\delta\eta_1^2-16\big)+87(\delta-1)\delta H_d^3\eta_1^2\Big) \\ C_2 = & \dfrac{3\eta_2\mathcal{G}^4 I_p^{(3)}\mu_p^7\left(G_d^2-H_d^2\right)(m_p+M_{*_1}+M_{*_2})^4}{16m_pG_d^2L_d^6L_a\left(G_d^2\left((\delta-1)\delta\eta_1^2-4\right)-3(\delta-1)\delta H_d^2\eta_1^2\right)}, \\ & \Big(G_d^2\,\big(223(\delta-1)\delta\eta_1^2-16\big)+87(\delta-1)\delta H_d^2\eta_1^2\Big) \\ C_3 = & -\dfrac{3(\delta-1)\delta\mathcal{G}^2H_d\mu_p^3\eta_1^2(m_p+M_{*_1}+M_{*_2})^2}{L_d^2\left(G_d^2\left((\delta-1)\delta\eta_1^2-4\right)-3(\delta-1)\delta H_d^2\eta_1^2\right)}, \\ C_4 = & \dfrac{3\eta_2\mathcal{G}^4 I_p^{(3)}\mu_p^7(m_p+M_{*_1}+M_{*_2})^4}{16m_pG_d^2L_d^6L_a\left(G_d^2\left((\delta-1)\delta\eta_1^2-4\right)-3(\delta-1)\delta H_d^2\eta_1^2\right)} \\ & \Big(G_d^4\,\big(85(\delta-1)\delta\eta_1^2+16\big)+2G_d^2H_d^2\,\big(95(\delta-1)\delta\eta_1^2-24\big)+261(\delta-1)\delta H_d^4\eta_1^2\Big) \end{cases} \tag{2.51}$$

By definition, we know that H_a is the angular momentum's orthogonal projection to $\boldsymbol{E}_3$ axis and $h_d - h_a$ is the phase difference between the two precessions (the precession of the planet's orbit and the precession of the angular momentum). Thus X, h together give the approximated dynamics of the spin.

The Hamiltonian corresponding to the dynamics above is Eq. (2.50)

$$\begin{aligned} H(X,h) =& C_1\sqrt{1-X^2}X\cos(h) + C_2\left(1-X^2\right)\cos(2h) \\ &+ C_3X + C_4X^2. \end{aligned} \tag{2.52}$$

The averaged dynamics has an $\mathcal{O}\left(\max(\eta_1, \sqrt{\eta_2})\right)$ error at least till time $\mathcal{O}\left(\frac{1}{\max(\eta_1,\sqrt{\eta_2})}\right)$, and numerically, we observed that the approximation remains accurate over an even longer time span.

2.5 Conclusion and Discussions

Stellar binaries are common in the Solar neighborhood, and thus to understand the habitability of their planets, if any, is as important as that for single star systems. We adopt a dynamics approach, in which we show that the planetary spin evolution can be very different when a second host star is present. This is because torques from both stars act on the spin-axis of its planets. This fact, together with more significant deviations from Keplerian orbits, makes it difficult to directly apply existing theory for obliquity designed for single star systems. We investigated the circumbinary case both analytically and numerically in this article, and discovered that the spin angle variations of circumbinary planets residing near their stellar binary orbital plane are typically much less pronounced than their single star analogue. Thus, their obliquity (spin-orbit misalignment) variations are mostly low.

The physical intuition behind our result is, planets orbiting around a stellar binary reside in a gravitational potential composed of a large quadrupole component due to the stellar separation. The quadrupole potential result in a shorter oscillation period. This detunes systems from spin-orbit resonances and avoid large spin-axis variations. The same result holds for both single planet and multi-planetary systems. However, we note that large moons ($\gtrsim$ mass of our own Moon) close to the planets (slightly outside its Laplace radius ~ 15 Earth-radius) would significantly increase the precession rate [52], and this can lead to enhanced spin-axis variations. Thus, contrary to the Moon of Earth, moons of circumbinary planets could lead to large obliquity variations.

In our secular theory, we neglected mean motion resonances between the stellar binary and the planetary orbit. This simplification was made because lower order mean motion resonances (within $3:1$) reside inside the region of orbital instability [53], and higher order mean motion resonances are weak. We also assumed no resonance between planetary rotation and its orbital mean motion, which could lead to interesting dynamics [54]. This is also a reasonable assumption – circumbinary planets typically orbit wide stellar binaries beyond orbital period of 7 days, with planet orbital period $\gtrsim 30$ days in the orbital stable region [27]. Thus, planet orbital periods tend to be much longer than Earth-like rotation periods, and low order resonances between planet spin and orbital mean motion are thus absent. In addition, over long ($\sim$Gyr) timescales, tidal interactions could aligned the planetary orbit with that of the stellar binary and could affect the planet spin-axis dynamics [55]. As we focus on how planet spin affects insolation and hence habitability, which is already a meaningful subject of investigation in shorter timescales, our study neglected tidal interactions.

Our assumption on the interested planet being Earth-like (with Earth's mass, inertia tensor and rotation rate) is not required by either our theory or numerical methods, but only a choice for starting the investigation. Recently, observational techniques of planetary obliquity and oblateness have been proposed [56, 57, 58], and the first constraint on obliquities of planets outside of the Solar System has been made [59]. In addition, the spin-rate of planets have been measured and revealed to us important clues on planetary formation [60]. Such progress in observation will help us better understand planetary spin properties to constrain obliquity variations. However, so far, these observational techniques could only be applied to massive Jupiter-sized planets, and the spin measurement of terrestrial planets is still beyond the limit of observational sensitivity. For planets whose obliquity observation is available, our work also opens up a possibility of probing otherwise-hard-to-observe planet properties from its obliquity dynamics, as one can apply our tools in an inverse problem setup. In this sense, the application of our theory and numerical tool is beyond constraining habitability.

CHAPTER 3

STRUCTURE-PRESERVING INTEGRATION OF RIGID BODY DYNAMICS

To carefully study effects of spin-orbit coupling (e.g. the spin-axis dynamics of circumbinary planets in Chapter 2), we develop symplectic algorithm ("Gravitationally interacting Rigid-body InTegrator", `GRIT`) starting with the first-principal rigid body dynamics, so that the mutual interactions between spin and orbital dynamics can be accurately accounted for.

3.1 Introduction

Among thousands of detected exoplanetary systems, a significant fraction of them involve planets with close-in orbits. In particular, the occurrence rate for the compact systems (e.g., multiple planets with periods of less than 10 days) are estimated to be $\sim 20 - 30\%$ [61, 62]. The close separations between the planets allow strong planetary interactions that could lead to rich features in the dynamical evolution of the compact planetary systems.

In particular, spin-axis dynamics becomes very interesting in compact planetary systems. For instance, rotational and tidal distortion of the planets can lead orbital precession due to planet spin-orbit coupling, and this causes variations in transit timing. Recently, [63] showed that the transit timing variations due to spin-orbit coupling could be detectable for Trappist-I, which could in turn help one constrain physical properties of the planets. In addition, although tidal effects are strong for planets with close-in orbits, strong interactions between the planets could push these planets (with orbital periods in ~ 10 days) out of synchronized states [64]. Moreover, secular resonance-driven spin-orbit coupling could drive to large obliquity variations and lead to obliquity tides. This sculpts the exoplanetary systems: the obliquity tide could explain the overabundance of planet pairs that reside just wide of the first order mean-motion resonances [65].

Integrator involving spin-axis coupling have been developed to study these effects.

There are mainly two different approaches: 1) evolving the orbital dynamics separately from the spin-axis evolutions [e.g., 5, 43, 64]; 2) evolving the spin and orbit evolution simultaneously [e.g, 66, 67, 68, 6, 69, 70, 65]. In the first approach, orbital evolution of the systems are first integrated using N-body simulation packages assuming the objects are point-mass particles, and then spin-axis dynamics are computed using the results of the orbital evolution. This approach assumes that the effects of the spin on orbital dynamics are weak. In the second approach, additional force due to spin-orbit coupling is included in the N-body simulation package, which could affect the orbital evolution as well as the spin-axis evolution.

Symplectic Lie-Poisson integrator for rigid body has already been constructed in the seminal work of [71] for systems with near Keplerian orbits, focusing on a system with 1 rigid body (for systems with more than 1 rigid bodies, the spin dynamics of each rigid body is considered separately under it's own frame). However, for systems that involve close-encounters, the orbits of object are no longer Keplerian. The original version of the method in [71] was not high-order (in the time step) either. Building upon the existing progress, we no longer assume near Keplerian orbits for wider applicability, and our package includes several high-order implementations. Moreover, we put all the bodies under the same inertia frame such that the spin orbit interactions are all considered altogether in one Hamiltonian framework. We also note that symplectic integrator for secular spin-orbit dynamics have been developed by [72], while our method is based on direct (non-secular) numerical simulations and therefore suitable for resonant situations.

The development of our integrator is tightly based on the profound field of geometric integration. This is because rigid body dynamics can be intrinsically characterized by mechanical systems on Lie groups. More precisely, the phase space is $T^*\mathsf{SE}(3)^{\otimes n}$, where n is the number of interacting bodies and the special Euclidean group $\mathsf{SE}(3)$ is where the center of mass and rotational orientation of each body lives. How to properly simulate such systems in a structure preserving way, so that symplecticity can be conserved and the dynamics remain on the Lie group, has been extensively studied. See e.g., [73, 74, 75] for general Lie group integrators, and more broadly, [76, 77, 78, 79] for monographs on geometric integration.

Regarding rigid body integrators in particular, the following is an incomplete list in addition to [71]. Firstly, the work of [80] used a splitting approach (similar to [71] in essence, however split differently) to construct symplectic and Lie-group-preserving integrators for rigid molecules. The main idea is to split the Hamiltonian into a free rigid body part, including both translational and rotational kinetic energies, plus a potential part. The latter can be exactly integrated, and the former too when the rigid body is axial symmetric; otherwise, it is further split into a symmetric top and a correction term, both of which can be exactly integrated in a cheap way (without using special functions). Methods in the proposed package (which are explicit, high-order integrators) are largely based on this idea. Secondly, we note various splitting schemes for integrating free rigid bodies were compared in [81]. Recall that the free rigid body is integrable, and its numerical simulation based on multiple ways of expressing the exact solution were also proposed (e.g., [82, 83]), but the exact expressions involve special functions (unless the bodies are axial symmetric), which can be computationally expensive. Moreover, the 'exact' solutions are not exact due to round-off errors, and this complication is studied (and remedied) in [84]. For simple and robust arithmetic, the free-rigid-body part of our method will be based on a sub-splitting into an axial-symmetric part and a small correction, as most rotating celestial bodies relevant to this study are (almost) axial-symmetric. Also worth mentioning is, geometric integrators for (non-free but) gravitationally interacting rigid bodies have also been proposed; besides [71], [85] constructed variational integrators using elegant geometric treatments; however, those integrators are implicit, and computational efficiency is hence not optimal.

As we are interested in gravitationally interacting rigid bodies, `GRIT` uses tailored splitting schemes. This way, the existence of small parameters and separation of timescales in the system is utilized so that a better trade-off between efficiency and accuracy can be achieved (see section 3.6.1 for details). Our treatment is of course based on extensive existing studies of splitting methods, and some more general discussions on splitting methods can be found, e.g., in [86, 87, 88].

This chapter is organized as the following: section 3.2 describes the rigid body formulation adopted in the algorithms, section 3.4 derives the equations of motion of the N-rigid-body dynamics and section 3.6 presents our symplectic algorithms. We then show consistency between our simulation using `GRIT` and secular theories in section 3.7, for the case of a moonless Earth and the case of a hypothetical Earth-Moon system that include tidal interactions.

3.2 Rigid Body Representation

As we are interested in the dynamics of the planet's spin-axis, the planet is modeled as a rigid body to account for its finite size and rotation. Other than the spatial position and the linear momentum, the rotational orientation and angular momentum of the rigid body are necessary to represent its state. The above are 12-dimensional in total, and besides the spatial position (3-dim) and its conjugate linear momentum (3-dim), we still need a set of variables (6-dim) to represent the orientation and the rotation of the rigid body.

3.2.1 The Body Frame and The Rotation Matrix

Under a specific fixed reference frame of Euclidean space $\mathbb{R}^3$ with basis $(\boldsymbol{E}_1, \boldsymbol{E}_2, \boldsymbol{E}_3)$, the spatial position and the translational speed of a body can be expressed by vectors in $\mathbb{R}^3$. On the other hand, the body frame (Fig. 3.1) attached to the body gives fixed coordinates of each small particle of the body. As can be seen from Fig. 3.1, orthogonal bases $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ form a body frame of this rigid body. As the body moving along the dashed trajectory following arrows as well as self rotating from time t_0 to time t_1, coordinates of points P, Q under the body frame stay the same, without subjecting to the motion of the rigid body.

The configuration of a rigid body is described by both the position of its center of mass

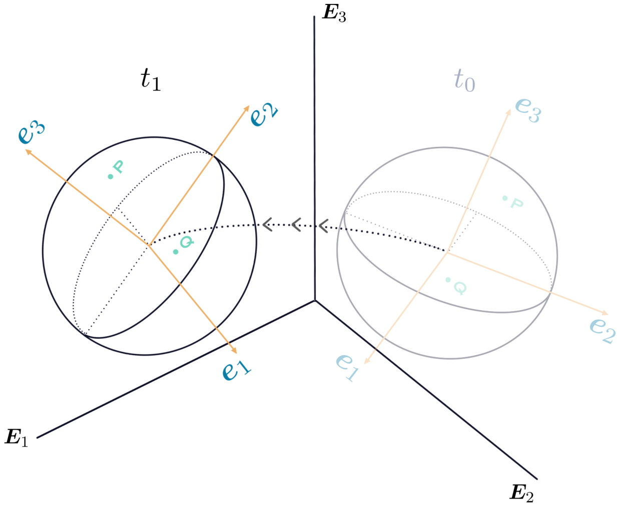


Figure 3.1: The body frame.

and its rotational orientation. The orientation in the reference frame can be expressed as an rotation by an orthogonal matrix $\boldsymbol{R}(t) \in \mathsf{SO}(3)$ from the body frame (e.g., z-axis of the body frame at time t will be $\boldsymbol{R}(t) \cdot \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$ in the reference frame). To switch between the inertia frame and the body frame, one can simply left multiply the rotation matrix $\boldsymbol{R}$ or $\boldsymbol{R}^{-1}$. Note that $\boldsymbol{R} \in \mathsf{SO}(3)$ and if a numerical method can keep $\boldsymbol{R}$ exactly in this Lie group, its inverse will be equal to its transpose, i.e. $\boldsymbol{R}^{-1} = \boldsymbol{R}^T$.

3.2.2 The Angular Velocity and the Angular Momentum

Denoting $\boldsymbol{\Omega} = \begin{bmatrix} \Omega_1 & \Omega_2 & \Omega_3 \end{bmatrix}^T \in \mathbb{R}^3$ the angular velocity of the rigid body under the body frame, then the direction of $\boldsymbol{\Omega}$ matches the rotational axis and $\|\boldsymbol{\Omega}\|_2$ represents the rotational speed. Consider a mass point $\boldsymbol{x} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}^T$ in one rigid body under the body frame, its speed under the body frame can be expressed as

$$\boldsymbol{\Omega} \times \boldsymbol{x} = \begin{bmatrix} 0 & -\Omega_3 & \Omega_2 \\ \Omega_3 & 0 & -\Omega_1 \\ -\Omega_2 & \Omega_1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \hat{\boldsymbol{\Omega}} \boldsymbol{x}, \tag{3.1}$$

where the hat-map $\hat{\cdot}$ is an isomorphism from the Lie algebra $\mathfrak{so}(3)$ to 3-by-3 skew-symmetric matrices, defined by

$$\hat{\boldsymbol{\Omega}} := \begin{bmatrix} 0 & -\Omega_3 & \Omega_2 \\ \Omega_3 & 0 & -\Omega_1 \\ -\Omega_2 & \Omega_1 & 0 \end{bmatrix}. \tag{3.2}$$

In addition, the inverse map of $\hat{\cdot}$ is denoted by $\check{\cdot}$.

With the angular velocity, the rotational kinetic energy of this rigid body can be expressed as

$$\begin{aligned} T^{rot}(\boldsymbol{\Omega}) &= \int_{\mathcal{B}} \rho(\boldsymbol{x}) \frac{1}{2} \left\| \boldsymbol{\Omega} \times \boldsymbol{x} \right\|_2^2 \, d\boldsymbol{x} \\ &= \int_{\mathcal{B}} \rho(\boldsymbol{x}) \frac{1}{2} \left\| \hat{\boldsymbol{x}} \boldsymbol{\Omega} \right\|_2^2 \, d\boldsymbol{x} \\ &= \frac{1}{2} \boldsymbol{\Omega}^T \left[\int_{\mathcal{B}} \rho(\boldsymbol{x}) \hat{\boldsymbol{x}}^T \hat{\boldsymbol{x}} \, d\boldsymbol{x} \right] \boldsymbol{\Omega} \\ &= \frac{1}{2} \boldsymbol{\Omega}^T \boldsymbol{J} \boldsymbol{\Omega} \end{aligned} \tag{3.3}$$

with $\boldsymbol{J}$ the (standard) moment of inertia tensor. Specifically, for an ellipsoid with semiaxes a, b, c and mass M, choosing the principal axes as the body frame such that x, y, z-axes matches semi-axes and taking the integral, we have the following moment of inertia tensor for a uniform density object.

$$\begin{aligned} \boldsymbol{J} &= \int_{\mathcal{B}} \rho(\boldsymbol{x}) \hat{\boldsymbol{x}}^T \hat{\boldsymbol{x}} \, d^3\boldsymbol{x} \\ &= \begin{bmatrix} \frac{1}{5} M (b^2 + c^2) & 0 & 0 \\ 0 & \frac{1}{5} M (a^2 + c^2) & 0 \\ 0 & 0 & \frac{1}{5} M (a^2 + b^2) \end{bmatrix}. \end{aligned} \tag{3.4}$$

Note that one may substitute this with the principal moment of inertia directly.

Alternatively, the rotational kinetic energy can also be expressed as

$$\begin{aligned}T^{rot}\left(\hat{\boldsymbol{\Omega}}\right) &= \int_{\mathcal{B}} \rho(\boldsymbol{x})\frac{1}{2}\left\|\hat{\boldsymbol{\Omega}}\boldsymbol{x}\right\|_2^2\, d\boldsymbol{x} \\ &= \frac{1}{2}\int_{\mathcal{B}} \rho(\boldsymbol{x}) Tr\left[\hat{\boldsymbol{\Omega}}\boldsymbol{x}\boldsymbol{x}^T\hat{\boldsymbol{\Omega}}^T\right]\, d\boldsymbol{x} \\ &= \frac{1}{2} Tr\left[\hat{\boldsymbol{\Omega}}\boldsymbol{J}_d\hat{\boldsymbol{\Omega}}^T\right].\end{aligned} \tag{3.5}$$

with $\boldsymbol{J}^{(d)} = \int_{\mathcal{B}} \rho(\boldsymbol{x})\boldsymbol{x}\boldsymbol{x}^T\, d\boldsymbol{x}$ (nonstandard) moment of inertia. We also have $\boldsymbol{J}^{(d)} = \frac{1}{2} Tr\left[\boldsymbol{J}\right]\boldsymbol{I}_{3\times3} - \boldsymbol{J}$ ($\boldsymbol{J} = Tr\left[\boldsymbol{J}^{(d)}\right]\boldsymbol{I}_{3\times3} - \boldsymbol{J}^{(d)}$).

By definition, the angular momentum in the body frame is $\boldsymbol{\Pi} = \boldsymbol{J}\boldsymbol{\Omega}$. Left multiplying the rotation matrix $\boldsymbol{R}$, the angular velocity and the angular momentum in the inertia frame are $\boldsymbol{\omega} = \boldsymbol{R}\boldsymbol{\Omega}$ and $\boldsymbol{\pi} = \boldsymbol{R}\boldsymbol{\Pi}$ respectively.

3.2.3 The Relation between the Rotation Matrix and the Angular Velocity

Express $\boldsymbol{R}(t) = \left[\boldsymbol{c}_1(t)|\boldsymbol{c}_2(t)|\boldsymbol{c}_3(t)\right]$ where $\boldsymbol{c}_1(t)$, $\boldsymbol{c}_2(t)$, $\boldsymbol{c}_3(t)$ are columns of $\boldsymbol{R}(t)$. We have $\boldsymbol{c}_1(t)$, $\boldsymbol{c}_2(t)$, $\boldsymbol{c}_3(t)$ representing directions of three axes of the body in the reference frame respectively. By the definition of angular velocity, we have $\dot{\boldsymbol{c}}_i(t) = \hat{\boldsymbol{\omega}}c_i(t)$ for $i = 1, 2, 3$, thus

$$\dot{\boldsymbol{R}}(t) = \hat{\boldsymbol{\omega}}\boldsymbol{R}(t). \tag{3.6}$$

Multiplying both sides of Eq. (3.6) with $\boldsymbol{R}(t)^T$, we have $\dot{\boldsymbol{R}}(t)\boldsymbol{R}(t)^T = \hat{\boldsymbol{\omega}}$ which is a skew-symmetric matrix. Considering the speed of an arbitrary mass point $\boldsymbol{x}$, $\boldsymbol{v}_{\boldsymbol{x}} = \boldsymbol{R}\hat{\boldsymbol{\Omega}}\boldsymbol{x} = \hat{\boldsymbol{\omega}}\left[\boldsymbol{R}\boldsymbol{x}\right]$. Thus $\boldsymbol{R}\hat{\boldsymbol{\Omega}} = \hat{\boldsymbol{\omega}}\boldsymbol{R}$, which implies $\hat{\boldsymbol{\Omega}} = \boldsymbol{R}^T\hat{\boldsymbol{\omega}}\boldsymbol{R} = \boldsymbol{R}^T\dot{\boldsymbol{R}}$.

To summarize, for a rigid body, it's angular velocity and angular momentum in different frames are denoted as the following,

with $\boldsymbol{\Pi} = \boldsymbol{J}\boldsymbol{\Omega}$ and $\boldsymbol{\pi} = \boldsymbol{J}\boldsymbol{\omega}$. Specifically, we have $\hat{\boldsymbol{\Omega}} = \boldsymbol{R}^T\dot{\boldsymbol{R}}$ and $\hat{\boldsymbol{\omega}} = \dot{\boldsymbol{R}}\boldsymbol{R}^T$.

The rotation matrix $\boldsymbol{R}$ and the angular momentum $\boldsymbol{\Pi}$ will be utilized to describe a rigid body when we design an N-rigid-body integrator later (details can be found in section 3.6).

	The inertia frame (fixed)	The body frame (moving)
Angular velocity	$\boldsymbol{\omega}\,(=\boldsymbol{R}\boldsymbol{\Omega})$	$\boldsymbol{\Omega}$
Angular momentum	$\boldsymbol{\pi}\,(=\boldsymbol{R}\boldsymbol{\Pi})$	$\boldsymbol{\Pi}$

3.3 The Constrained Hamiltonian of an N-rigid-body System

Denote m_i the mass of the i-th body; $\boldsymbol{q}_i \in \mathbb{R}^3$ the position of the i-th body; $\boldsymbol{p}_i \in \mathbb{R}^3$ the linear momentum of the i-th body; $\boldsymbol{R}_i \in \mathsf{SO}(3)$ the rotation matrix of the i-th body; $\boldsymbol{\Pi}_i \in \mathbb{R}^3$ the angular momentum of the i-th body; $\boldsymbol{J}_i \in \mathbb{R}^{3\times 3}$ the (standard) moment of inertia tensor for the i-th body.

The Hamiltonian of this system consists of the linear kinetic energy $T^{linear} = \sum_i \frac{1}{2}\boldsymbol{p}_i^T\boldsymbol{p}_i/m_i$, the rotational kinetic energy $T^{rot} = \sum_i \frac{1}{2}\boldsymbol{\Pi}_i^T\boldsymbol{J}_i^{-1}\boldsymbol{\Pi}_i$ and the potential energy

$$V(\mathsf{q},\mathsf{R}) = \sum_{i<j} V_{ij}\left(\boldsymbol{q}_i,\boldsymbol{q}_j,\boldsymbol{R}_i,\boldsymbol{R}_j\right). \tag{3.7}$$

Denote $\mathsf{q} = \{\boldsymbol{q}_1,\boldsymbol{q}_2,\ldots,\boldsymbol{q}_N\}$, $\mathsf{p} = \{\boldsymbol{p}_1,\boldsymbol{p}_2,\ldots,\boldsymbol{p}_N\}$, $\mathsf{\Pi} = \{\boldsymbol{\Pi}_1,\boldsymbol{\Pi}_2,\ldots,\boldsymbol{\Pi}_N\}$, $\mathsf{R} = \{\boldsymbol{R}_1,\boldsymbol{R}_2,\ldots,\boldsymbol{R}_N\}$. The Hamiltonian can be expressed as

$$\begin{aligned} H(\mathsf{q},\mathsf{p},\mathsf{\Pi},\mathsf{R}) = &\sum_i \frac{1}{2}\boldsymbol{p}_i^T\boldsymbol{p}_i/m_i \\ &+\sum_i \frac{1}{2}\boldsymbol{\Pi}_i^T\boldsymbol{I}_i^{-1}\boldsymbol{\Pi}_i + V(\mathsf{q},\mathsf{R}) \qquad \text{with } \boldsymbol{R}_i \in \mathsf{SO}(3). \end{aligned} \tag{3.8}$$

The true potential energy between i-th body and j-th body is

$$\int_{\mathcal{B}_i}\int_{\mathcal{B}_j} -\frac{\mathcal{G}\rho(\boldsymbol{x}_i)\rho(\boldsymbol{x}_j)}{\|\,(\boldsymbol{q}_i+\boldsymbol{R}_i\boldsymbol{x}_i)-(\boldsymbol{q}_j+\boldsymbol{R}_j\boldsymbol{x}_j)\,\|}\,d\boldsymbol{x}_j\,d\boldsymbol{x}_i. \tag{3.9}$$

We may approximate it as V_{ij} (in Eq. (3.7)) by Taylor expanding the denominator. Expanding to the 2nd order with respect to the radius of the planet over the distance between two bodies (see appendix section B.1), the approximated potential is,

$$\begin{aligned}
&\int_{\mathcal{B}_i}\int_{\mathcal{B}_j} -\frac{\mathcal{G}\rho(\boldsymbol{x}_i)\rho(\boldsymbol{x}_j)}{\|\,(\boldsymbol{q}_i+\boldsymbol{R}_i\boldsymbol{x}_i)-(\boldsymbol{q}_j+\boldsymbol{R}_j\boldsymbol{x}_j)\,\|}\,d\boldsymbol{x}_j\,d\boldsymbol{x}_i \\
&\approx -\frac{\mathcal{G}m_im_j}{\|\boldsymbol{q}_i-\boldsymbol{q}_j\|} - \frac{\mathcal{G}\left(m_iTr[\boldsymbol{J}_i]+m_jTr[\boldsymbol{J}_j]\right)}{2\,\|\boldsymbol{q}_i-\boldsymbol{q}_j\|^3} \\
&+ \frac{3\mathcal{G}(\boldsymbol{q}_i-\boldsymbol{q}_j)^T\left(m_j\boldsymbol{R}_i\boldsymbol{J}_i\boldsymbol{R}_i^T+m_i\boldsymbol{R}_j\boldsymbol{J}_j\boldsymbol{R}_j^T\right)(\boldsymbol{q}_i-\boldsymbol{q}_j)}{2\,\|\boldsymbol{q}_i-\boldsymbol{q}_j\|^5}
\end{aligned} \tag{3.10}$$

with $-\frac{\mathcal{G}m_im_j}{\|\boldsymbol{q}_i-\boldsymbol{q}_j\|}$ being the potential of purely point mass interactions and the rest part being the corrections of the potential due to the body i and j being not point masses. If we further expand the potential to the 4th order (see appendix section B.1), rigid body – rigid body interactions will also be included as higher order corrections.

3.4 Review: Equations of Motion of an N-rigid-body System

We will first review the derivation of equations of motion of one rigid body in a potential using two approaches then apply it to the N-rigid-body systems.

3.4.1 Equations of Motion of One Rigid Body in a Potential

The Lagrangian for a system consisting of one rigid body is a function of $\boldsymbol{R}(t)$ and $\dot{\boldsymbol{R}}(t)$ by plugging in $\hat{\boldsymbol{\Omega}} = \boldsymbol{R}^T\dot{\boldsymbol{R}}$ in Eq. (3.5),

$$L\left(\boldsymbol{R},\dot{\boldsymbol{R}}\right) = \frac{1}{2}Tr\left[\dot{\boldsymbol{R}}\boldsymbol{J}_d\dot{\boldsymbol{R}}^T\right] - V(\boldsymbol{R}). \tag{3.11}$$

Utilizing the constraint $\boldsymbol{R}^T\boldsymbol{R}-\boldsymbol{I}=\boldsymbol{0}$ (approach 1 below) or using the variational principle of Hamilton's for Lie group (approach 2 below), one can derive the equations of motion

$$\begin{cases} \dot{\boldsymbol{R}} = \boldsymbol{R}\widehat{\boldsymbol{J}^{-1}\boldsymbol{\Pi}}, \\ \dot{\boldsymbol{\Pi}} = \boldsymbol{\Pi} \times \boldsymbol{J}^{-1}\boldsymbol{\Pi} - \left(\boldsymbol{R}^T \frac{\partial V(\boldsymbol{R})}{\partial \boldsymbol{R}} - \left(\frac{\partial V(\boldsymbol{R})}{\partial \boldsymbol{R}} \right)^T \boldsymbol{R} \right)^{\vee}. \end{cases} \tag{3.12}$$

Approach 1

We can view $\boldsymbol{R}$ to be in the embedded Euclidean space $\mathbb{R}^{3\times 3} \hookleftarrow \mathsf{SO}(3)$ and use $\boldsymbol{R} \in \mathsf{SO}(3)$ as a holonomic constraint. The Lagrangian L (Eq. (3.11)) has 9-DOF before applying the constraint $\boldsymbol{R} \in \mathsf{SO}(3)$. The conjugate variable of $\boldsymbol{R}(t)$ will be denoted by $\boldsymbol{P}(t)$.

The constraint of a system forces the evolution of the system in a specific manifold, and the manifold can be directly calculated from the constraint (one may refer Chapter VII of [89] for details). For a rigid body dynamics represented by a rotation matrix $\boldsymbol{R}(t)$, the constraint is $\boldsymbol{R}(t)^T\boldsymbol{R}(t) - \boldsymbol{I}_{3\times 3} = \boldsymbol{0}_{3\times 3}$. [90], [89] have shown the procedure of finding equations of motion by utilizing the constraint for a rigid body system with a $\boldsymbol{R}$ dependent potential. Using Lagrange multipliers [89] for the constraint $\boldsymbol{R}^T\boldsymbol{R} - \boldsymbol{I}_{3\times 3} = \boldsymbol{0}$, we have the following Lagrangian,

$$L\left(\boldsymbol{R}, \dot{\boldsymbol{R}}\right) = \frac{1}{2}Tr\left[\dot{\boldsymbol{R}}\boldsymbol{J}_d\dot{\boldsymbol{R}}^T\right] - V(\boldsymbol{R}) - \frac{1}{2}Tr\left(\boldsymbol{\Lambda}^T\left(\boldsymbol{R}^T\boldsymbol{R} - \boldsymbol{I}_{3\times 3}\right)\right), \tag{3.13}$$

with 6-dim Lagrange multipliers $\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & \lambda_4 & \lambda_6 \\ \lambda_4 & \lambda_2 & \lambda_5 \\ \lambda_6 & \lambda_5 & \lambda_3 \end{bmatrix} \in \mathbb{R}^{3\times 3}$ a symmetric matrix.

Doing Legendre transform for Eq. (3.13), we have

$$\boldsymbol{P} = \frac{\partial L\left(\boldsymbol{R}, \dot{\boldsymbol{R}}\right)}{\partial \dot{\boldsymbol{R}}} = \dot{\boldsymbol{R}}\boldsymbol{J}_d, \tag{3.14}$$

and the corresponding Hamiltonian,

$$H(\boldsymbol{R},\boldsymbol{P}) = \frac{1}{2}Tr\left[\boldsymbol{P}\boldsymbol{J}_d^{-1}\boldsymbol{P}^T\right] + V(\boldsymbol{R}) + \frac{1}{2}Tr\left(\boldsymbol{\Lambda}^T\left(\boldsymbol{R}^T\boldsymbol{R} - \boldsymbol{I}_{3\times3}\right)\right). \tag{3.15}$$

As the constraint for $\boldsymbol{R}$ is $\boldsymbol{R}^T\boldsymbol{R} = \boldsymbol{I}_{3\times3}$, according to [89], the constraint for $\boldsymbol{P}$ can be obtained by taking time derivative for $\boldsymbol{R}^T\boldsymbol{R} - \boldsymbol{I}_{3\times3} = \boldsymbol{0}_{3\times3}$, i.e. $\boldsymbol{J}_d^{-1}\boldsymbol{P}^T\boldsymbol{R} + \boldsymbol{R}^T\boldsymbol{P}\boldsymbol{J}_d^{-1} = \boldsymbol{0}_{3\times3}$.

So,

$$\begin{cases} \dot{\boldsymbol{R}} = \dfrac{\partial H}{\partial \boldsymbol{P}} = \boldsymbol{P}\boldsymbol{J}_d{}^{-1}, \\ \dot{\boldsymbol{P}} = -\dfrac{\partial H}{\partial \boldsymbol{R}} = -\dfrac{\partial V(\boldsymbol{R})}{\partial \boldsymbol{R}} - \boldsymbol{R}\boldsymbol{\Lambda}, \end{cases} \tag{3.16}$$

on the manifold

$$\mathcal{M} = \left\{(\boldsymbol{R},\boldsymbol{P})\,|\boldsymbol{R}^T\boldsymbol{R} = \boldsymbol{I}_{3\times3}, \boldsymbol{J}_d^{-1}\boldsymbol{P}^T\boldsymbol{R} + \boldsymbol{R}^T\boldsymbol{P}\boldsymbol{J}_d^{-1} = \boldsymbol{0}_{3\times3}\right\}. \tag{3.17}$$

Note that $\hat{\boldsymbol{\Omega}} = \boldsymbol{R}^T\boldsymbol{P}\boldsymbol{J}_d^{-1}$ with $\boldsymbol{\Omega}$ being the body's angular velocity. Taking time derivative for $\hat{\boldsymbol{\Omega}}$, we have

$$\dot{\hat{\boldsymbol{\Omega}}} = \boldsymbol{J}_d^{-1}\boldsymbol{P}^T\boldsymbol{P}\boldsymbol{J}_d^{-1} + \boldsymbol{R}^T\left(-\frac{\partial V(\boldsymbol{R})}{\partial \boldsymbol{R}} - \boldsymbol{R}\boldsymbol{\Lambda}\right)\boldsymbol{J}_d^{-1}. \tag{3.18}$$

Physically, we want to find dynamics of $\boldsymbol{R}$ and the body's angular momentum $\boldsymbol{\Pi}$. Since $\hat{\boldsymbol{\Pi}} = \widehat{\boldsymbol{J}\boldsymbol{\Omega}} = Tr\left[\boldsymbol{J}_d\right]\hat{\boldsymbol{\Omega}} - \widehat{\boldsymbol{J}_d\boldsymbol{\Omega}} = \hat{\boldsymbol{\Omega}}\boldsymbol{J}_d - \boldsymbol{J}_d\hat{\boldsymbol{\Omega}}^T$ (see appendix section B.2), we may find dynamics of $\boldsymbol{\Pi}$,

$$\dot{\hat{\boldsymbol{\Pi}}} = \left(\boldsymbol{J}_d^{-1}\boldsymbol{P}^T\boldsymbol{P} - \boldsymbol{P}^T\boldsymbol{P}\boldsymbol{J}_d^{-1}\right) + \left(\left(\frac{\partial V(\boldsymbol{R})}{\partial \boldsymbol{R}}\right)^T\boldsymbol{R} - \boldsymbol{R}^T\frac{\partial V(\boldsymbol{R})}{\partial \boldsymbol{R}}\right), \tag{3.19}$$

with the symmetric $\boldsymbol{\Lambda}$ vanished [1].

---

[1] Since Λ is symmetric, applying $\dot{\hat{\boldsymbol{\Omega}}} \in \mathfrak{so}(3)$, Λ can actually be solved from Eq. (3.18).

As $\boldsymbol{P} = \boldsymbol{R}\hat{\boldsymbol{\Omega}}\boldsymbol{J}_d$, properties of hat-map (see appendix section B.2) lead to

$$\begin{aligned}\widehat{\dot{\boldsymbol{\Pi}}} &= \left(\hat{\boldsymbol{\Omega}}^T\hat{\boldsymbol{\Omega}}\boldsymbol{J}_d - \boldsymbol{J}_d\hat{\boldsymbol{\Omega}}^T\hat{\boldsymbol{\Omega}}\right) + \left(\left(\frac{\partial V(\boldsymbol{R})}{\partial \boldsymbol{R}}\right)^T \boldsymbol{R} - \boldsymbol{R}^T\frac{\partial V(\boldsymbol{R})}{\partial \boldsymbol{R}}\right)\\ &= \widehat{\boldsymbol{\Omega}\times \boldsymbol{J}_d}\boldsymbol{\Omega} + \left(\left(\frac{\partial V(\boldsymbol{R})}{\partial \boldsymbol{R}}\right)^T \boldsymbol{R} - \boldsymbol{R}^T\frac{\partial V(\boldsymbol{R})}{\partial \boldsymbol{R}}\right).\end{aligned} \tag{3.20}$$

Thus

$$\begin{aligned}\dot{\boldsymbol{\Pi}} &= \boldsymbol{\Omega}\times\boldsymbol{J}_d\boldsymbol{\Omega} - \left(\boldsymbol{R}^T\frac{\partial V(\boldsymbol{R})}{\partial \boldsymbol{R}} - \left(\frac{\partial V(\boldsymbol{R})}{\partial \boldsymbol{R}}\right)^T\boldsymbol{R}\right)^\vee\\ &= \boldsymbol{\Omega}\times\left(Tr[\boldsymbol{J}_d] - \boldsymbol{J}\right)\boldsymbol{\Omega} - \left(\boldsymbol{R}^T\frac{\partial V(\boldsymbol{R})}{\partial \boldsymbol{R}} - \left(\frac{\partial V(\boldsymbol{R})}{\partial \boldsymbol{R}}\right)^T\boldsymbol{R}\right)^\vee\\ &= -\boldsymbol{\Omega}\times\boldsymbol{J}\boldsymbol{\Omega} - \left(\boldsymbol{R}^T\frac{\partial V(\boldsymbol{R})}{\partial \boldsymbol{R}} - \left(\frac{\partial V(\boldsymbol{R})}{\partial \boldsymbol{R}}\right)^T\boldsymbol{R}\right)^\vee\\ &= \boldsymbol{\Pi}\times\boldsymbol{J}^{-1}\boldsymbol{\Pi} - \left(\boldsymbol{R}^T\frac{\partial V(\boldsymbol{R})}{\partial \boldsymbol{R}} - \left(\frac{\partial V(\boldsymbol{R})}{\partial \boldsymbol{R}}\right)^T\boldsymbol{R}\right)^\vee\end{aligned} \tag{3.21}$$

So, equations of motion with respect to $\boldsymbol{R}$ and $\boldsymbol{\Pi}$ for one rigid body system are

$$\begin{cases}\dot{\boldsymbol{R}} = \boldsymbol{R}\widehat{\boldsymbol{J}^{-1}\boldsymbol{\Pi}},\\ \dot{\boldsymbol{\Pi}} = \boldsymbol{\Pi}\times\boldsymbol{J}^{-1}\boldsymbol{\Pi} - \left(\boldsymbol{R}^T\frac{\partial V(\boldsymbol{R})}{\partial \boldsymbol{R}} - \left(\frac{\partial V(\boldsymbol{R})}{\partial \boldsymbol{R}}\right)^T\boldsymbol{R}\right)^\vee.\end{cases} \tag{3.22}$$

Approach 2

How to obtain Euler-Lagrange equation for the Hamilton's variational principle on a Lie group has been well studied (e.g., [91, 92]). Here we summarize the results for the special case of rigid bodies from the expository part of [93].

Denote the infinitesimally varied rotation by $\boldsymbol{R}_\epsilon = \boldsymbol{R}\exp(\epsilon\hat{\boldsymbol{\eta}})$ with $\epsilon \in \mathbb{R}$ and $\boldsymbol{\eta} \in \mathbb{R}^3$, where $\exp(\cdot)$ is a mapping from $\mathfrak{so}(3)$ to $\mathsf{SO}(3)$. The varied angular velocity is

$$\begin{aligned}\hat{\boldsymbol{\Omega}}_\epsilon = \boldsymbol{R}_\epsilon^T \dot{\boldsymbol{R}}_\epsilon &= e^{-\epsilon\hat{\boldsymbol{\eta}}}\boldsymbol{R}^T\left(\dot{\boldsymbol{R}}e^{\epsilon\hat{\boldsymbol{\eta}}} + \boldsymbol{R}\cdot e^{\epsilon\hat{\boldsymbol{\eta}}}\epsilon\hat{\dot{\boldsymbol{\eta}}}\right) \\ &= e^{-\epsilon\hat{\boldsymbol{\eta}}}\hat{\boldsymbol{\Omega}}e^{\epsilon\hat{\boldsymbol{\eta}}} + \epsilon\hat{\dot{\boldsymbol{\eta}}} = \hat{\boldsymbol{\Omega}} + \epsilon\left\{\hat{\dot{\boldsymbol{\eta}}} + \hat{\boldsymbol{\Omega}}\hat{\boldsymbol{\eta}} - \hat{\boldsymbol{\eta}}\hat{\boldsymbol{\Omega}}\right\} + \mathcal{O}(\epsilon^2).\end{aligned} \tag{3.23}$$

Consider the action

$$S\left(\boldsymbol{\Omega}, \boldsymbol{R}\right) = \int_{t_0}^{t_1} L\left(\boldsymbol{\Omega}, \boldsymbol{R}\right)\, dt = \int_{t_0}^{t_1} \frac{1}{2}Tr\left[\hat{\boldsymbol{\Omega}}\boldsymbol{J}_d\hat{\boldsymbol{\Omega}}^T\right] - V(\boldsymbol{R})\, dt. \tag{3.24}$$

Taking the variation of the action S, we have

$$\begin{aligned}S_\epsilon\left(\boldsymbol{\Omega}, \boldsymbol{R}\right) =& S\left(\boldsymbol{\Omega}_\epsilon, \boldsymbol{R}_\epsilon\right) \\ =& S\left(\boldsymbol{\Omega}, \boldsymbol{R}\right) + \epsilon\int_{t_0}^{t_1}\frac{1}{2}Tr\Big[-\hat{\dot{\boldsymbol{\eta}}}\left(\boldsymbol{J}_d\hat{\boldsymbol{\Omega}} + \hat{\boldsymbol{\Omega}}\boldsymbol{J}_d\right) \\ &+ \hat{\boldsymbol{\eta}}\hat{\boldsymbol{\Omega}}\left(\boldsymbol{J}_d\hat{\boldsymbol{\Omega}} + \hat{\boldsymbol{\Omega}}\boldsymbol{J}_d\right) - \hat{\boldsymbol{\eta}}\left(\boldsymbol{J}_d\hat{\boldsymbol{\Omega}} + \hat{\boldsymbol{\Omega}}\boldsymbol{J}_d\hat{\boldsymbol{\Omega}}\right)\Big] \\ &+ Tr\left[\hat{\boldsymbol{\eta}}\boldsymbol{R}^T\frac{\partial\boldsymbol{R}}{\partial\boldsymbol{R}}\right]\, dt + \mathcal{O}(\epsilon^2).\end{aligned} \tag{3.25}$$

Using *Hamilton's Principle*, we have $\frac{d}{d\epsilon}\Big|_{\epsilon=0} S_\epsilon = 0$, i.e.

$$\frac{1}{2}\int_{t_0}^{t_1} Tr\left[\hat{\boldsymbol{\eta}}\left\{\widehat{\boldsymbol{J}\dot{\boldsymbol{\Omega}}} + \widehat{\boldsymbol{\Omega}\times\boldsymbol{J}}\Omega + 2\boldsymbol{R}^T\frac{\partial V}{\partial\boldsymbol{R}}\right\}\right] = 0 \tag{3.26}$$

for any $\boldsymbol{\eta}\in\mathbb{R}^3$. Therefore, $\left\{\widehat{\boldsymbol{J}\dot{\boldsymbol{\Omega}}} + \widehat{\boldsymbol{\Omega}\times\boldsymbol{J}}\Omega + 2\boldsymbol{R}^T\frac{\partial V}{\partial\boldsymbol{R}}\right\}$ must be skew-symmetric, which gives us

$$\widehat{\boldsymbol{J}\dot{\boldsymbol{\Omega}}} = -\widehat{\boldsymbol{\Omega}\times\boldsymbol{J}}\boldsymbol{\Omega} + \left(\frac{\partial V}{\partial\boldsymbol{R}}^T\boldsymbol{R} - \boldsymbol{R}^T\frac{\partial V}{\partial\boldsymbol{R}}\right). \tag{3.27}$$

Thus

$$\hat{\dot{\boldsymbol{\Pi}}} = \boldsymbol{\Pi}\widehat{\times\boldsymbol{J}^{-1}}\boldsymbol{\Pi} + \left(\frac{\partial V}{\partial\boldsymbol{R}}^T\boldsymbol{R} - \boldsymbol{R}^T\frac{\partial V}{\partial\boldsymbol{R}}\right). \tag{3.28}$$

3.4.2 Equations of Motion of an N-rigid-body System

Deriving the translational equations of motion directly from the Hamiltonian Eq. (3.8) and applying Eq. (3.12) to each object for the rotational dynamics, the equations of motion of the N-rigid-body system for the Hamiltonian (Eq. (3.8)) are,

$$\begin{cases} \dot{\boldsymbol{q}}_i = \dfrac{\boldsymbol{p}_i}{m_i}, \\ \dot{\boldsymbol{p}}_i = -\dfrac{\partial V}{\partial \boldsymbol{q}_i}, \\ \dot{\boldsymbol{R}}_i = \boldsymbol{R}_i \widehat{\boldsymbol{J}_i^{-1} \boldsymbol{\Pi}_i}, \\ \dot{\boldsymbol{\Pi}}_i = \boldsymbol{\Pi}_i \times \boldsymbol{J}_i^{-1} \boldsymbol{\Pi}_i - \left(\boldsymbol{R}_i^T \dfrac{\partial V}{\partial \boldsymbol{R}_i} - \left(\dfrac{\partial V}{\partial \boldsymbol{R}_i} \right)^T \boldsymbol{R}_i \right)^{\vee}. \end{cases} \tag{3.29}$$

3.5 Review: Symplectic Integrators of Hamiltonian Systems via Splitting and Composition Methods

Symplectic integrators are suitable for the Hamiltonian system as it preserves the symplectic structure, which benefits the integrators with conservation properties, nice long time global accuracy under some general assumptions [89] etc. And a widely used method to construct symplectic integrators of a Hamiltonian system is the splitting and composition method, which firstly divide the Hamiltonian into several integrable parts, then compose the flows together. For our system Eq. (3.8), different ways of splittings will be shown in section 3.6. Given the flows of the split sub-Hamiltonians, then the composition method is employed to construct different orders of schemes via different composition coefficients. We list the composition methods used in the paper below for the general $H = A + B$ and perturbative Hamiltonian $H = A + \varepsilon B$ (e.g. near-integrable Hamiltonians) in Table. 3.1 and Table. 3.2 respectively.

Table 3.1: Composition methods $\mathcal{C}(\cdot,\cdot)$ of general $H = A + B$. φ_h^A and φ_h^B are flows of A and B respectively.

composition method	order
$\mathcal{C}_{\text{Euler}}\left(\varphi_h^A, \varphi_h^B\right) := \varphi_h^A \circ \varphi_h^B$	(1)
$\mathcal{C}_{\text{Verlet}}\left(\varphi_h^A, \varphi_h^B\right) := \varphi_{h/2}^A \circ \varphi_h^B \circ \varphi_{h/2}^A$	(2)
$\mathcal{C}_{\text{TriJump}}\left(\varphi_h^A, \varphi_h^B\right) := \varphi_{\gamma_1 h} \circ \varphi_{\gamma_2 h} \circ \varphi_{\gamma_1 h}$ with $\varphi_h := \mathcal{C}_{\text{Verlet}}(\varphi_h^A, \varphi_h^B)$ and $\gamma_1 = {}^{1}/_{(2-2^{1/3})}$, $\gamma_2 = 1 - 2\gamma_1$[94].	(4)
$\mathcal{C}_{S6}\left(\varphi_h^A, \varphi_h^B\right) := \varphi_{a_1 h} \circ \varphi_{a_2 h} \circ \varphi_{a_3 h} \circ \varphi_{a_4 h} \circ \varphi_{a_3 h} \circ \varphi_{a_2 h} \circ \varphi_{a_1 h}$ with $\varphi_h := \mathcal{C}_{\text{Verlet}}(\varphi_h^A, \varphi_h^B)$ and $a_1 = 0.784513610477560$, $a_2 = 0.235573213359357$, $a_3 = -1.17767998417887$, $a_4 = 1 - 2(a_1 + a_2 + a_3)$[95]	(6)

Table 3.2: Composition methods $\mathcal{C}(\cdot,\cdot)$ of perturbative $H = A + \varepsilon B$. φ_h^A and φ_h^B are flows of A and εB respectively.

composition method	order
$\mathcal{C}_{BAB22}(\varphi_h^A, \varphi_h^B) = \varphi_{h/2}^B \circ \varphi_h^A \circ \varphi_{h/2}^B$	$(\cdot, 2, 2)$
$\mathcal{C}_{ABA22}(\varphi_h^A, \varphi_h^B) = \varphi_{h/2}^A \circ \varphi_h^B \circ \varphi_{h/2}^A$	$(\cdot, 2, 2)$
$\mathcal{C}_{ABA42}(\varphi_h^A, \varphi_h^B) = \varphi^A_{\frac{(3-\sqrt{3})h}{6}} \circ \varphi^B_{\frac{h}{2}} \circ \varphi^A_{\frac{h}{\sqrt{3}}} \circ \varphi^B_{\frac{h}{2}} \circ \varphi^A_{\frac{(3-\sqrt{3})h}{6}}$. ($\mathcal{SABA}_2$ in [96] or equivalently the order $(4,2)$ ABA method with $s = 2$ in [97])	$(\cdot, 4, 2)$

3.6 Rigid Body Simulation: Structure-Preserving Algorithms via Specially Designed Splittings and Compositions

In this section, we will design symplectic integrators of the N-rigid-body system using splitting methods. The splitting method is basically to view the Hamiltonian (Eq. (3.8)) as the sum of several integrable parts, and then to compose the flow of each part over some pre-designed time duration to achieve a certain order of local error. In the following, we will introduce the Hamiltonian, build the symplectic integrators and analyze the accuracy of integrators step by step. In addition, we will provide a way to incorporate non-conservative forces into the integrators, such as the tidal force and post Newtonian effects.

3.6.1 The $\mathcal{T}$-series, $\mathcal{M}$-series and $\mathcal{K}$-series Methods for the System with Axis-symmetric Bodies via Different Splitting and Composition Strategies

In this section, we utilize the splitting method to construct symplectic integrators. A diverse range of symplectic integrators with different accuracy and time complexities can be designed as the splitting method is quite flexible in terms of splitting and composition. Based on our Hamiltonian of the N-rigid-body system, we will explore three different types of integrators, namely the $\mathcal{T}$-series methods, the $\mathcal{M}$-series methods and the $\mathcal{K}$-series methods – the $\mathcal{T}$-series methods split the Hamiltonian into two parts with comparable size while the $\mathcal{M}$-series and the $\mathcal{K}$-series methods split the Hamiltonian into, respectively, four and three parts corresponding to various magnitudes and hence different timescales.

In terms of the shape of rigid bodies, we make the axis-symmetric assumption in this section for simplicity. That is, without loss of generality, $\boldsymbol{J}_i = \begin{bmatrix} J_i^{(1)} & 0 & 0 \\ 0 & J_i^{(1)} & 0 \\ 0 & 0 & J_i^{(3)} \end{bmatrix}$. For general rigid bodies that are not axis-symmetric, different splitting mechanisms can be applied (see section 3.6.3).

Classical Splitting for Rigid-Body ($\mathcal{T}$-series Methods): $H = H_1 + H_2$ *with* $\frac{H_1}{H_2} = \mathcal{O}(1)$

One way of splitting is $H = H_1 + H_2$ following [80], with

$$\begin{cases} H_1(\mathsf{q}, \mathsf{p}, \Pi, \mathsf{R}) = \sum_i \frac{1}{2} \boldsymbol{p}_i^T \boldsymbol{p}_i / m_i + \sum_i \frac{1}{2} \boldsymbol{\Pi}_i^T \boldsymbol{J}_i^{-1} \boldsymbol{\Pi}_i, \\ H_2(\mathsf{q}, \mathsf{p}, \Pi, \mathsf{R}) = V(\mathsf{q}, \mathsf{R}) = \sum_{i \neq j} V_{ij} \left(\boldsymbol{q}_i, \boldsymbol{q}_j, \boldsymbol{R}_i, \boldsymbol{R}_j \right). \end{cases} \tag{3.30}$$

For H_1, the equations of motion are

$$\begin{cases} \dot{\boldsymbol{q}}_i = \frac{\boldsymbol{p}_i}{m_i}, \\ \dot{\boldsymbol{p}}_i = 0, \\ \dot{\boldsymbol{R}}_i = \boldsymbol{R}_i \widehat{\boldsymbol{J}_i^{-1} \boldsymbol{\Pi}_i}, \\ \dot{\boldsymbol{\Pi}}_i = \boldsymbol{\Pi}_i \times \boldsymbol{J}_i^{-1} \boldsymbol{\Pi}_i. \end{cases} \tag{3.31}$$

In Eq. (3.31), the 4th equation is the Euler equation for a free rigid body. It is exactly solvable, and the solution expression is particularly simple for axial-symmetric bodies:

$$\boldsymbol{\Pi}_i(t) = \exp\left(-\theta t \widehat{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}} \right) \boldsymbol{\Pi}_i(0) := \mathrm{R}_z^T(\theta t) \boldsymbol{\Pi}_i(0) \tag{3.32}$$

with $\theta = \left(\frac{1}{J_i^{(3)}} - \frac{1}{J_i^{(1)}} \right) \boldsymbol{\Pi}_i^T(0) \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ and R_z being the rotation matrix. Take $\boldsymbol{\Pi}_i(t)$ back to the 3rd equation of Eq. (3.31), we can obtain $\boldsymbol{R}_i(t)$ too.

Therefore, the flow $\phi_t^{[1]}$ of H_1 is,

$$\begin{cases} \boldsymbol{q}_i(t) = \boldsymbol{q}_i(0) + \dfrac{\boldsymbol{p}_i}{m_i} t, \\ \boldsymbol{p}_i(t) = \boldsymbol{p}_i(0), \\ \boldsymbol{R}_i(t) = \boldsymbol{R}_i(0) \, \mathrm{R}_{\boldsymbol{\Pi}(0)} \left(\dfrac{\|\boldsymbol{\Pi}(0)\|}{J_i^{(1)}} t \right) \mathrm{R}_z \left(\theta t\right), \\ \boldsymbol{\Pi}_i(t) = \mathrm{R}_z^T \left(\theta t\right) \, \boldsymbol{\Pi}(0), \end{cases} \tag{3.33}$$

with R_z and $\mathrm{R}_{\boldsymbol{\Pi}(0)}$ being rotation matrices representing the rotations around the z-axis and $\boldsymbol{\Pi}(0)$ respectively.

For H_2, the equations of motion are

$$\begin{cases} \dot{\boldsymbol{q}}_i = 0, \\ \dot{\boldsymbol{p}}_i = -\dfrac{\partial V}{\partial \boldsymbol{q}_i}, \\ \dot{\boldsymbol{R}}_i = 0, \\ \dot{\boldsymbol{\Pi}}_i = -\left(\boldsymbol{R}_i^T \dfrac{\partial V}{\partial \boldsymbol{R}_i} - \left(\dfrac{\partial V}{\partial \boldsymbol{R}_i} \right)^T \boldsymbol{R}_i \right)^\vee. \end{cases} \tag{3.34}$$

As $\boldsymbol{q}_i$ and $\boldsymbol{R}_i$ stay constants, we have $\boldsymbol{p}_i$ and $\boldsymbol{\Pi}_i$ changing at constant rates. Therefore, the flow $\phi_t^{[2]}$ for H_2 is given by

$$\begin{cases} \boldsymbol{q}_i(t) = \boldsymbol{q}_i(0), \\ \boldsymbol{p}_i(t) = \boldsymbol{p}_i(0) - \dfrac{\partial V}{\partial \boldsymbol{q}_i} t, \\ \boldsymbol{R}_i(t) = \boldsymbol{R}_i(0), \\ \boldsymbol{\Pi}_i(t) = \boldsymbol{\Pi}_i(0) - \left(\boldsymbol{R}_i^T \dfrac{\partial V}{\partial \boldsymbol{R}_i} - \left(\dfrac{\partial V}{\partial \boldsymbol{R}_i} \right)^T \boldsymbol{R}_i \right)^\vee t. \end{cases} \tag{3.35}$$

We may compose $\phi_t^{[1]}$ and $\phi_t^{[2]}$ via $\mathcal{C}$ to construct different symplectic integrators [see e.g., 86] (see Fig. 3.2). To name a few, set $\mathcal{C}$ as $\mathcal{C}_{\text{Euler}}$, $\phi_h = \phi_h^{[1]} \circ \phi_h^{[2]}$ is a 1st order scheme

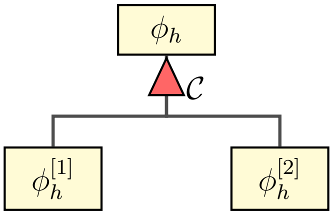


Figure 3.2: Composition of $\phi_h^{[1]}$ and $\phi_h^{[2]}$. The root node represent the final scheme. The leaves represent the basic ingredients of the composition which are exact flows. The red arrow represent the composition method specialized in composing two child flows with comparable scales.

with h being the step size (see appendix section 3.5 for $\mathcal{C}_{\text{Euler}}$ and the following composition methods $\mathcal{C}_{\text{Verlet}}$ and $\mathcal{C}_{S6}$).

Applying symmetric composition $\mathcal{C}_{\text{Verlet}}$, a 2nd order integrator $\mathcal{T}_2$ is in the form of

$$\phi_h^{\text{Verlet}} := \mathcal{C}_{\text{Verlet}}\left(\phi_h^{[1]}, \phi_h^{[2]}\right) = \phi_{\frac{h}{2}}^{[1]} \circ \phi_h^{[2]} \circ \phi_{\frac{h}{2}}^{[1]} \tag{3.36}$$

Applying ϕ_h^{TriJump} [94], we have the following 4th-order scheme $\mathcal{T}_4$,

$$\phi_{\gamma_1 h}^{\text{Verlet}} \circ \phi_{\gamma_2 h}^{\text{Verlet}} \circ \phi_{\gamma_3 h}^{\text{Verlet}}, \tag{3.37}$$

with $\gamma_1 = \gamma_3 = \frac{1}{2-2^{1/3}}$, $\gamma_2 = 1 - 2\gamma_1$. Similarly, a 6th-order scheme $\mathcal{T}_6$ can be constructed by composing $\phi_h^{[1]}$, $\phi_h^{[2]}$ with $\mathcal{C}_{S6}$

In the package, $\mathcal{T}_2$, $\mathcal{T}_4$ and $\mathcal{T}_6$ are implemented.

Tailored Splitting I ($\mathcal{M}$-series Methods): $H = H_1 + H_2 + H_3 + H_4$ *with* $\frac{H_1}{H_2} = \mathcal{O}(1)$ *and* $\frac{H_3}{H_1}, \frac{H_4}{H_2} = \mathcal{O}(\varepsilon)$

Different from point mass systems which can already exhibit dynamics over multiple timescales, the N-rigid-body system can have additional timescales created by the rotational dynamics.

Thus, we further split the Hamiltonian into more terms of different magnitudes, which produce flows at different timescales, and then carefully compose them[2]. More specifically,

---

[2] Similar techniques have already been employed; see e.g., [88] and references therein. The structure of

consider $H = H_1 + H_2 + H_3 + H_4$ with

$$\begin{cases} H_1(\mathsf{q},\mathsf{p}) = \sum_i \frac{1}{2}\boldsymbol{p}_i^T\boldsymbol{p}_i/m_i \\ H_2(\mathsf{q},\mathsf{p}) = -\sum_{i<j}\frac{\mathcal{G}m_im_j}{\|\boldsymbol{q}_i-\boldsymbol{q}_j\|}, \\ H_3(\Pi) = \sum_i \frac{1}{2}\boldsymbol{\Pi}_i^T\boldsymbol{J}_i^{-1}\boldsymbol{\Pi}_i, \\ H_4(\mathsf{q},\mathsf{R}) = V(\mathsf{q},\mathsf{R}) - H_2. \end{cases} \tag{3.38}$$

Here, H_1, H_2 have comparable size and $\frac{H_3}{H_1}, \frac{H_4}{H_2} = \mathcal{O}(\varepsilon)$ with ε being a small scaling parameter determined by the properties of the system. Based on scales of the dynamics, we denote $H^{\text{fast}} = H_1 + H_2$ and $H^{\text{slow}} = H_3 + H_4$. For example, consider the solar system, setting all the bodies to be point masses except the Earth, $\varepsilon \approx 10^{-6}$.

The flows $\left\{\varphi_t^{[1]}, \varphi_t^{[2]}, \varphi_t^{[3]}, \varphi_t^{[4]}\right\}$ of $\{H_1, H_2, H_3, H_4\}$ can be derived similarly to section 3.6.1 and the schemes are build by hierarchically composing $\left\{\varphi_t^{[i]}\right\}_{i=1}^4$ together. Specifically, as shown in Fig. 3.3, we firstly group the flows of the fast dynamics ($\varphi_t^{[1]}$ and $\varphi_t^{[2]}$) as a sub-scheme φ_h^{fast} via $\mathcal{C}_{\text{fast}}$ and the flows of the slow dynamics ($\varphi_t^{[3]}$ and $\varphi_t^{[4]}$) as a sub-scheme φ_h^{slow} via $\mathcal{C}_{\text{slow}}$ respectively. Then composing φ_h^{fast} and φ_h^{slow} together as the final scheme φ_h^{multi} via $\mathcal{C}_{\text{multi}}$. $\mathcal{C}_{\text{fast}}$ and $\mathcal{C}_{\text{slow}}$ are composition methods of composing two Hamiltonian flows with comparable scales [86]. $\mathcal{C}_{\text{multi}}$ is a composition method specialized in perturbative Hamiltonian systems of the form $H = A + \varepsilon B$ [97, 96, 88]. Note that the flows φ_h^{fast}, φ_h^{slow} are not exact, so the order of φ_h^{multi} is not the same as the order of $\mathcal{C}_{\text{multi}}$ applied for exact flows. In fact, the global error of φ_h^{multi} is the summation of the global errors of all three methods $\mathcal{C}_{\text{fast}}$, $\mathcal{C}_{\text{slow}}$ and $\mathcal{C}_{\text{multi}}$ (see appendix section B.3 for proof).

For example, if we set $\mathcal{C}_{\text{fast}}$, $\mathcal{C}_{\text{slow}}$ and $\mathcal{C}_{\text{multi}}$ as $\mathcal{C}_{S6}$, $\mathcal{C}_{\text{Verlet}}$ and $\mathcal{C}_{ABA42}$ (see appendix section 3.5) respectively. The global error of the above method is $\mathcal{O}(h^6) + \mathcal{O}(\varepsilon^2 h^2) + \mathcal{O}(\varepsilon h^4 + \varepsilon^2 h^2)$, i.e. $\mathcal{O}\left(h^6 + \varepsilon h^4 + \varepsilon^2 h^2\right)$. We name it $\mathcal{M}_{642}$ scheme with $6, 4, 2$ representing

---

our system, however, is new (due to the rigid-body part) and thus so is our specific splitting.

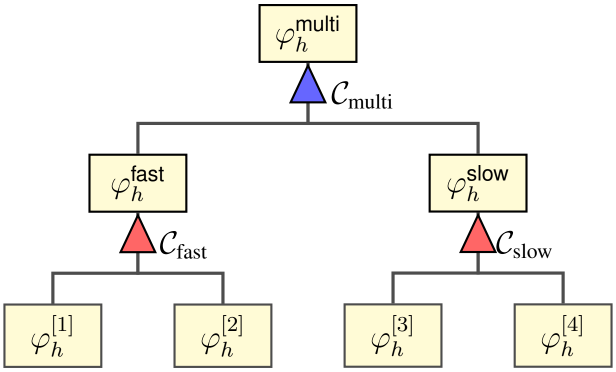


Figure 3.3: Hierarchical composition tree. The root node represents the final scheme. The leaves represent the basic ingredients of the composition which are exact flows. Nodes in the middle represent the intermediate composition flows. Red arrows represent the composition methods specialized in composing two child flows with similar scales. The blue arrow represent the composition methods specialized in composing two child flows with different scales.

Table 3.3: Comparisons of different schemes with respect to the number of dominating expensive stages and the global error order. The number of expensive stages are counted in an isolated step without considering the concatenation of the last stage with the first stage of the next step. The notation in the 'order' column is explained in the main text.

scheme	expensive stages	order
$\mathcal{T}_2$	3	(2)
$\mathcal{T}_4$	7	(4)
$\mathcal{T}_6$	15	(6)
$\mathcal{M}_{42}$	3	$(4,2)$
$\mathcal{M}_{642}$	6	$(6,4,2)$

the power of h of each term in the order and $\mathcal{M}$ representing multiscale splitting.

Similarly, we design the $\mathcal{M}_{42}$ scheme by choosing $\mathcal{C}_{\text{fast}}$, $\mathcal{C}_{\text{slow}}$ and $\mathcal{C}_{\text{multi}}$ as $\mathcal{C}_{\text{TriJump}}$, $\mathcal{C}_{\text{Verlet}}$ and $\mathcal{C}_{ABA22}$ respectively with the global error being $\mathcal{O}(h^4+\varepsilon h^2)$.

Compared with schemes in section 3.6.1, tailored splitting is able to mixing the fast and slow flows flexibly, thus being able control the time complexity. In fact, $T(\phi_h^{[1]}) = T(\varphi_h^{[1]})+T(\varphi_h^{[3]})$, $T(\phi_h^{[2]}) = T(\varphi_h^{[2]})+T(\varphi_h^{[4]})$ with $T(\cdot)$ being the number of operations of the one-step forward flow and evolving $\varphi_h^{[3]}$, $\varphi_h^{[4]}$ are much more expensive than evolving $\varphi_h^{[1]}$, $\varphi_h^{[2]}$. Since $\varphi_h^{[3]}$ and $\varphi_h^{[4]}$ are (expensive) slow dynamics that can be evolved with less effort (e.g. larger step size, less stages) than fast dynamics when evolving together and tailored splitting makes it possible to control the number of expensive stages. To compare, the number of expensive stages and the global errors of all schemes mentioned (in section 3.6.1 and section 3.6.1) are listed in Table. 3.3. In Table. 3.3, the order index $(o_0, o_1, \ldots)$ represents the power of h in front of $\varepsilon^0, \varepsilon^1, \ldots$ (e.g. a scheme of order (o_0, o_1, o_2) has a global error of $\mathcal{O}(h^{o_1} + \varepsilon h^{o_2} + \varepsilon^2 h^{o_3})$).

Moreover, since the hierarchical composition is a general framework, one can easily extend the family of numerical schemes, such as to construct higher order schemes, by applying a variety of existing splitting and composition methods.

Tailored Splitting II ($\mathcal{K}$-series Methods): $H = K_1 + K_2 + K_3$ with $\frac{K_3}{K_1}, \frac{K_2}{K_1} = \mathcal{O}(\varepsilon_{_K})$

We also provide an option to use the popular Wisdom-Holman [98] scheme for the orbital part, which works well for the specific but common setup of near Keplerian orbits; such systems usually correspond to $N-1$ well-separated bodies orbiting around a massive central body (indexed by 1 in our following description). This method is similar to the approach by [71], except that their coordinates are set using the body-frame and we provided a higher-order implementation.

Isolating the Keplerian dynamics as K_1, combining the rotational kinetic energy with the rest translational kinetic energy as K_2, and putting the rest potential energy to K_3, $H = K_1 + K_2 + K_3$ with

$$\begin{cases} K_1(\mathsf{q}, \mathsf{p}) = H_{Kepler}(\mathsf{Q}, \mathsf{P}) = \displaystyle\sum_{i=2}^{N} \frac{1}{2} \boldsymbol{P}_i^T \boldsymbol{P}_i / m_i - \frac{\mathcal{G} m_1 m_i}{\|\boldsymbol{Q}_i\|}, \\ K_2(\mathsf{p}, \mathsf{\Pi}) = \displaystyle\sum_{i=1}^{N} \frac{1}{2} \boldsymbol{\Pi}_i^T \boldsymbol{J}_i^{-1} \boldsymbol{\Pi}_i + \frac{\left\| \boldsymbol{p}_1 - \frac{m_1}{m_{tot}} \sum_{i=1}^{N} \boldsymbol{p}_i \right\|^2}{2 m_1}, \\ K_3(\mathsf{q}, \mathsf{R}) = V(\mathsf{q}, \mathsf{R}) + \displaystyle\sum_{i=2}^{N} \frac{\mathcal{G} m_1 m_i}{\|\boldsymbol{q}_i - \boldsymbol{q}_1\|}, \end{cases} \tag{3.39}$$

and $K_2, K_3 \ll K_1$. Here, $V(\mathsf{q}, \mathsf{R})$ is defined in Eq. (3.7). Note that K_1 represents Keplerian orbits in Q, P variables, which are canonical democratic heliocentric variables [99] with

$$\boldsymbol{Q}_i = \begin{cases} \boldsymbol{q}_i - \boldsymbol{q}_1 & i \neq 1, \\ \dfrac{\sum_{j=1}^{N} m_j \boldsymbol{q}_j}{m_{tot}} & i = 1, \end{cases} \tag{3.40}$$

and

$$\boldsymbol{P}_i = \begin{cases} \boldsymbol{p}_i - \dfrac{m_i}{m_{tot}} \displaystyle\sum_{j=1}^{N} \boldsymbol{p}_i & i \neq 1, \\ \displaystyle\sum_{j=1}^{N} \boldsymbol{p}_i & i = 1. \end{cases} \tag{3.41}$$

So when evolving K_1 dynamics, additional steps of switching back and force between (q, p) and (Q, P) coordinates are necessary. In terms of compositions, similarly, we first compose the flows of K_2 and K_3 together as $\varphi_h^{K,slow}$ via $\mathcal{C}_{\text{slow}}^K$, then compose the flow of K_1 ($\varphi_h^{K,\text{fast}}$) with $\varphi_h^{K,\text{slow}}$ via a multiscale compositing method $\mathcal{C}_{\text{multi}}^K$. The error of such composition is the summation of the global errors of two methods $\mathcal{C}_{\text{slow}}^K, \mathcal{C}_{\text{multi}}^K$ (and the numerical error of evolving Keplerian orbits).

For instance, $\mathcal{K}_{\cdot 2}$ method in our package is based on choosing $\mathcal{C}_{\text{slow}}^K, \mathcal{C}_{\text{multi}}^K$ as $\mathcal{C}_{\text{Verlet}}$ and $\mathcal{C}_{ABA22}$, and its global error is $\mathcal{O}(\varepsilon_{_K} h^2)$.

Which One to Use, the $\mathcal{T}$-series, the $\mathcal{M}$-series, or the $\mathcal{K}$-series Methods?

In general, the orders of the $\mathcal{T}$-series methods are only h dependent, while the $\mathcal{M}$-series and $\mathcal{K}$-series methods are (h, ε) dependent and $(h, \varepsilon_{_K})$ dependent respectively. Here, ε and $\varepsilon_{_K}$ are system specific, and they affect the choice of method. For example, $\varepsilon \approx 10^{-6}$ and $\varepsilon_{_K} \approx 10^{-3}$ in Solar system simulations with Earth being the only rigid body – note that $\varepsilon_{_K}$ represents the scale of the orbital planetary interactions while the ε in section 3.6.1 represents the scale of the spin and the potential correction due to rigidity, so in practice, $\varepsilon \ll \varepsilon_{_K}$. With the small parameters incorporated, the tailored splitting methods are usually more efficient. In general, the $\mathcal{K}$-series methods specialize in near-Keplerian problems, while the $\mathcal{M}$-series methods are more generic and at the same time almost always faster than the $\mathcal{T}$-series methods with nearly no trade-offs of the accuracy; in fact, oftentimes the $\mathcal{M}$-series methods are both more accurate and more efficient due to delicate splittings and compositions[3]. However, $\mathcal{T}$-series methods are recommended for extreme cases with large ε and $\varepsilon_{_K}$ (e.g., a super fast spinning body might contribute to a large ε).

3.6.2 Adding Non-Conservative Forces

Non-conservative forces such as tidal forces and post Newtonian corrections are incorporated in the package. As the implemented schemes are based on symmetric splitting and composition the corresponding non-conservative momentum update is inserted in the middle of the composition. This is similar to how dissipative forces were added in `REBOUNDx` [100].

Tidal Forces

We model the tidal dissipation between each pair of bodies using the constant time lag equilibrium tide model, following [66, 67]. Note that we only adopted the dissipative component in the tidal force here. The expression of the acceleration of the tidal force is

$$\begin{aligned}\boldsymbol{a}^{tidal}_{\text{host,guest}} &= -\frac{9\sigma m^2_{\text{guest}}A^2}{2\mu_{\text{host,guest}}d^{10}}\Big[3\boldsymbol{d}\left(\boldsymbol{d}\cdot\dot{\boldsymbol{d}}\right) \\ &+\left(\left(\boldsymbol{d}\times\dot{\boldsymbol{d}}\right)-\boldsymbol{\omega}d^2\right)\times\boldsymbol{d}\Big].\end{aligned} \tag{3.42}$$

Here, $m_{\text{host}}, m_{\text{guest}}$ denote the masses of the host and the guest body respectively; $\boldsymbol{d} = \boldsymbol{q}_{\text{guest}} - \boldsymbol{q}_{\text{host}}$ denotes the relative position of the guest body; $d = \|\boldsymbol{d}\|$ denotes the distance between two bodies; $\mu_{\text{host,guest}} = \frac{m_{\text{host}}\cdot m_{\text{guest}}}{m_{\text{host}}+m_{\text{guest}}}$ denotes the reduced mass; $\boldsymbol{\omega}$ denotes the

---

[3] One should not be misled to think an error like $\mathcal{O}(h^4+\varepsilon h^2)$ is larger than $\mathcal{O}(h^4)$; for example, if $\varepsilon = h^2$, the former may actually be smaller due to different constant factors; see section 3.7.1 for practical illustrations.

angular velocity of the host body under the reference frame (the inertia frame); the constant σ denotes the dissipation rate; A is defined as

$$A \ = \ \frac{d^5 Q}{1-Q}, \tag{3.43}$$

with Q the constant that measures quadrupolar deformability of the objects.

The dissipation rate σ is related to the time lag τ by the following formula,

$$\tau \ = \ \frac{3\sigma d^5}{4\mathcal{G}} \cdot \frac{Q}{1-Q}. \tag{3.44}$$

We may integrate the tidal acceleration $\boldsymbol{a}_{i,j}^{tidal}$ to our integrator after each time step by considering all pairs of bodies under tidal interactions. Note that each $\boldsymbol{a}_{i,j}^{tidal}$ only calculates the force of each (host, guest) pair, where each pair (i, j) treat i as the extended object and j as the point mass object. Thus, the equations of motion due to tidal dissipation are listed below:

$$\begin{cases} \boldsymbol{p}_i \ = \ \boldsymbol{p}_i + h \sum\limits_{j \neq i} \left(-\mu_{i,j} \, \boldsymbol{a}_{i,j}^{tidal} + \mu_{j,i} \, \boldsymbol{a}_{j,i}^{tidal} \right), \\ \boldsymbol{\Pi}_i \ = \ \boldsymbol{\Pi}_i - h \sum\limits_{j \neq i} \mu_{i,j} \, \boldsymbol{R}_i^T \left((\boldsymbol{q}_j - \boldsymbol{q}_i) \times \boldsymbol{a}_{i,j}^{tidal} \right). \end{cases} \tag{3.45}$$

General Relativistic Effects

We added the first-order post-Newtonian correction for general relativistic effects following e.g., [101]. For planetary systems, we assumed the central object (the host star) is much more massive comparing to the surrounding objects (the planets). Thus, we only included the correction due to the star. The acceleration can be expressed as the following [e.g., 102, 103]:

$$\mathbf{a} = \frac{GM_{star}}{r^3 c^2} \Big[\Big(\frac{4GM_{star}}{r} - \mathbf{v}^2 \Big) \mathbf{r} + 4(\mathbf{v} \cdot \mathbf{r}) \mathbf{v} \Big] \tag{3.46}$$

3.6.3 Asymmetric Case

For planets with close-in orbits, both rotational flattening and tidal force distort the shape of the planets, and lead to non-axial symmetric distortions. Thus, we include the option to study non-axial symmetric planets here, where one could specify the principal moment of inertia or the semi-axes of the planets directly. In this case, $J_i^{(1)} \neq J_i^{(2)} \neq J_i^{(3)}$ in $\boldsymbol{J}_i$, and our splitting of the Hamiltonian is modified as the previous Hamiltonian plus $H_{asymmetric}$, where

$$H_{asymmetric}(\mathsf{R}, \mathsf{\Pi}) = \sum_i \left(\frac{1}{J_i^{(2)}} - \frac{1}{J_i^{(1)}} \right) \cdot \frac{{\Pi_i^{(2)}}^2}{2}, \tag{3.47}$$

and $H_{asymmetric} \ll H_3$ in Eq. (3.38).

The dynamics of Eq. (3.47) is

$$\begin{cases} \boldsymbol{R}_i(t) = \mathrm{R}_y \left(\delta \Pi_i^{(2)} t \right) \boldsymbol{R}_i(0), \\ \boldsymbol{\Pi}_i(t) = \mathrm{R}_y(-\delta \Pi_i^{(2)} t)\, \boldsymbol{\Pi}_i(0), \end{cases} \tag{3.48}$$

with $\delta = \frac{1}{J_i^{(2)}} - \frac{1}{J_i^{(1)}}$. Based on the symmetric schemes in section 3.6.1, we simply evolve $H_{asymmetric}$ half step at the beginning and the end of each step.

3.7 Code Validation

3.7.1 Numerical Tests

Conservation Properties

The conservation properties of the integrators are tested for $\mathcal{T}_4$ and $\mathcal{M}_{42}$ schemes in the *Sun-Earth-Moon* system with all three bodies being rigid. As shown in Fig. 3.4, both schemes conserve linear momentum and angular momentum (except there are arithmetic inaccuracies due to machine precision), and the energies exhibit no drift but only fluctuate at magnitudes $\mathcal{O}(h^4)$ and $\mathcal{O}(h^4 + \varepsilon h^2)$ for $\mathcal{T}_4$ and $\mathcal{M}_{42}$ respectively. In the simulations, tides

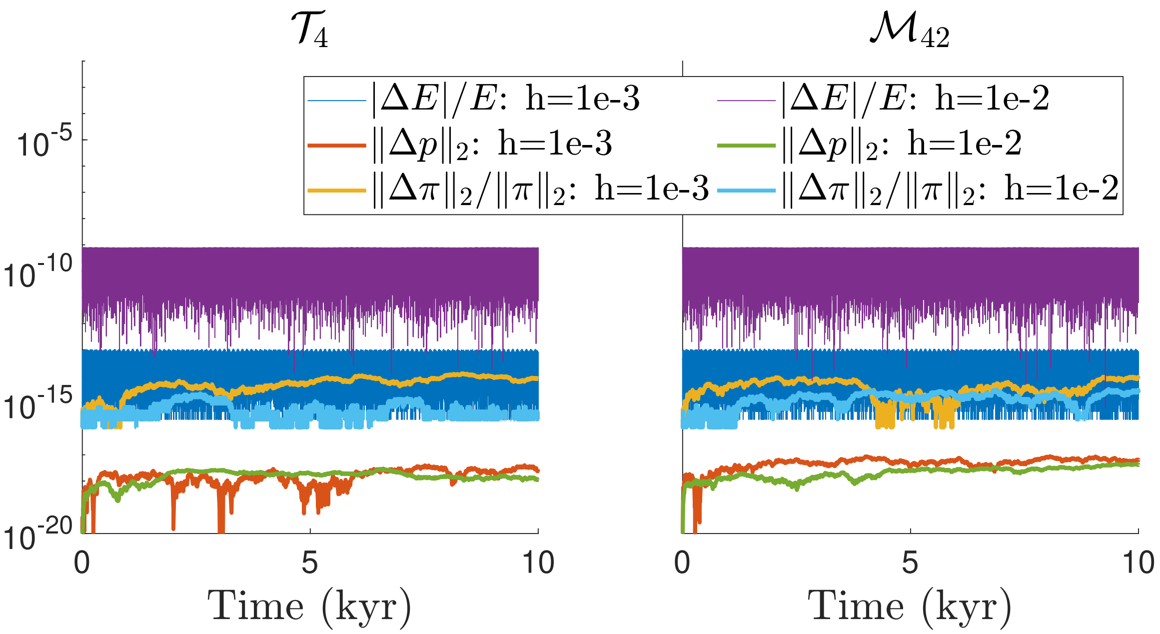


Figure 3.4: Conservation of momentum maps and near conservation of energy by our methods. Relative error of energy E, error of the total linear momentum $\boldsymbol{p}$ and relative error of the total angular momentum $\boldsymbol{\pi}$ are measured. $\boldsymbol{p}(0) = [0, 0, 0]$. The potential order is set to be 2.

are not included (otherwise the system is no longer conservative) and initial conditions are set to be the data of epoch J2000 from JPL HORIZONS System.

Here floating-point format is set to be double-precision, although our package can also use long-double or single.

Our integrators also (exactly) preserve symplecticity when tidal dissipation is excluded, because they are Hamiltonian splitting schemes. The definition of symplecticity in a non-Euclidean setup is not completely trivial, but the symplecticity of splitting approaches considered here has been established in, e.g., [104] (with $r(t) = 0$; otherwise one gets a more general result, namely conformal symplecticity).

Convergence Tests and Accuracy Comparisons

We now numerically illustrate how the integration error depends on h for different numerical schemes, which include both methods we implemented in `GRIT` and `SMERCURY-T`. `SMERCURY-T` is a concurrent simulation package that can evolve an object's spin-axis under obliquity tide [105]. It is based on the `Mercury` simulation package [106]. Specifically, it includes a subroutine to evolve the spin-axis dynamics following the procedure outlined in [6], which is based on the Lie-Poisson integrator of rigid-body dynamics developed by [71]. In addition, it includes a subroutine for obliquity tide following the algorithms outlined in [69]. The model for tidal interaction of `SMERCURY-T` is different from what we included in `GRIT`, which natually contains both obliquity tide and tidal effects due to non-tidally synchronized orbits. Thus, we focus on the rigid-body dynamics here, where we do not include tidal interactions in our convergence test. We also turned off, in comparisons presented here, our rigid-body rigid-body interaction option, which is mainly for accurate simulations of rigid bodies' close encounters, because such interactions are supported only in `GRIT`.

We first test on the *Sun-Earth-Moon* system (Fig. 3.5). One observation in this case is, if the step size is too large so that splitting into $H_1 + H_2 + H_3 + H_4$ (`GRIT`'s $\mathcal{M}_{42}$, $\mathcal{M}_{642}$) doesn't work, `SMERCURY-T` doesn't work either (unlike expected by some). More precisely, with $h = 2 \cdot 10^{-2}$ yr, $\mathcal{M}_{42}$ and `SMERCURY-T` cannot resolve the the motion of the Moon orbiting around the Earth, whose period is a month, and even the performance of the 6th order method $\mathcal{M}_{642}$ is not ideal, and significant errors are observed in all methods. Accuracy is improved for stepsizes below this stability limit, and the rate of improvement is, as expected, dependent on the order of the numerical scheme. Consequently, higher order methods such as $\mathcal{M}_{42}$ and $\mathcal{M}_{642}$ show substantially smaller errors when smaller step sizes are applied (readers interested in understanding this together with computational costs are referred to section 3.7.1).

We then test on a non-Keplerian system (note `SMERCURY-T` performs well for near Keplerian problems as designed): an Earth-like planet orbiting around two stars alternatively in a stellar binary system (Fig. 3.6). As there is no single body that has the dominant mass of the system and the the planet is alternatively captured by the two stars, the planetary orbit is not nearly Keplerian, and splitting into $H_1 + H_2 + H_3 + H_4$ is more accurate

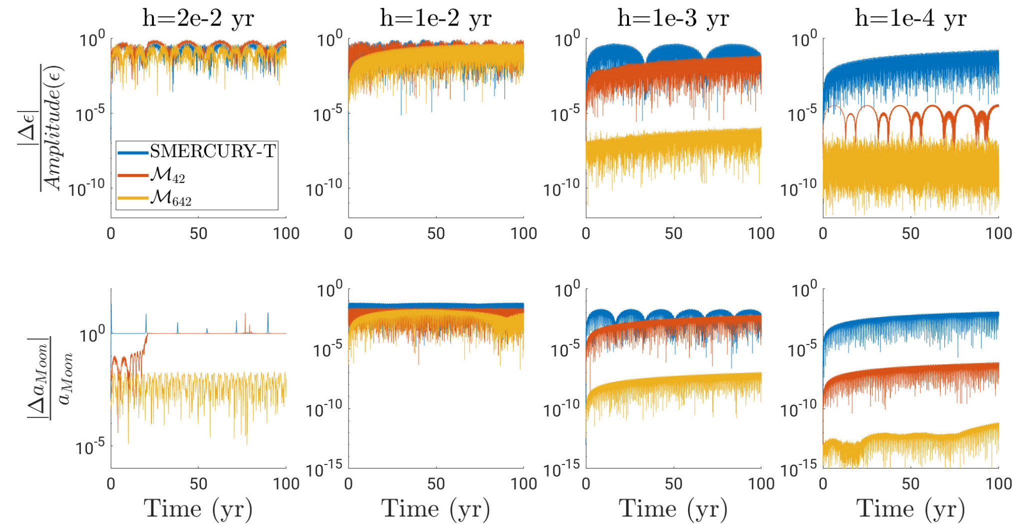


Figure 3.5: Error of Earth's obliquity (ϵ) over the range of ϵ's fluctuation and the relative error of the semi-major axis of the Moon for the *Sun-Earth-Moon* system. Earth, Sun, Moon are rigid body, point mass and point mass respectively. The benchmark is simulated using the $\mathcal{T}_6$ scheme with $h = 10^{-5}$ yr.

than `SMERCURY-T` for all choices of step sizes here. Specifically, as shown in Fig. 3.6, the orbital position of `SMERCURY-T` saturates to $\mathcal{O}(1)$ relative error after a relatively short period of time, no matter if $h = 10^{-3}, 10^{-4}$, or 10^{-5} yr. The orbital inaccuracy naturally affects the spin angle as well. Meanwhile, $\mathcal{M}_{42}$ and $\mathcal{M}_{642}$ do not have this issue.

For reproducibility, the initial condition used is

$$\begin{aligned}
\boldsymbol{q}_{star_1} &= \begin{bmatrix} -0.5 & 0 & 0 \end{bmatrix}^T, & \boldsymbol{v}_{star_1} &= \begin{bmatrix} 0 & -0.0086012119 & 0 \end{bmatrix}^T, \\
\boldsymbol{q}_{star_2} &= \begin{bmatrix} 0.5 & 0 & 0 \end{bmatrix}^T, & \boldsymbol{v}_{star_2} &= \begin{bmatrix} 0 & 0.0086012119 & 0 \end{bmatrix}^T, \\
\boldsymbol{q}_{planet} &= \begin{bmatrix} 1.16 & 0 & 0 \end{bmatrix}^T, & \boldsymbol{v}_{planet} &= \begin{bmatrix} 0 & 0.0164271047 & 0 \end{bmatrix}^T
\end{aligned}$$

in units of AU and AU/day, and $m_{star_1} = m_{star_2} = 0.5 m_{\odot}$, $m_{planet} = m_{\oplus}$.

Investigation of Efficiency

We now demonstrate the improved computational efficiency of the tailored splitting schemes. A comparison of the time efficiency among the traditional splitting method $\mathcal{T}_4$, $\mathcal{T}_6$ and the

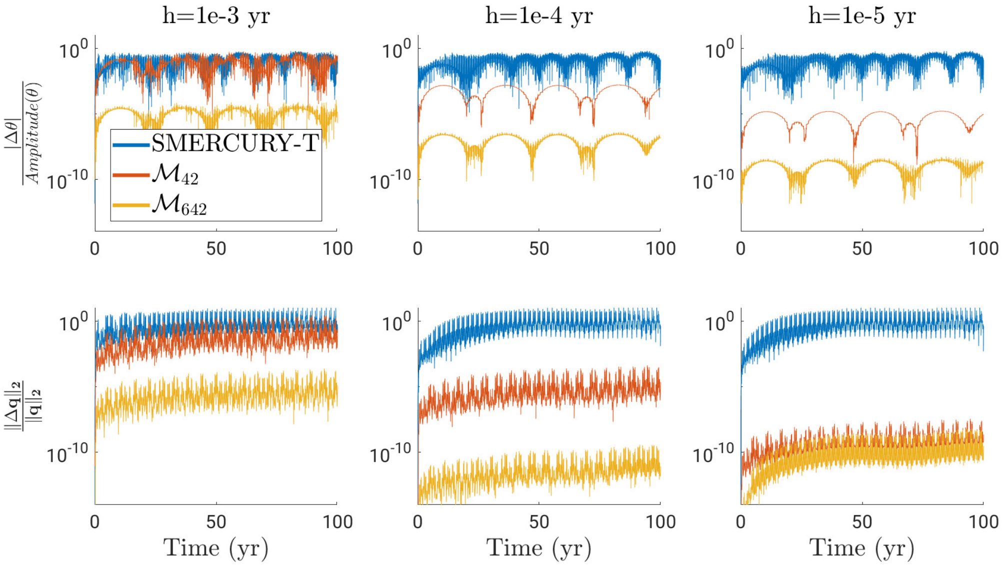


Figure 3.6: Error of spin angle (the angle between the angular momentum and the z-axis of the inertia frame) and position for an Earth-like planet orbiting around two stars alternatively. The Earth-like planet, star 1 and star 2 are set to be rigid body, point mass and point mass respectively. The benchmark is simulated using the $\mathcal{T}_6$ scheme with h=1e-5.

tailored splitting scheme $\mathcal{M}_{42}, \mathcal{M}_{642}, \mathcal{K}_{\cdot 2}$ in the 10 rigid body (Sun with 8 planets and the Moon) is shown in Table. 3.4. $\mathcal{M}_{42}$ ($\mathcal{M}_{642}$) is about twice the speed of $\mathcal{T}_4$ ($\mathcal{T}_6$) with comparable integrating accuracy. Note `SMERCURY-T` cannot be compared against here, because its currently available version[4] can only set one of the objects as rigid-body.

To gain additional understanding of the performance of `GRIT`, complementary results that include comparisons to `SMERCURY-T` are also provided. For a fair comparison, we continue using the Solar system example, which is a near Keplerian problem that `SMERCURY-T` specializes in, but we had to alter it by setting only the Earth to be a rigid body and all others as point masses. Results are in Table. 3.5, where $\mathcal{M}_{42}$ shows improved accuracy over `SMERCURY-T`, while $\mathcal{M}_{642}$ is even more accurate however with traded-off time complexity.

Also for the sake of fairness, note that wall-clock counts are platform dependent and therefore should only be used as a qualitative (not quantitative) indicator. Experiments reported here are conducted on a machine with AMD Ryzen 7 3700X 8-Core Processor, 16

---

[4] https://github.com/SMKreyche/SMERCURY-T/tree/cbc25299825559f255cee096c7650f379af41aa5

Table 3.4: Efficiency comparison among scheme $\mathcal{T}_4$, $\mathcal{T}_6$, $\mathcal{M}_{42}$, $\mathcal{M}_{642}$ and $\mathcal{K}_{\cdot 2}$. The Solar system with 8 planets and the Moon (10 rigid bodies in total) is simulated till 1000 years with $h = 10^{-3}$ yr and $h = 10^{-4}$ yr for all schemes using a single thread. The benchmark is simulated using the $\mathcal{T}_6$ scheme with $h = 10^{-5}$ yr and long-double precision. Mean absolute errors (MAE) of the Earth's obliquity (rad) are measured. Data is output every 0.1 yr.

$h = 10^{-3}$ yr	Wall time (s)	MAE of Earth's Obliquity
$\mathcal{T}_4$	30.573	1.996646e-05
$\mathcal{M}_{42}$	15.488	1.997454e-05
$\mathcal{T}_6$	72.55	1.728156e-08
$\mathcal{M}_{642}$	40.626	4.365093e-10
$\mathcal{K}_{\cdot 2}$	14.673	2.186379e-05
SMERCURY-T	N/A	N/A
$h = 10^{-4}$ yr		
$\mathcal{T}_4$	273.71	2.680897e-09
$\mathcal{M}_{42}$	140.58	3.817091e-09
$\mathcal{T}_6$	708.26	8.689218e-11
$\mathcal{M}_{642}$	395.52	2.039980e-10
$\mathcal{K}_{\cdot 2}$	131.56	1.292609e-05
SMERCURY-T	N/A	N/A
$h = 10^{-5}$ yr		
$\mathcal{K}_{\cdot 2}$	1299.9	1.378428e-07

GB memory and the Linux distribution of openSUSE Leap 15.2. `GRIT` was compiled using GNU C++ compiler and `SMERCURY-T` using GNU Fortran compiler, both with the default compilation options. Single-thread is used for experiments in both Tables 3.4 and 3.5 for fairness (note a parallelization option is available in `GRIT`; we recommend turning it on when the simulated system has large numbers of rigid objects). We also noted that `SMERCURY-T` slows down more significantly than `GRIT` when its integration is outputted more frequently, and thus chose a large output step size to reduce `SMERCURY-T`'s I/O overhead so that the focus can be on the integration time itself.

Summary of Section 3.7.1 Numerical Tests

In general, `GRIT` suits not only near-Keplerian orbits but also non-Keplerian ones. Multiple splitting and composition options are provided in `GRIT` too so that, if preferred, a user can choose the classical Wisdom-Holman scheme for the orbital part which specializes in

Table 3.5: Efficiency comparison among scheme $\mathcal{M}_{42}$, $\mathcal{M}_{642}$ and SMERCURY-T. The Solar system with 8 planets and the Moon (9 point masses and 1 rigid body (the Earth) in total) is simulated till 1000 years with $h = 10^{-3}$ yr and $h = 10^{-4}$ yr for all schemes using a single thread. The benchmark is simulated using the $\mathcal{T}_6$ scheme with $h = 10^{-5}$ yr and long-double precision. Mean absolute errors (MAE) of the Earth's obliquity (rad) are measured. Data is output every 0.1 yr.

$h = 10^{-3}$ yr	Wall time (s)	MAE of Earth's Obliquity
$\mathcal{M}_{42}$	6.408	1.997119e-05
$\mathcal{M}_{642}$	23.122	3.841649e-10
`SMERCURY-T`	8.638	2.157662e-05
$h = 10^{-4}$ yr		
$\mathcal{M}_{42}$	53.782	3.833661e-09
$\mathcal{M}_{642}$	216.09	1.990336e-10
`SMERCURY-T`	39.079	1.903458e-05

near-Keplerian orbits (e.g., $\mathcal{K}_{\cdot 2}$). Furthermore, equipped with higher order methods, `GRIT` integrations have errors that decrease very rapidly as step size decreases in a reasonable range.

3.7.2 Comparison with Secular Results

To further verify the accuracy of our integration package, we compare our simulation results to secular theory here. We include two examples: the first one integrate the obliquity variation of a moon-less Earth without the influence of tidal interactions, and the second example considers tidal interactions between a hypothetical Earth-Moon system. We find good agreement between our simulation package with the results of the secular theory.

Obliquity Variations of a Moon-less Earth

Spin-orbit resonances lead to large obliquity variations for a moonless Earth [15], and this classical example can serve as a test case for our simulation package. Specifically, planetary companions of the Earth (from Mercury to Neptune) all perturb Earth's orbit and lead to forced oscillations in the orbital plane of Earth. At the same time, torquing from the Sun leads to precession of Earth's spin axis. The natural precession frequency coincides with

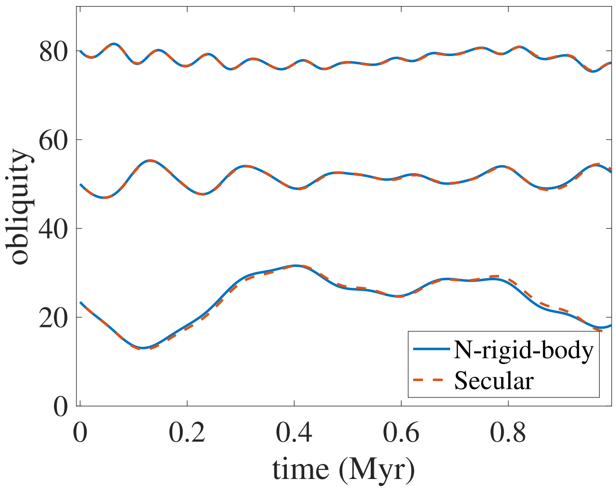


Figure 3.7: Obliquity variations of a moon-less Earth. The solid lines represent the rigid-body simulations, and the dashed lines represent the secular results following [5]. The results of our simulation package agree well with the secular theory.

the forcing frequencies and drives resonant obliquity variations of Earth. Tidal interactions are weak in this case, so we neglected tidal effects in our code and considers the dynamical coupling between the planetary spin axes and its orbit.

We include the eight Solar System planets in this system, and we adopt the position and velocity of the Solar System planets from JPL database [107]. We only treat the Earth as a rigid object with oblateness of 0.00335, and set the other planets and the Sun as point particles.

Figure Figure 3.7 shows the comparison of the obliquity variations of the moon-less Earth with that from the secular theory shown in [15, 43]. We included three examples starting with different initial obliquities, and all of them show good agreement with the secular results. In particular, below $\sim 40^\circ$, large obliquity variations can be seen due to the spin-orbit resonances. We chose a time step of 10^{-4}yrs, in order to resolve the spin of the Earth. The fractional change in energy is at the order of 10^{-14} and the fractional change in angular momentum is at the order of 10^{-12} for all the three runs with different initial obliquities.

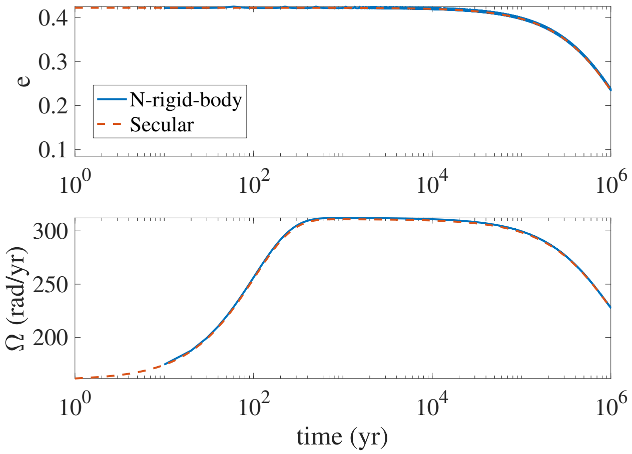


Figure 3.8: Tidal interaction in a hypothetical Earth-Moon system. The spin rate (Ω) increases rapidly to the pseudo-synchronized state, which is then followed by a much slower decay as the orbit circularizes under tide. The solid lines represent the simulation results and the dashed lines represent the secular results. The results of our simulation package agree well with the secular theory.

Tidal Interactions of a Hypothetical Earth-Moon System

To illustrate the accuracy of our simulation package including tidal interactions, we use a simple hypothetical Earth-Moon two-body system here. We set the initial semi-major axis and eccentricity to be 0.0018AU and 0.4. For the Earth, we set the spin period to be 1 day, oblateness to be 0.00335, love number to be 0.305 and tidal time lag to be 698sec. For the Moon, we set the spin period to be 14 days, oblateness to be 0.0012, love number to be 0.02416 and tidal time lag to be $8,639$sec.

Figure Figure 3.8 shows the agreement between our simulation package (solid lines) with the secular results (dashed lines). The secular results are obtained following [67]. The upper panel plots orbital eccentricity versus time and the lower panel plots the spin rate of the Moon versus time. It shows that the spin rate of the Moon increases to the pseudo-synchronized state within a few hundred years, and then slowly decreases as orbital eccentricity decays due to tide. We chose a time step of 10^{-4}yr to resolve the spin of the Earth, and the total fractional change in angular momentum is 7×10^{-12}.

3.8 Applications to Trappist-I

Spin-orbit coupling leads to profound dynamics in planetary systems, in particular for planets with close-in orbits. For Trappist-I, it is shown that tidal and rotational deformation of the planets leads to orbital precession that can be detected in the TTV measurements [63]. In addition, strong interactions between planets in resonant chains can push habitable zone Trappist-I planets into non-synchronous states [64].

To illustrate the effects of the spin-orbit coupling, we use our numerical package to simulate the long-term dynamics of spin-axis variations, as well as the short-term effects on TTV for Trappist-I. We note that both our numerical package and *POSIDONIUS* [70, 63] consider tidal effects and spin-orbit coupling, beyond point mass dynamics based on Newtonian interactions and GR corrections. In particular, [63] obtained both dissipative and non-dissipative forces from tidal dissipation and tidal torquing separately, and considered forcing due to planetary rotational deformation.

Using our numerical package, we find that the habitable zone planets can indeed allow large spin-state variations, consistent with the findings by [64]. In addition, we find that allowing the non-synchronized states could lead to significantly larger TTVs, which could reach a magnitude of $\sim min$ in ten-year timescale.

3.8.1 System set up

We use the same orbital initial condition and physical properties for the planets in Trappist-I following [63] (Table A.2 in [63]), in order to compare the magnitude of TTVs, and we use the same reference tidal parameters for the star and the planet (e.g., $k_{2f,*} = 0.307$, $k_{2f,p} = 0.9532$, $\Delta\tau_p = 712.37$sec). The Q coefficient in the tidal model can then be calculated ($k_2 = Q/(1-Q)$) [67, 108, 109].

To calculate the moment of inertia along the three principal axes (A, B, C), we follow the derivation by [110], assuming a homogeneous model for simplicity and assuming the rotation velocity of the planet is close to the orbital velocity. Specifically, the moment of inertia can be expressed as the following:

$$
\begin{aligned}
A &= I_0(1 - \frac{1}{3}\alpha - \frac{1}{2}\beta) \\
B &= I_0(1 - \frac{1}{3}\alpha + \frac{1}{2}\beta) \\
C &= I_0(1 + \frac{2}{3}\alpha)
\end{aligned}
$$

where

$$
\begin{aligned}
\alpha &= \frac{5}{4}q(1 + k_f) \\
\beta &= \frac{3}{2}q(1 + k_f)
\end{aligned}
$$

and k_f is the love number, and q is the ratio of the centrifugal acceleration to the gravitational acceleration. We assume all the planets have the same radius of gyration squared $rg_p^2 = 0.3308$ following [63], and we include in Table. 3.6 the moment of inertia of the planets.

Moreover, because the planets are very close to their host star, general relativistic precession plays a non-negligible role in the transit time. Thus, we also included the first order post-Newtonian correction in our simulation code (see §subsubsection 3.6.2).

Table 3.6: Principal moment of inertia of the Trappist-I planets.

Planet	A ($M_\odot km^2$)	B ($M_\odot km^2$)	C ($M_\odot km^2$)
b	50.5245	50.8474	50.955
c	54.3319	54.4432	54.4803
d	7.6321	7.6396	7.6421
e	26.384	26.391	26.3933
f	39.9836	39.9898	39.9918
g	58.8644	58.8698	58.8716
h	7.8901	7.8904	7.8905

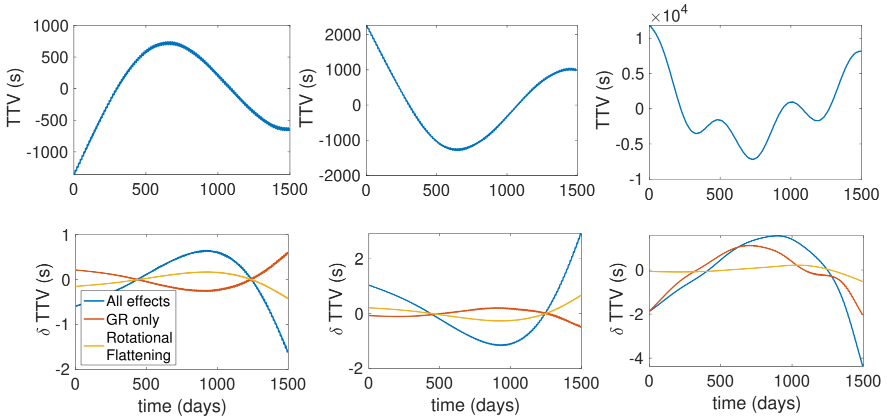


Figure 3.9: Transit-timing variations (TTVs) of planets (b, c, d) in Trappist-I. The upper panels show TTVs of the planets assuming they are point mass particles and neglect effects due to GR. The lower panels show the differences in TTVs due to GR, rotational flattening of the planets and all the effects (GR, rotational flattening, tidal precession and tidal dissipation combined). The differences due to GR and rotational flattening are consistent with the results in [63], while assuming the planets to be rigid bodies, the TTV differences are larger.

3.8.2 Transit-timing Variations

The measurement of transit-timing variations (TTVs) is a powerful method to derive physical properties of planets, in particular masses and eccentricity of planets [111]. Most studies consider only point-mass dynamics. However, full-body dynamics including tidal effects and distortion of the planets could also play an important role [112, 113, 114, 115]. It is recently shown that new measurements of the TTV of the Trappist-I system lead to significant increase in the mass estimate for planet b and c, which may be due to unaccounted physical processes including tidal effects and rotational distortion of the planets [116, 117, 63]. Thus, we use our simulation package to estimate the TTV of the inner planets in Trappist-I here as an example, in comparison with the study by [63].

We include the result of the transit-timing variations for Trappist-I b,c and d in 1500 days in Figure Figure 3.9, to compare our results with those in [63]. Similar to Figure 1 in [63], the upper panels show the transit timing variations assuming the planets are all point-mass particles, and the lower panels show the differences in the TTVs due to different

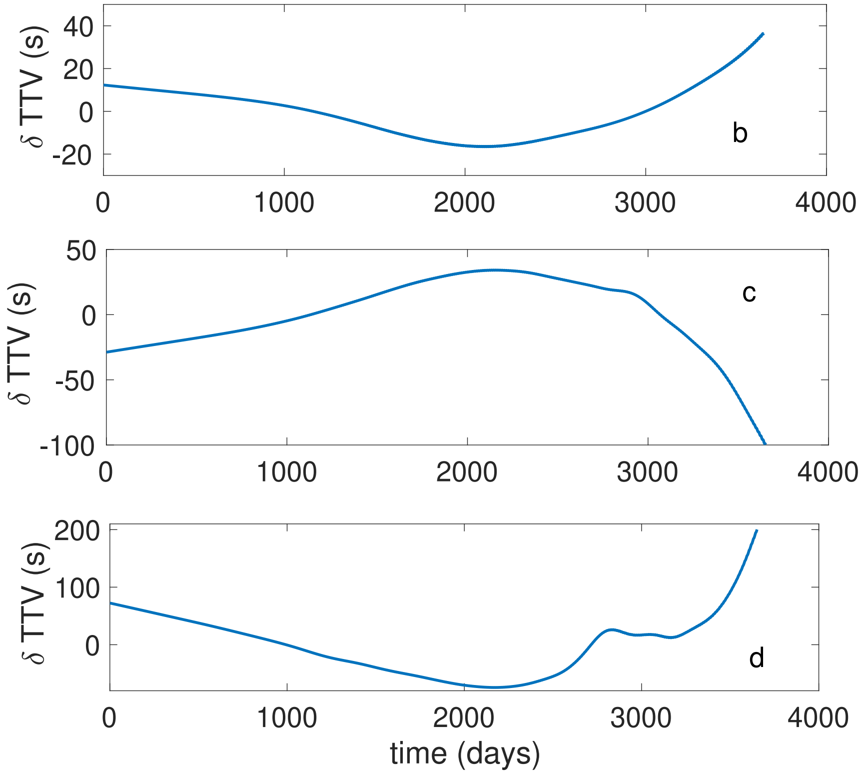


Figure 3.10: Differences in transit-timing variations (TTVs) of planets (b, c, d) in Trappist-I over ten year measurements. With larger spin-misalignment, transit-timing variations could reach $\sim$ mins.

effects. The differences due to GR and rotational flattening of the planets computed using our simulation package are agreeable with that in [63]. Different from [63], we assume the objects are rigid-bodies when considering tidal interactions with the central star using our rigid-body simulator. This leads to slightly larger TTV differences. We note that the magnitude of the differences in the TTVs depend on the misalignment between the elongated principal axis and the location direction of the planet from the central star. For the illustrative example, we assume the planets all start with their long-axes perfectly aligned to the direction of the central star.

As the system evolve further, the misalignment could be excited to larger values (as discussed further in section subsection 3.8.3). The differences in TTVs could reach ~ 5sec

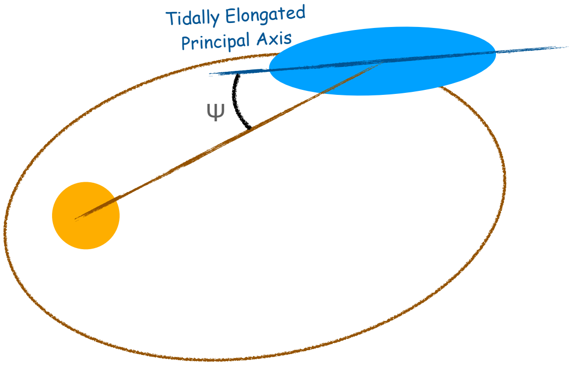


Figure 3.11: Illustration of the long-axis misalignment. Low variations in ψ correspond to a tidally locked planet.

for 1500 days, and a few minutes in 10 year measurements, shown in Figure Figure 3.10. A detailed study of how the TTVs depend on the physical properties of the planets (e.g., the love number, tidal time lag, etc.) is out of the scope of this paper, and will be discussed in a follow up project.

3.8.3 Long-term dynamics

Long-term dynamics of spin-axes of planets, in particular their synchronized states, play an important role in the atmosphere circulation of the planets. When the planets are tidally locked, the extreme temperature differences on one side of the planet facing the star from the other side may lead to the collapse of planetary atmosphere [118, 119, 120]. For Trappist-I, [64] developed a framework studying the spin-axis variations of the planets and found that the mean motion resonant chain could drive the habitable zone planets out of the synchronized state.

Specifically, [64] evolves the longitude of the substellar point separately based on results of orbital evolution of Trappist-I using the *Rebound* simulation package [121]. This does not include effects of the variation of the spin-axis on the orbits, and the developed framework neglected the 3-D variations of the planetary spin-axis (i.e., assuming zero plan-

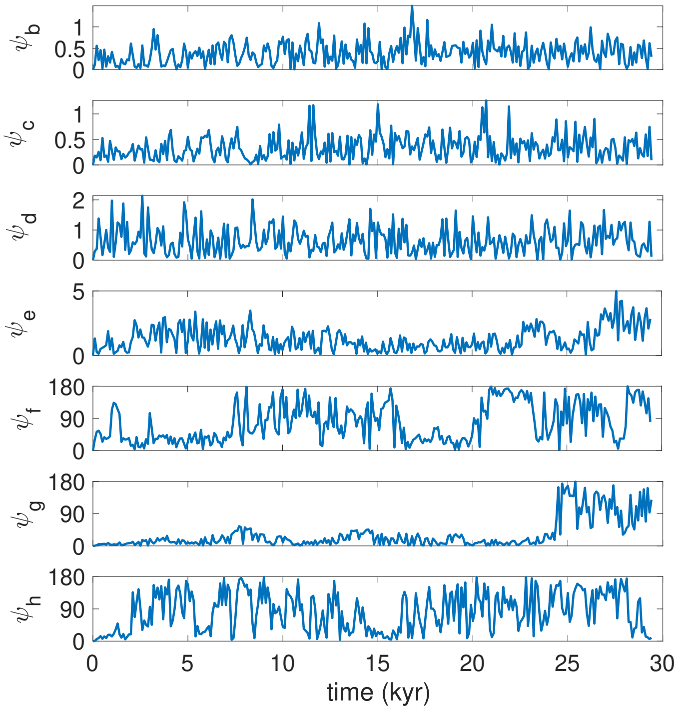


Figure 3.12: Spin-axis misalignment as a function of time. Planet f, g and h all have large long-axes variations, and are not tidally locked.

etary obliquities) for simplicity. To evaluate the spin-axis dynamics more accurately, we use our simulation package, which allows backreactions of the spin-axis dynamics on the orbit, as well as the full 3-D dynamics of the planetary spin-axis.

We use the same initial condition as those in section subsection 3.8.2 for the long-term dynamical simulation over $100,000$yrs. We start the planets in synchronized configurations and we calculate the misalignment between the long axes of the planets and their radial direction from the host star, which is ψ illustrated in Figure Figure 3.11.

Figure Figure 3.12 shows this misalignment (ψ) of the planets. Planet b, c, d and e are closer to the host star, and allow stronger tidal interactions. This leads to low variations in the long-axes of the planets. However, planet f, g and h are further away, where planetary interactions could compete with tidal re-alignment and drive larger spin-axis variations. We note that the obliquities of these planets still remain low (within a few degrees). The detailed dependence of the spin-axis variations on the parameters of the planets are beyond the scope of this article, and will be invested in a follow up paper.

3.9 Conclusions

In this article, we developed symplectic integrators and provided a package "`GRIT`" for studying the spin-orbit coupling of N-rigid-body systems. We split the Hamiltonian into four parts with different evolution timescales (tailored splitting), and compose the four parts together in a hierarchical way so that the expensive slow scale evolution is more efficient. In general, the tailored splitting is more flexible and efficient than the traditional splitting.

To illustrate the validity of the integrator, we showed that it provides results consistent with the secular theories for the obliquity variation of a moonless Earth, and the tidal evolution of a hypothetical Earth-Moon system. This allowed us to confidently apply it to the less well understood system Trappist-I, and show that the differences in transit-timing variations could reach a few seconds for a four year measurements, and planetary interactions could push planet f, g and h out of the synchronized states, which are consistent with [63] and [64].

We assume the objects are rigid bodies in our simulation package. This is a good approximation when the deformation of the objects are slow. Thus, our simulation package can be applied for objects with a slow change of rotation rate or tidal distortion. When the deformation rate is faster than the orbital variation timescales, spin-orbit coupling using hydrodynamical simulations could provide more accurate results [e.g., 122].

CHAPTER 4

DATA-DRIVEN PREDICTION OF GENERAL HAMILTONIAN DYNAMICS VIA LEARNING EXACTLY-SYMPLECTIC MAPS

4.1 Introduction

Given a collection of sequences, each being a multidimensional time series produced by the same latent mechanism, we consider learning this mechanism and predicting a sequence's future evolution. More precisely, suppose there is an unknown (possibly highly nonlinear) map ϕ that evolves any initial condition in discrete time i according to

$$x_{i+1} = \phi(x_i).$$

Provided with training data $\{x_{i,j}\}$, where each $j \in [M]$ corresponds to such a sequence ($i \in \{0, \dots, N_j\}$), we'd like to learn ϕ purely from the data, or more precisely an approximation of its image $\tilde{\phi}(x) \approx \phi(x)$ for any x in the problem domain. This way, one can for example perform continuation of existing sequences via $\tilde{x}_{i+1,j} = \tilde{\phi}(\tilde{x}_{i,j})$ for $i \geq N_j$ and $\tilde{x}_{N_j,j} = x_{N_j,j}$, or predict a sequence evolved from a new initial condition via $y_{i+1} = \tilde{\phi}(y_i)$ for $i > 0$.

This task of course appears in many contexts (see e.g., section 4.2). This article considers a very specific one, for which the latent map ϕ is assumed to be *symplectic*: for simplicity we will work with finite dim. vector spaces equipped with canonical symplectic structure, which means each x can be written as $x = [\boldsymbol{p}, \boldsymbol{q}]$ where $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{R}^d$, and the Jacobian of ϕ satisfies

$$(\phi')^T J \phi' = J,$$

where $J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$ is a $2d$-by-$2d$ matrix with 0 and I being d-by-d blocks.

The consideration of symplectic evolution maps is largely motivated by the learning and prediction of mechanical behaviors, which recently attracted significant attention (see section 4.2 3rd paragraph). More precisely, if the latent evolution mechanism is provided by a Hamiltonian mechanical system, each time series is given by a solution to the Hamiltonian ODE system sampled at discretized time points. That is, $x_i = [\boldsymbol{p}_i, \boldsymbol{q}_i]$, $\boldsymbol{p}_i = \boldsymbol{p}(ih), \boldsymbol{q}_i = \boldsymbol{q}(ih)$, with

$$\dot{\boldsymbol{p}}(t) = -\frac{\partial H}{\partial \boldsymbol{q}}(\boldsymbol{p}(t), \boldsymbol{q}(t)),\ \dot{\boldsymbol{q}}(t) = \frac{\partial H}{\partial \boldsymbol{p}}(\boldsymbol{p}(t), \boldsymbol{q}(t)), \tag{4.1}$$

where $H(\cdot, \cdot)$ is the latent Hamiltonian function and $h > 0$ is the sampling time step. In this case, the latent ϕ we're trying to learn is the flow map of ODE (Equation 4.1), defined as

$$\phi[\boldsymbol{p}(t), \boldsymbol{q}(t)] := [\boldsymbol{p}(t+h), \boldsymbol{q}(t+h)], \quad \forall t.$$

Given any $h > 0$, the corresponding ϕ is a symplectic map (e.g., [123]), and that is why the learning of symplectic maps is relevant.

Worth noting is, a popular and successful line of thoughts is based on learning the right hand side of the latent ODE, which in our case (Eq. (4.1)) amounts to either generic approaches that learn the vector-valued function $f(\boldsymbol{p}, \boldsymbol{q}) := [-\partial H/\partial \boldsymbol{q}, \partial H/\partial \boldsymbol{p}]$ (see section 4.2 1st paragraph), or specialized methods that utilize the problem structure and learn the scalar-valued function $H(\boldsymbol{p}, \boldsymbol{q})$ (see section 4.2 3rd paragraph). This article, however, is based on a different idea, namely to directly **learn the evolution map** ϕ.

The advantages of doing so include: (i) Generality: it works no matter whether there is an underlying ODE system (see Sec.subsection 4.5.4 for an example where there isn't). (ii) Local Accuracy: for purely data driven problems, learning the map has *one* source of error, namely the approximation error of the map, whereas learning vector field generally has *three*: first one has to estimate the vector field from data, for example by finite difference which incurs error, then there is approximation error of the vector field, and finally the vector field needs to be numerically integrated in order to make predictions and this creates error too. (iii) Symplecticity (and Global Accuracy): we will propose a simple way to exactly maintain the symplecticity of ϕ despite using an approximation[1]; we will also rigorously show, when considering how local error accumulates for making long time predictions, exact symplecticity can help significantly by keeping the error accumulation additive, so that global error grows linearly instead of exponentially.

Example 1. *To make things concrete, consider latent Hamiltonian dynamics* $\dot{x} = Ax$ *where* $A = -A^T$. *A vector-field-based method aims at learning the function* Ax, *but if discrete time-series are the only available data, it actually learns* $\tilde{A}x$ *instead, where* $\tilde{A} = (\exp(Ah) - I)/h$ *if 1st-order finite difference is used for estimating the vector-field. A map-based method seeks the map* $\exp(Ah)x$ *instead. When later making predictions, the vector-field-based method numerically integrates the vector field* $\tilde{A}x$, *which oftentimes corresponds to constructing a polynomial in* h *approximation of* $\exp(\tilde{A}h)$; *on the contrary, a map-based method requires no numerical integration. A similar comparison holds for nonlinear cases too.*

In order to learn a *symplectic* evolution map, we use a tool known as generating functions, which have one-to-one correspondence with symplectic maps. We use a Neural Network, however not for approximating the latent symplectic map, but to approximate its corresponding Generating Function (the method is thus called GFNN). By doing so, the associated evolution map is always symplectic, whether or not it is a good approximation of the latent evolution map, and an appropriate neural network, even just a feedforward one, will be a good approximation after training (see Rmk. 3). This symplectic map representation is intrinsic, purely due to the symplectic structure, and no regularization is used.

Moreover, the guaranteed symplecticity originated from the generating function technique allows us to obtain a nontrivial, linearly growing bound on the prediction error:

Theorem 1 (Informal version of Thm.Equation 4)**.** *Consider latent dynamics far from chaos (more precisely, being integrable). If the latent generating function is approximated*

---

[1] Vector-field-based methods can also be designed to make symplectic predictions; see Sec.section 4.2 4th paragraph.

with $\leq \varepsilon$ *error in first derivatives, then except for a set of initial conditions whose measure goes to 0 as* $\varepsilon \downarrow 0$*, the deviation between the predicted sequence* $(\boldsymbol{p}_0, \boldsymbol{q}_0), (\boldsymbol{p}_1, \boldsymbol{q}_1,), \ldots$ *and the true sequence satisfies*

$$\begin{cases} \|\boldsymbol{p}_n - \boldsymbol{p}(nh)\|_2 \leq C \cdot (nh) \cdot \varepsilon, \\ \|\boldsymbol{q}_n - \boldsymbol{q}(nh)\|_2 \leq C \cdot (nh) \cdot \varepsilon, \end{cases} \quad \forall n \leq h^{-1}\varepsilon^{-1},$$

for some constant $C > 0$*, where* n *is the number of prediction steps and* h *is the sampling time step of the data.*

The merit of this bound lies in long time predictions: note n can be arbitrarily large as ε can be infinitesimal (h is fixed by the training data, and nh is the physical prediction time).

A brief summary of main contributions:

- (Algorithm) *Learn map instead of vector field. Exact symplecticity guaranteed by generating function representation.*
- (Theory) *Linear bound on long-time prediction error.*
- (Validation) *Systematic empirical investigations.*

4.2 Related Works and Discussions

Learning and then predicting dynamics from data is an extremely active research direction. While it is impossible to review all important works, we first mention the classical area of time series (e.g., [124, 125, 126, 127]), where latent differential equations may or may not be assumed. For cases where a latent differential equation is believed to exist, which may correspond to a complex and/or un-modeled underlying dynamical process, some seminal works include [128, 129, 130], and more recent progress include those based on learning (part of) the vector field via sparse regression of a library (e.g., [131, 132, 133, 134, 135, 136, 137]), learning the vector field via neural network (e.g., [138, 139, 140, 141]), and learning the vector field via other approaches such as Gaussian processes (e.g., [142]).

‘Model-free’ approaches that are based on machine learning techniques for sequences have also been proposed, such as [143] (vanilla RNN), [144] (LSTM), [145] (reservoir computing), [146] (CNN), and [147] (transformer).

Faced with the extreme success of these generic methods, interests have also been growing in incorporating domain knowledge and specific structures of the underlying problems into the otherwise-black-box schemes (see e.g., [148]). In terms of mechanical problems modeled by Hamiltonian systems, seminal progress include HNN [149] and an independent work [150], SRNN [151], SympNets [152], and [153, 154, 155, 156, 157], all of which, except SympNets, are related to learning some quantity that produces the Hamiltonian vector field.

In particular, both HNN and SRNN are based on the great idea of learning (using a neural network) the Hamiltonian that generates the vector field (VF), instead of learning the VF itself; this improves accuracy as the Hamiltonian structure of the VF will not be lost due to approximation. HNN learns the Hamiltonian by matching its induced VF with the latent VF (when such information is unavailable, for example in a purely data driven context, data-based approximation such as finite-difference is needed). Then it predicts by numerically integrating the learned VF, and for this we note a Hamiltonian VF doesn’t guarantee the symplecticity of its integration[2]. SRNN, on the other hand, learns the Hamiltonian by matching its symplectic integration with the training sequences, and its prediction is then given by symplectic integration of the learned Hamiltonian. It is therefore the closer to GFNN as it essentially learns a symplectic map; it is just that SRNN represents this map by a symplectic integration of a neural-network-approximated Hamiltonian, whereas GFNN represents it by a neural-network-approximated generating function. Because of this, SRNN doesn’t need finite-difference approximation and has good prediction accuracy, but it only works for symplectic maps originated from Hamiltonian ODEs, and its

---

[2] Unless a symplectic integrator is used. Note the seminal work of HNN used RK45 which is not symplectic, however with small error tolerance (thus good precision but high computation cost).

accuracy is hampered if the latent Hamiltonian is nonseparable[3].

In comparison, GFNN is not based on Hamiltonian vector fields. It is purely data driven, always symplectic, and works the same for separable-, nonseparable-, or even non-Hamiltonian latent systems.

Worth mentioning is the clever recent work of SympNets [152], which also enjoys most of the aforementioned qualitative features of GFNN. It complements GFNN and echoes with our view that directly approximating symplectic maps (instead of Hamiltonian vector fields) in an exactly symplectic way is advantageous (note SRNN can also be seen as a (different) way of doing so). *Algorithmically*, SympNets stack up triangular maps (inspired by symplectic integrator) to construct specialized (new) neural networks, which represent only symplectic maps, and then use them to directly approximate the latent evolution map; GFNN on the other hand uses generating function to indirectly represent the evolution map, and because of its mathematical structure, exact symplecticity is automatically guaranteed, and no special neural network is needed for representing the generating function. Consequently, the *theory* of SympNets is devoted to a universal approximation theorem that characterizes the local prediction error, whereas we focus on the global prediction error (i.e., error after many steps of prediction, instead of one) and rigorously show a nontrivial fact that local errors only accumulate linearly into global error; no approximation theorem needed as it's already established for generic networks. In terms of *performance*, we observe SympNets outperforming vector-field-based approaches (as reasoned above), but GFNN has further improved performance; see e.g., Fig. 4.2, for which we tried up to 30 layers with 10 sublayers using SympNets' code (both LA- and G-SympNets) and plotted its best result, namely LA-SympNets with 30 layers and 10 sublayers (c.f., here GFNN used 5 layers). We feel SympNets generally require a significantly deeper network than GFNN to achieve high approximation power, but then training and computational challenges may arise.

---

[3] The original SRNN is based on symplectic integrators for separable Hamiltonians, and nonsymplectic integrators for nonseparable ones; see Footnote footnote 6 for additional information.

One more remark is, a good amount of existing work considered predicting chaotic dynamics, but a major part of this work is concerned with structured, non-chaotic dynamics, for which controlled long time (strong) accuracy in individual trajectories becomes possible. Predicting chaos is nontrivial, but sometimes the existence of a chaotic attractor makes the system forgiving, and because of that, prediction errors do not accumulate as much as they can in non-chaotic systems. Besides, one often cares more about statistical accuracy for chaotic systems (e.g., [158]), as opposed to strong accuracy in trajectory (which usually grows too fast in chaos; see Rmk. 4). Meanwhile, accurately predicting the trajectory of non-chaotic systems is desirable in numerous applications. Nevertheless, GFNN's predictive power for chaos will be empirically confirmed too.

4.3 Methods

4.3.1 Symplectic Map and Generating Function

As we do not assume or seek a latent ODE system but directly approximate the evolution map, a representation of this symplectic map is essential. Instead of directly approximating it, which has the extra difficulty of losing symplecticity (which has to be exact), we use a mathematical tool known as generating function. Let us be more specific:

Firstly, given a (type-2) generating function, there is an associated symplectic map (a.k.a. canonical transformation):

Lemma 1. *Consider a differentiable function $F(\boldsymbol{q}, \boldsymbol{P})$ which shall be called a generating function. The map $[\boldsymbol{p}, \boldsymbol{q}] \mapsto [\boldsymbol{P}, \boldsymbol{Q}]$ implicitly defined by $\boldsymbol{p} = \frac{\partial F}{\partial \boldsymbol{q}}(\boldsymbol{q}, \boldsymbol{P})$, $\boldsymbol{Q} = \frac{\partial F}{\partial \boldsymbol{P}}(\boldsymbol{q}, \boldsymbol{P})$, is a symplectic map.*

Proof. See e.g., [123]. □

The converse is also true, as long as the latent map doesn't correspond to an evolution time too long (otherwise singularities can be developed):

Lemma 2. *For any infinitesimal symplectomorphism (i.e., symplectic map) on $T^*\mathbb{R}^d$ (i.e., vector phase space), there is a corresponding generating function.*

Proof. This is because the first cohomology group of $T^*\mathbb{R}^d$ is trivial; see e.g., [159]. □

Remark 1 (generating functions and Hamiltonian system)**.** *We do not assume the latent map that generates the data in discrete time corresponds to an underlying Hamiltonian ODE system in continuous time. There are symplectic maps that do not have such correspondence (see e.g., Sec.subsection 4.5.4).*

On the other hand, given a Hamiltonian system, its flow map, defined as $\phi^t : [\boldsymbol{p}(0), \boldsymbol{q}(0)] \mapsto [\boldsymbol{p}(t), \boldsymbol{q}(t)]$, is symplectic for any t. Therefore, there is a family of corresponding generating functions $F(\boldsymbol{q}, \boldsymbol{P}, t)$, each of which generates the symplectic map $[\boldsymbol{P}, \boldsymbol{Q}] = \phi^t[\boldsymbol{p}, \boldsymbol{q}]$. Moreover, the relation between the Hamiltonian H and F can be made more direct via the Hamilton-Jacobi PDE: $H\left(\frac{\partial F}{\partial \boldsymbol{q}}, \boldsymbol{q}, t\right) + \frac{\partial F}{\partial t} = 0$.

Because of their 1-to-1 correspondence, instead of approximating the symplectic evolution map $\phi : [\boldsymbol{p}, \boldsymbol{q}] \mapsto [\boldsymbol{P}, \boldsymbol{Q}]$, we use a Feedforward Neural Network to approximate the corresponding generating function $F(\boldsymbol{q}, \boldsymbol{P})$. This way, no matter how much error the FNN has in approximating F, it always gives to an evolution map that is *exactly* symplectic.

4.3.2 Learning Based on Generating Function Training

The type-2 generating function corresponding to a h-time flow map is $F(\boldsymbol{q}, \boldsymbol{P}) = \boldsymbol{q} \cdot \boldsymbol{P} + \mathcal{O}(h)$, and what varies across different problems is inside the $\mathcal{O}(h)$ term. Therefore, for easier training we learn an equivalent, modified generating function S_h, defined through $F(\boldsymbol{q}, \boldsymbol{P}) = \boldsymbol{q} \cdot \boldsymbol{P} + h \cdot S_h(\boldsymbol{q}, \boldsymbol{P})$. It generates a sequence via iteration

$$\begin{cases} \boldsymbol{p}_i = \boldsymbol{p}_{i+1} + h \cdot \partial_1 S_h(\boldsymbol{q}_i, \boldsymbol{p}_{i+1}), \\ \boldsymbol{q}_{i+1} = \boldsymbol{q}_i + h \cdot \partial_2 S_h(\boldsymbol{q}_i, \boldsymbol{p}_{i+1}), \end{cases} \tag{4.2}$$

as long as an initial condition $[\boldsymbol{p}_0, \boldsymbol{q}_0]$ is provided.

To learn the latent S_h, GFNN uses a neural-network approximation S_h^θ, and trains for a good parameterization θ to best satisfy (Equation 4.2). See Algorithm. 1.

Algorithm 1: GFNN

Data: The data set $\left\{\left([\boldsymbol{p}_{i,j}, \boldsymbol{q}_{i,j}]_{i=0}^{N_j}\right)_{j=1}^{M}\right\}$ is observed from sequences generated by a symplectic map ϕ^h, with $[\boldsymbol{p}_{i,j}, \boldsymbol{q}_{i,j}] \in \mathcal{D} \subseteq \mathbb{R}^d \times \mathbb{R}^d \cong T^*\mathbb{R}^d$ and $\boldsymbol{p}_{i+1,j}, \boldsymbol{q}_{i+1,j} = \phi^h\left(\boldsymbol{p}_{i,j}, \boldsymbol{q}_{i,j}\right)$.

Training: Optimize the loss function

$$\begin{aligned}\mathcal{L}_{GFNN} = &\frac{1}{\sum_{j=1}^{M} N_j} \sum_{j=1}^{M} \sum_{i=0}^{N_j-1} \\ &\Big(\left\| h\partial_2 S_h^\theta(\boldsymbol{q}_{i,j}, \boldsymbol{p}_{i+1,j}) - (\boldsymbol{q}_{i+1,j} - \boldsymbol{q}_{i,j}) \right\|_2^2 \\ &+ \left\| h\partial_1 S_h^\theta(\boldsymbol{q}_{i,j}, \boldsymbol{p}_{i+1,j}) - (\boldsymbol{p}_{i,j} - \boldsymbol{p}_{i+1,j}) \right\|_2^2 \Big).\end{aligned} \tag{4.3}$$

with respect to neural network parameters θ (see Appendix for our experimental details).

Prediction: Given any initial condition $(\boldsymbol{q}_0, \boldsymbol{p}_0) \in \mathcal{D}$, one step evolution to $(\tilde{\boldsymbol{q}}_1, \tilde{\boldsymbol{p}}_1)$ can be solved from

$$\begin{cases} \boldsymbol{p}_0 = \tilde{\boldsymbol{p}}_1 + h \cdot \partial_1 S_h^\theta\left(\boldsymbol{q}_0, \tilde{\boldsymbol{p}}_1\right), \\ \tilde{\boldsymbol{q}}_1 = \boldsymbol{q}_0 + h \cdot \partial_2 S_h^\theta\left(\boldsymbol{q}_0, \tilde{\boldsymbol{p}}_1\right). \end{cases} \tag{4.4}$$

This can be iterated.

4.4 Global Error Analysis

We now show that, under reasonable assumptions, GFNN's prediction will be close to the true sequence (continued by the latent ϕ) for a very long time, as a linearly growing long time error bound will be established. This will be contrasted with an obtainable exponentially growing error bound for generic vector-field-based methods. The latter methods are of course more versatile but they do not utilize the special symplectic structure. Proofs are based on normal form and KAM-type techniques and deferred to Appendix.

The main condition needed for this mild error growth is integrability, which, very roughly speaking, requires the latent system to be far from chaos, but it could still be highly nonlinear; see e.g., [160]. In order to make it precise, some mathematical preparations are needed, but one can jump to Thm.Equation 4 for the main results if preferred.

Definition 1. *A function* $g(\boldsymbol{p}, \boldsymbol{q})$ *is called a 1st-integral or a constant of motion of the dynamics if it remains constant as* $\boldsymbol{p}, \boldsymbol{q}$ *evolves in (continuous or discrete) time.*

Definition 2. *The (canonical) Poisson bracket of two arbitrary functions* $f(\boldsymbol{p}, \boldsymbol{q}), g(\boldsymbol{p}, \boldsymbol{q})$ *is another function defined as* $\{f, g\} := \langle \partial f/\partial q, \partial g/\partial p\rangle - \langle \partial f/\partial p, \partial g/\partial q\rangle$*.*

Theorem 2 (Arnold-Liouville)**.** *Consider a* d*-degree-of-freedom Hamiltonian system. Assume there exist d independent* 1*st integrals in the sense that the Poisson bracket of every pair is* 0*. If the* d*-dim. surfaces implicitly defined by the level sets of those* 1*st integrals are compact, then there exists a canonical transformation from* $\boldsymbol{p}, \boldsymbol{q}$ *to* $\boldsymbol{I}, \boldsymbol{\varphi}$*, such that* $\boldsymbol{\varphi}$ *can be defined on the* d*-torus, and in the new variables the Hamiltonian only depends on* $\boldsymbol{I}$*. In this case,* $\boldsymbol{I}$*,* $\boldsymbol{\varphi}$*, and the Hamiltonian are respectively called the action, angle variables, and an integrable Hamiltonian.*

Proof. See e.g., [160]. ☐

Remark 2. *In an integrable system, the action variables* $\boldsymbol{I}$ *remain constants (they are canonical versions of the first integrals), while the angle variables* $\boldsymbol{\varphi}$ *evolve on an invariant torus* $\left\{\boldsymbol{I} = \boldsymbol{I}(0), \boldsymbol{\varphi} \in \mathbb{T}^d\right\}$*, where*

$$\mathbb{T}^d = \mathbb{R}^d/(2\pi\mathbb{Z}^d) = \{(\varphi_1, \ldots, \varphi_d) \mod 2\pi; \varphi_i \in \mathbb{R}\}.$$

It is easy to show that the fixed time step generating function for an integrable system takes the form of $S_h(\boldsymbol{I}_0, \boldsymbol{\varphi}_1) = H(\boldsymbol{I}_0)$*, as the exact time-h flow is defined as the following*

$$\begin{cases} \boldsymbol{I}_1 = \boldsymbol{I}_0, \\ \boldsymbol{\varphi}_1 = \boldsymbol{\varphi}_0 + h\,\partial_1 S_h(\boldsymbol{I}_0, \boldsymbol{\varphi}_1) = \boldsymbol{\varphi}_0 + h\,\nabla H(\boldsymbol{I}_0). \end{cases} \tag{4.5}$$

Denote $\nabla H(\boldsymbol{I})$ *by* $\boldsymbol{\omega}(\boldsymbol{I}) = [\omega_1(\boldsymbol{I}), \ldots, \omega_d(\boldsymbol{I})]$*. It can be directly seen from Eq.* (4.5) *that* $\omega_i(\boldsymbol{I})$ *represents the change rate (i.e., frequency) of the angle variable* φ_i*.*

Our theory works for almost all initial frequencies, and to describe what are the exceptions we need the following definition, which generalizes irrational numbers in some sense.

Definition 3 ((γ, ν)-Diophantine condition[4])**.** *Frequency vector* $\boldsymbol{\omega} = \{\omega_1, \omega_2, \ldots, \omega_d\}$ *satisfies* (γ, ν)*-Diophantine condition iff* $|\boldsymbol{k} \cdot \boldsymbol{\omega}| \geq \gamma \cdot \|\boldsymbol{k}\|_1^{-\nu}$, $\forall \boldsymbol{k} \in \mathbb{Z}^d$, $\boldsymbol{k} \neq \mathbf{0}$*, for some* $\gamma > 0, \nu > 0$.

Definition 4 (ε-neighborhood condition)**.** $(\boldsymbol{p}, \boldsymbol{q}) \in \mathbb{R}^d \times \mathbb{R}^d$ *of an integrable system satisfies* ε*-neighborhood condition if there exists* $\boldsymbol{I}^* \in \mathbb{R}^d$*, such that* $\boldsymbol{\omega}(\boldsymbol{I}^*)$ *satisfies the* (γ, ν)*-Diophantine condition (definition 3), and* $\|\boldsymbol{I}(\boldsymbol{p}, \boldsymbol{q}) - \boldsymbol{I}^*\|_2 \leq c \cdot |\log \varepsilon|^{-\nu-1}$ *for some* ε *independent constant* c *(defined in the Appendix) with* $\boldsymbol{I}(\boldsymbol{p}, \boldsymbol{q})$ *being the actions of the system.*

With these preparations, we see the action and angle variables $\boldsymbol{I}, \boldsymbol{\varphi}$ form a new coordinate system alternative to $\boldsymbol{p}, \boldsymbol{q}$ (note even if the system is not integrable and/or time is no longer continuous, one is still free to perform any canonical coordinate transformation; it's just doing so may or may not reveal structured dynamics any more). In fact, they give finer estimates of the prediction error:

Theorem 3 (GFNN's long-time prediction error in actions and angles)**.** *Consider an integrable Hamiltonian system written in action-angle variables, whose exact time-h flow map corresponds to generating function* $S_h(\cdot, \cdot)$*. Predict its trajectory using GFNN with learned generating function* $S_h^\theta(\cdot, \cdot)$ *in a bounded data domain* $\mathcal{D} = \mathcal{D}_1 \times \mathbb{T}^d \subseteq \mathbb{R}^d \times \mathbb{T}^d$*.* $\exists\, \varepsilon > 0, \rho > 0$*, such that if the learned generating function* S_h^θ *(extended in a complex*

---

[4] also known as strong non-resonance condition

neighborhood of $\mathcal{D}$) is analytic and satisfies

$$\sum_{i=1,2} \left\| \partial_i S_h^\theta (\cdot, \cdot) - \partial_i S_h (\cdot, \cdot) \right\|_\infty \leq C_1 \varepsilon,$$

for some ε independent constant C_1, where the L^∞ norm is defined over the ε independent complex neighborhood $\mathcal{B}_\rho(\mathcal{D})$ of $\mathcal{D}$, then, $\forall (\boldsymbol{I}_0, \boldsymbol{\varphi}_0) \in \mathcal{D}$ that satisfies ε-neighborhood condition (definition 4), the predicted sequence $(\boldsymbol{I}_0, \boldsymbol{\varphi}_0), (\boldsymbol{I}_1, \boldsymbol{\varphi}_1,), \ldots$ generated by GFNN satisfies

$$\begin{cases} \left\| \boldsymbol{I}_n - \boldsymbol{I}(0) \right\|_2 \leq C \cdot \varepsilon, \\ \left\| \boldsymbol{\varphi}_n - \boldsymbol{\varphi}(nh) \right\|_2 \leq C \cdot (nh) \cdot \varepsilon, \end{cases} \quad \forall n \leq h^{-1} \varepsilon^{-1}, \tag{4.6}$$

for some constant C.

The intuition behind the proof of Thm. 3 (which is in Appendix) is the following: the predicted dynamics $(\boldsymbol{I}_n, \boldsymbol{\varphi}_n)$ and the true dynamics $(\boldsymbol{I}(nh), \boldsymbol{\varphi}(nh))$ deviate because each step of the prediction introduces some error due to inaccurate S_h^θ, but these errors accumulate in a very delicate way; in fact, earlier errors cannot be amplified too much in order for a linear bound to exist. The key reason, as the proof will recover, is that $\boldsymbol{I}_n$'s dynamics is mostly just oscillatory in time. We show this by decomposing the predicted dynamics into a macroscopic part plus microscopic oscillations. The macroscopic part can be proved to correspond to a barely changing action. The microscopic part, on the other hand, does not accumulate.

Specifically, we introduce a carefully-chosen near-identity canonical coordinate change $\mathcal{T} : [\boldsymbol{I}, \boldsymbol{\varphi}] \mapsto [\boldsymbol{J}, \boldsymbol{\theta}]$, $\mathcal{T} \approx id + \mathcal{O}(\varepsilon)$, and show that the new variables $[\boldsymbol{J}_n, \boldsymbol{\theta}_n]$ describe, roughly, the macroscopic part of the predicted dynamics.

We then prove, when compared to the true dynamics,

$$\begin{cases} \left\|\boldsymbol{I}(nh) - \boldsymbol{J}_n\right\|_2 = \mathcal{O}\left(\epsilon\right), \\ \left\|\boldsymbol{\varphi}(nh) - \boldsymbol{\theta}_n\right\|_2 = nh \cdot \mathcal{O}(\varepsilon), \end{cases} \qquad \forall nh = \mathcal{O}\left(\varepsilon^{-1}\right).$$

Since $\mathcal{T}$ is near-identity, $[\boldsymbol{J}_n, \boldsymbol{\theta}_n] = [\boldsymbol{I}_n, \boldsymbol{\varphi}_n] + \mathcal{O}(\varepsilon)$ for all n, and the triangle inequality then completes the proof. □

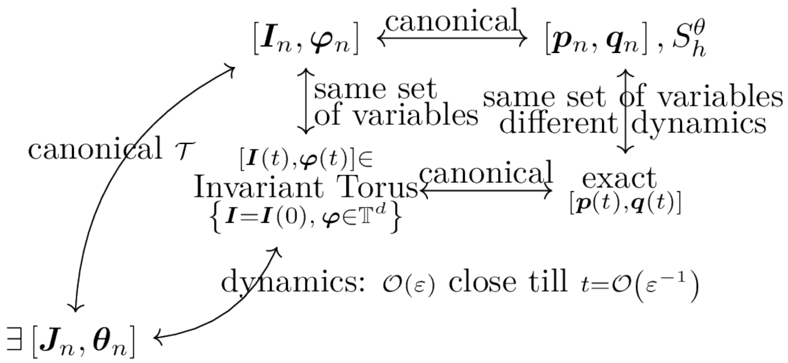


Figure 4.1: Main components in the proof of linear error growth.

Now we can relate the error bound in Thm. 3 back to that for the original variables $\boldsymbol{p}, \boldsymbol{q}$. The big picture is summarized by Fig.Figure 4.1. In the end, our theory only requires the existence of action and angle variables, and no knowledge about how to find the actions and angles is needed.

Theorem 4 (linear growth of GFNN long-time prediction error)**.** *Consider an integrable Hamiltonian system whose exact solution is denoted by $\boldsymbol{p}(t), \boldsymbol{q}(t) \in \mathbb{R}^d$. Denote by $S_h(\cdot,\cdot)$ the generating function corresponding to its exact time-h flow map. Consider predicting its trajectory using GFNN with learned generating function $S_h^\theta(\cdot,\cdot)$ in a bounded data domain $\mathcal{D} \subseteq \mathbb{R}^d \times \mathbb{R}^d$. $\exists\, \varepsilon > 0, \rho > 0$, such that if the learned generating function S_h^θ (extended in a complex neighborhood of $\mathcal{D}$) is analytic and satisfies*

$$\sum_{i=1,2} \left\|\partial_i S_h^\theta(\cdot,\cdot) - \partial_i S_h(\cdot,\cdot)\right\|_\infty \leq C_1 \varepsilon, \tag{4.7}$$

for some ε independent constant C_1, where the L^∞ norm is defined over the ε independent complex neighborhood $\mathcal{B}_\rho(\mathcal{D})$ of $\mathcal{D}$, then, $\forall (\boldsymbol{p}_0, \boldsymbol{q}_0) \in \mathcal{D}$ that satisfies ε-neighborhood condition (definition 4) with nonlinear frequency $\boldsymbol{\omega}(\cdot)$ being given by S_h, the predicted sequence $(\boldsymbol{p}_0, \boldsymbol{q}_0), (\boldsymbol{p}_1, \boldsymbol{q}_1,), \ldots$ generated by GFNN satisfies

$$\begin{cases} \|\boldsymbol{p}_n - \boldsymbol{p}(nh)\|_2 \leq C \cdot (nh) \cdot \varepsilon, \\ \|\boldsymbol{q}_n - \boldsymbol{q}(nh)\|_2 \leq C \cdot (nh) \cdot \varepsilon, \end{cases} \quad \forall n \leq h^{-1}\varepsilon^{-1}, \tag{4.8}$$

for some constant $C > 0$.

Remark 3. *It is known that neural networks can approximate functions and their derivatives with any precision; see e.g., the classical work [161] and a more recent discussion [162].* (Equation 4.7) *can thus be attained.*

Remark 4. *The integrability assumption in Thm.Equation 4 is nontrivial, however reasonable. This is because it rules out the possibility of a positive Lyapunov exponent, which by definition indicates that a deviation between two trajectories can exponentially grow in time (e.g., [163]). Naturally, if the latent system does have a positive Lyapunov exponent, then in general one should not expect a linearly growing prediction error, as an arbitrarily small approximation error, even if it's just made in one step, can be exponentially amplified.*

A simple illustration of this is a Hamiltonian system $\dot{x} = y, \dot{y} = x$, which is not integrable due to noncompactness (not even chaos). It has a Lyapunov exponent of $+1$. Consider predictions based on approximation $\dot{x} = y + \delta_x, \dot{y} = x + \delta_y$, then no matter how small δ_x and δ_y are, the difference between its solution and the original one grows like $\exp(t)$ *except for measure zero δ_x and δ_y values.*

As a comparison, if the prediction map is not symplectic, either due to nonsymplectic numerical integration, or because the learned vector field is no longer Hamiltonian, local prediction error (in each step) may get amplified and long time prediction error may grow exponentially:

Theorem 5. *Consider the latent dynamics $\dot{x} = f(x)$ and its prediction via an Euler integration of the learned vector field $x_{i+1} = x_i + h\tilde{f}(x_i)$, with consistent initial condition $x(0) = x_0$. Assume f is L-Lipschitz continuous, $\mathcal{C}^1$, the learned vector field is accurate up to δ in the sense that $\|\tilde{f} - f\|_\infty \leq \delta$, and the prediction remains bounded. Then the accuracy of the prediction at time $T = nh$ satisfies*

$$\|x(T) - x_n\| \leq \frac{\exp(LT) - 1}{L}(\delta + Lh/2).$$

Proof. See Appendix. $\square$

Remark 5. *This exponential growth with T (and n) is not an overestimation. A simple example that attains it is $f(x) = x$ and $\tilde{f}(x) = x + \delta$. This is of course because the latent dynamics is structurally bad and does not forgive past errors, but that is exactly our point: when the latent dynamics has specific structures such as being a symplectic flow, utilizing those structures in the prediction could lead to much better controlled accumulation of errors.*

Remark 6. *In the context of learning dynamics from data, two sources contribute to the difference between f and $\tilde{f}$. One is approximation error, for instance of the neural network; the other is because one doesn't have an oracle about the latent vector field f but only its approximation from the data, for example $f(x_i) \approx (x_{i+1} - x_i)/h$. A map based approach doesn't directly use f and thus can avoid the latter error, and it doesn't have numerical integration errors in the next phase of predictions either. A neural ODE type treatment [155] can avoid the latter error too, but integration errors in the prediction phase remain (unless computationally expensive small steps are used).*

4.5 Experiments

Let's now systematically (within the page limit) investigate the empirical performances of GFNN. It was conjectured that invariant sets of a smooth map with a dense trajectory are typically either periodic, quasiperiodic[5], or chaotic [164]. Thus, Sec.subsection 4.5.1-subsection 4.5.4 will study classical examples that respectively correspond to periodic, quasiperiodic+chaotic, quasiperiodic, and quasiperiodic+chaotic cases. We'll see smaller and linearly growing errors of GFNN in both periodic and quasiperiodic cases, even when the latent system is not integrable. In chaotic cases, GFNN will also exhibit pleasant statistical accuracy.

VFNN stands for: learning the Vector Field via a Neural Network (without caring about the Hamiltonian structure).

Details of data preparation and training are in section C.2.

4.5.1 An Integrable and Separable Hamiltonian: 2-Body Problem

Consider the motion of 2 gravitationally interacting bodies. Letting their distance be $\boldsymbol{q}(t)$ and the corresponding momentum be $\boldsymbol{p}(t)$, the problem can be equivalently turned into (after unit normalization) an ODE system governed by

$$H(\boldsymbol{p}, \boldsymbol{q}) = \|\boldsymbol{p}\|_2^2/2 - 1/\|\boldsymbol{q}\|_2 .$$

Despite its high nonlinearity, this is an integrable system. Analytical solutions known as Keplerian orbits exist and are periodic in bounded cases. Each solution is described by important physical quantities known as orbital elements, which include semi-major axis and eccentricity, that characterize the shape of the elliptic orbit. As shown in Fig. 4.2, GFNN outperforms other methods and keeps the errors of semi-major axis and eccentricity small

---

[5] A function $f(t)$ is quasiperiodic if $\exists$ some constants $n \in \mathbb{Z}^+$, $\Omega_1, \cdots, \Omega_n \in \mathbb{R}$, and some function F 1-periodic in each argument, s.t., $f(t) = F(\Omega_1 t, \cdots, \Omega_n t)$. An integrable system's solution is quasiperiodic if LCM$(\omega_1(\boldsymbol{I}), \cdots, \omega_d(\boldsymbol{I}))$ doesn't exist (see Rmk.Equation 2 for $\omega_i(\boldsymbol{I})$); otherwise it is periodic.

and bounded, which is consistent with Thm. 3 because semi-major axis and eccentricity are functions of actions known as Delaunay variables [50]. The advantage of GFNN can also been seen in the original variables (e.g., $\boldsymbol{q}$), and the zoomed-in plots in row 2 show that the next two top performers are SRNN (seq_len=2) and SRNN (seq_len=5); SympNets has notably larger error in the orbital phase but its accuracy in the orbital shape is actually comparable to SRNN.

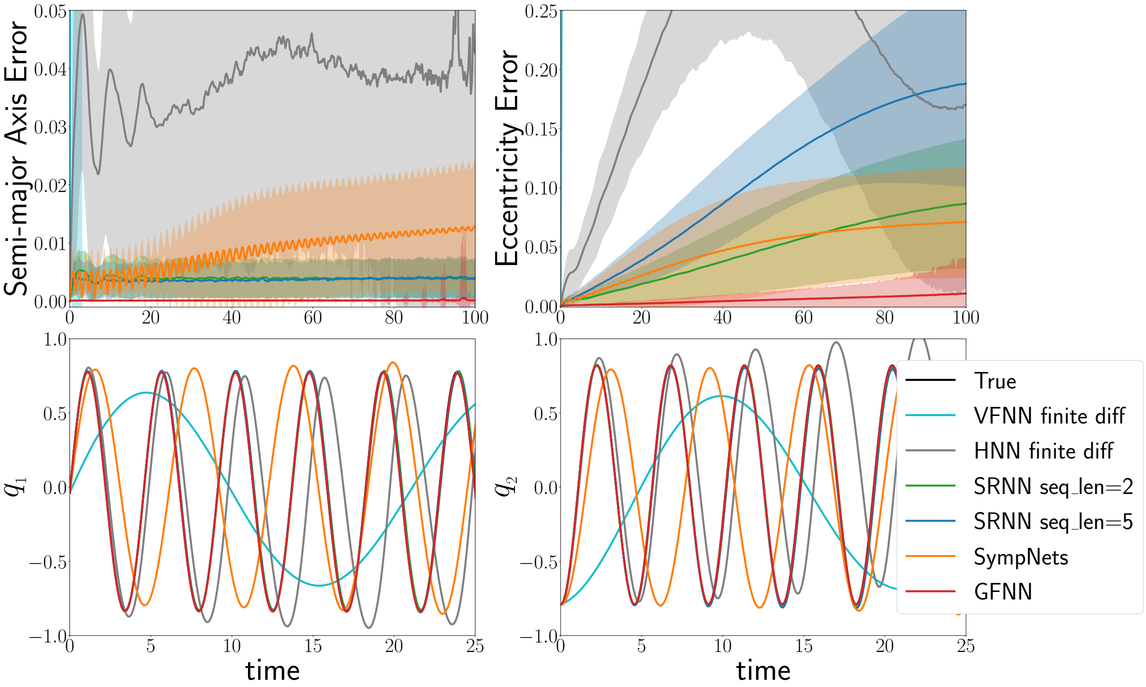


Figure 4.2: Comparison of 2D Keplerian orbit predicted by different methods. The 1st row is the error growth of two variables of physical importance, namely semi-major axis and eccentricity (for this problem, their true values are both constants). Mean values of prediction errors starting from 1,000 i.i.d. initial conditions are plotted with shades corresponding to 1 standard deviation. The 2nd row zooms in the position variables of one predicted sequence (in q_1 and q_2 respectively). Data sequences are prepared with time step 0.1.

4.5.2 A Non-integrable but Separable Hamiltonian System: Hénon-Heiles

The Hénon-Heiles system describes the motion of stars around a galactic center [165]. It is a classical non-integrable system with very complex dynamics, governed by Hamiltonian

$$H\left(p_1, p_2, q_1, q_2\right) = \frac{\left(p_1^2 + p_2^2\right)}{2} + \left(\frac{q_1^2 + q_2^2}{2} + q_1^2 q_2 + \frac{q_2^3}{3}\right).$$

Both chaotic and (quasi)-periodic solutions exist. Initial conditions corresponding to higher energy (i.e., H's value) are more likely to be chaotic. We investigate GFNN's performance in both cases.

Non-periodic but regular motions

For a non-chaotic initial condition, numerically observed was that the long time prediction error of GFNN still grows linearly even though the latent system is no longer integrable; see Fig. 4.3. Notably, SRNN also exhibits linear error growth (although at a higher rate), and this is consistent with our intuition as SRNN also learns a symplectic evolution map (indirectly via the symplectic integration of a Hamiltonian to-be-learned). HNN, on the other hand, has exponentially growing error which quickly saturates to maximum values (due to boundedness of trajectories).

GFNN's linear error growth despite non-integrability is due to the existence of mathematical objects known as KAM-tori (e.g., [166]). They correspond to part of the phase space where dynamics are topologically equivalent to integrable ones. A by-product is, solutions in this region are either quasiperiodic or periodic (see Footnote footnote 5).

Dynamics in chaotic sea

To visualize the prediction of chaotic dynamics, which take place in 4D, we use the standard tool of Poincaré section, which plots where an orbit intersects with a 2D slice of the 3D constant-energy manifold. Fig. 4.4 shows the Poincaré section produced by predictions

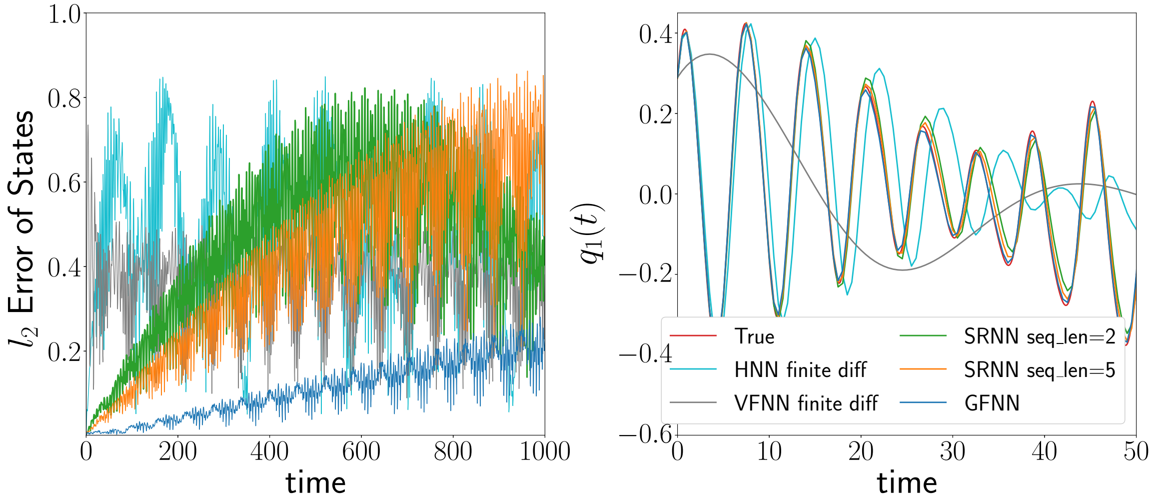


Figure 4.3: Error of predictions of a quasi-periodic trajectory with energy level near $\frac{1}{12}$ of the Hénon-Heiles system. Data sequences are prepared with time step 0.5.

of different methods, based on the same initial condition that leads to chaotic motion via the latent dynamics. The true chaotic motion is ergodic on a submanifold of the phase space, and when restricting to the Poincaré section, it gives intersections that are dense in a subset known as the chaotic sea. Therefore, the shapes of the dense area and the holes inside it (often corresponding to regular islands on which motions are (quasi)-periodic) are indicators of the prediction accuracy. Among methods tested in Fig. 4.4, only VFNN didn't produce a pattern similar to the truth. Quantitative comparisons are conducted by comparing the empirical distributions of points on the Poincaré section, and KL divergences between their marginals and the truth are annotated along with the histograms. GFNN has the smallest errors.

4.5.3 A Non-integrable and Non-separable Hamiltonian: Planar Circular Restricted 3-Body Problem (PCR3BP)

PCR3BP is a special case of the gravitational 3-body problem. In addition to a co-planar restriction, it assumes two bodies massive and the third infinitesimal, which models settings like mission design for a space shuttle near Earth and Moon [37], and understanding a

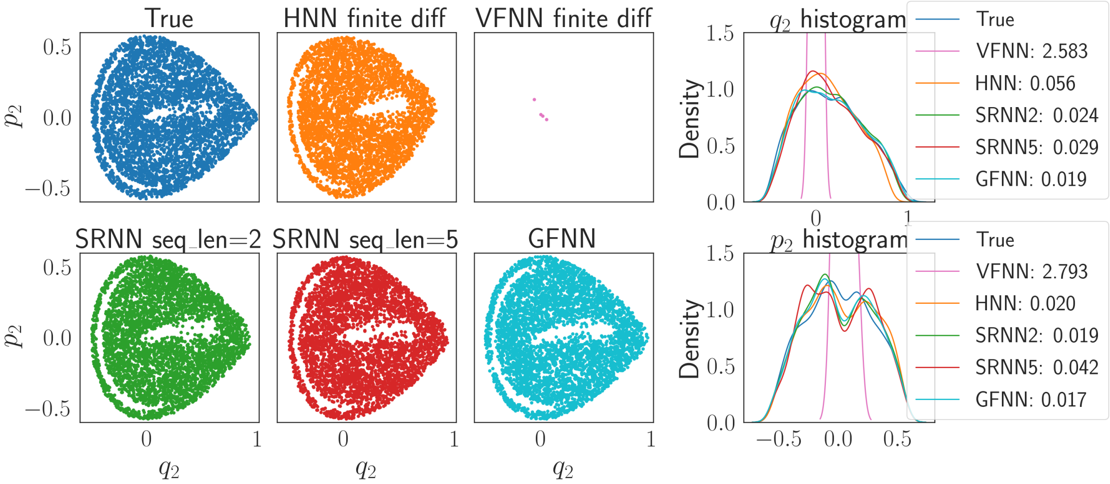


Figure 4.4: Quantifying the statistical accuracy in predicting a chaotic orbit of the Hénon-Heiles system. Left 3 columns: Poincaré section; right column: marginal distributions and their KL divergences from the truth. The plotted orbit corresponds to energy $\frac{1}{6}$; Poincaré section is given by $q_2(t)$, $p_2(t)$ at $q_1(t) = 0$. Data sequences are prepared with time step 0.5.

planet's motion around binary stars [30, 167]. Its Hamiltonian takes the form

$$\begin{aligned} H(\boldsymbol{p}, \boldsymbol{q}) = & \frac{p_1^2 + p_2^2}{2} + p_1 q_2 - p_2 q_1 \\ & - \frac{1-\mu}{\|(q_1 + \mu, q_2)\|_2} - \frac{\mu}{\|(q_1 + \mu - 1, q_2)\|_2}, \end{aligned}$$

with $\mu \in (0, 1)$ a constant mass parameter. Note it cannot be written as $K(\boldsymbol{p}) + V(\boldsymbol{q})$, hence nonseparable.

In order to focus on comparing with SOTA methods for trajectory accuracy, we predict solutions in the nearly-integrable (non-chaotic) regime of PCR3BP; see Fig.Figure 4.5. GFNN still has the smallest error among those experimented, and its growth is again linear. SRNN typically performs the best among tested existing approaches, but its published version loses symplecticity in this case due to non-separability[6], and its accuracy deteriorated. Note also that for methods that learn, in the separable case, $V(\boldsymbol{q})$ in the Hamiltonian or

---

[6] A possible remedy based on our nonseparable symplectic integrators [168] was mentioned in SRNN as a future direction. This remedy is implemented in a concurrent work [157], which successfully reduces the error of predicting nonseparable dynamics to the level of SRNN for separable dynamics.

$\nabla V(\boldsymbol{q})$ in the vector field, now they cannot just do so but have to learn the entire $H(\boldsymbol{p},\boldsymbol{q})$ in doubled dimensions.

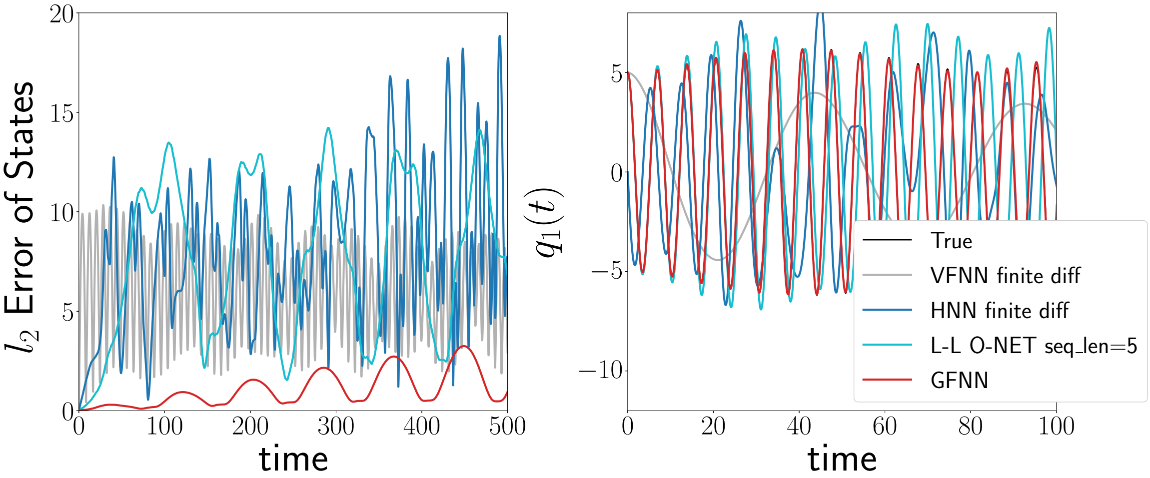


Figure 4.5: Comparison of PCR3BP orbit predicted by different methods. L-L O-NET (in SRNN paper) is selected instead of SRNN as the Hamiltonian is not separable. Data sequences are prepared with time step 0.1.

4.5.4 A Discrete-time Non-(Smooth-)Hamiltonian System: the Standard Map

The standard map is a classical model in accelerator physics. It is a chaotic system whose statistical property is (relatively) well understood. It is a symplectic map in 2D given by

$$\begin{cases} p_{n+1} &= p_n + K\sin(\theta_n), \\ \theta_{n+1} &= \theta_n + p_{n+1}. \end{cases} \tag{4.9}$$

The dynamics is genuinely in *discrete* time, as no smooth Hamiltonian ODE can produce a flow map like it[7]. K is a positive constant that controls the strength of nonlinearity, and it has been estimated that the region of initial conditions leading to chaos has size increasing with K [169].

Methods based on vector fields (e.g., VFNN) or Hamiltonian (e.g., HNN, SRNN) are

---

[7] This is because autonomous Hamiltonian systems in 2D are never chaotic (the Hamiltonian itself is a 1st integral) but the standard map is chaotic.

not very suitable for this prediction task because there is no latent continuous (Hamiltonian) dynamics. One can still apply these methods regardless, for example by using finite differences to construct a fictitious vector field for VFNN and HNN to learn, or just use SRNN without realizing that no Hamiltonian will be able to produce the training data. Their results (obtained using $h = 1$) will be compared with those of GFNN, which is still applicable here as it directly learns evolution maps.

Fig. 4.6 illustrates the predicted evolutions of a fixed initial condition in the chaotic sea (of the true dynamics, $K = 1.2$) by various methods. Note both θ and p have been mod 2π as this quotient compactifies the phase space into the 2-torus without affecting the dynamics (see Eq. (4.9)). Like before, the prediction quality can be inferred from the geometric shape of the set of plotted points, which should match that of the truth (i.e., the latent chaotic sea), and quantitative comparisons can be made using distances between empirical distributions of p, θ values collected along long time predictions (KL divergences from the truth are provided).

One can see GFNN is the only method that captures the major regular islands (the big holes), but even GFNN does not capture the minor regular islands well. The standard map seems to be a challenging problem; HNN and SRNN did not manage to reproduce any chaotic motion, and VFNN completely distorted the chaotic sea.

Fig. 4.7 on the other hand illustrates predictions in regular islands (of the true dynamics, $K = 0.6$). The two (not three, note periodic boundary conditions) elliptical shapes near $p \approx \pi$ and $\theta \approx 0, \pi$ correspond to quasiperiodic orbits, and GFNN is the only one that captures them: the exact trajectory is jumping back and forth between two islands, so does GFNN's prediction, while other methods tend to produce continuous trajectories without capturing the jumps.

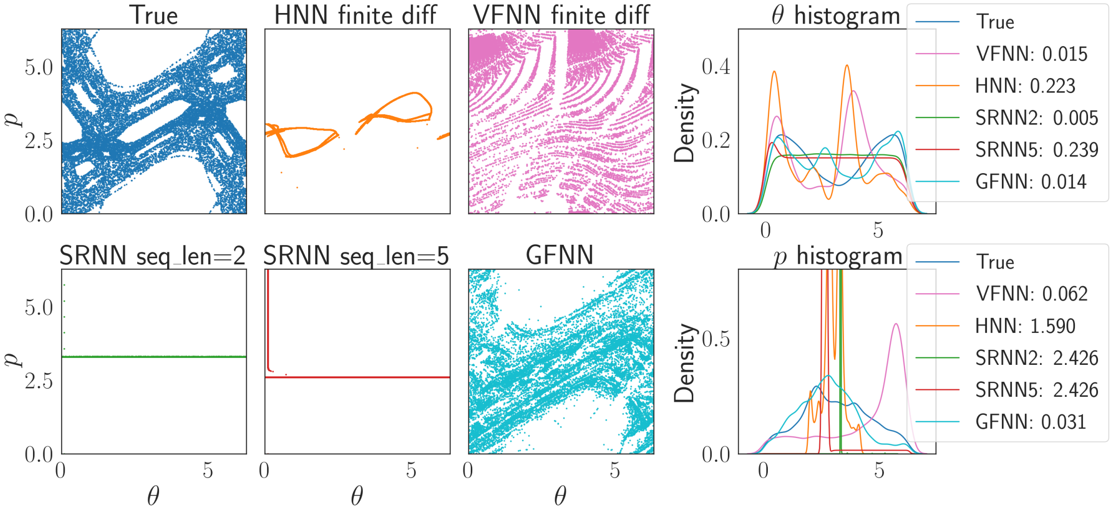


Figure 4.6: Predict a chaotic orbit of the standard map. Left 3 columns: the predicted orbit in phase space; right column: marginals of its empirical measure and their KL divergences from the truth.

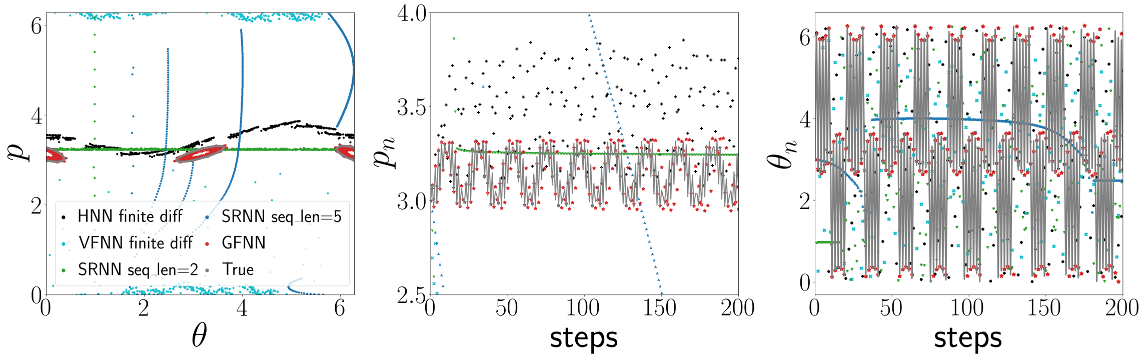


Figure 4.7: Predict a regular orbit of the standard map. The 1st plot is orbits predicted by various methods in phase space. The 2nd and 3rd plots shows how their two coordinates change with time.

Appendices

APPENDIX A
SUPPLEMENTARY MATERIALS OF CHAPTER 2

A.1 Supplementary Information: Perturbative Analysis for Combining Lagrange-Laplace Theory with Our Circumbinary Secular Theory

Theorem 6. *Given* $A = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}$ *and* $B = \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{bmatrix}$*, the eigenvalues of* $A + \epsilon B$ *are* $\lambda_i + \epsilon b_{ii} + \mathcal{O}(\epsilon^2)$ *for* $i = 1, \ldots, n$.

Proof. Let v_i be the i-th unit vector. Obviously $Av_i = \lambda_i v_i$ and $v_j^T A = \lambda_j v_j^T$. Matching $\mathcal{O}(\epsilon)$ terms in

$$(A + \epsilon B)(v_i + \epsilon\delta v_i) = (\lambda_i + \epsilon\delta\lambda_i)(v_i + \epsilon\delta v_i)$$

gives

$$Bv_i + A\delta v_i = \delta\lambda_i v_i + \lambda_i \delta v_i.$$

Multiplying v_i^T from the left gives

$$b_{ii} + v_i^T \lambda_i \delta v_i = \delta\lambda_i + \lambda_i v_i^T \delta v_i.$$

Therefore, $\delta\lambda_i = b_{ii}$. P.S. One can also find about entries δv_i by left multiplication by v_j^T. Multiplying v_j^T from the left gives

$$b_{ji} + v_j^T \lambda_j \delta v_i = \lambda_i v_j^T \delta v_i \Longrightarrow v_j^T \delta v_i = \frac{b_{ji}}{\lambda_j - \lambda_i} \quad \text{(we need } \frac{1}{\lambda_j - \lambda_i} = O(1)\text{)}.$$

□

A.2 Properties of the Dynamical System Corresponding to the Secular Theory

Note that the bifurcations of a physical system happen in the vicinity of $i_p = 90°$, one has to zoom in the i_p to capture the bifurcation. In addition, the topology of the phase portraits across the bifurcations are twisted. Therefore, as a complementary, we explore bifurcation of parameter C_4 by plotting the phase portraits and the bifurcation diagram of X, h under a set of unphysical C_1, C_2, C_3 in Figs. A.1 and A.2. We fix $C_1 = C_2 = C_3 = -1$ in Eq. (2.50) and varying C_4. It can be seen from Figs. A.1 and A.2 that bifurcations happen at $C_4 = -1$, $\frac{7\pm\sqrt{17}}{8}$ and the Hamiltonian pitchfork bifurcation at $C_4 = \frac{7+\sqrt{17}}{8}$ have similar topological change of the phase portraits (shown from Fig. A.2g to Fig. A.2h) with the Hamiltonain bifurcation in Fig. 2.4.

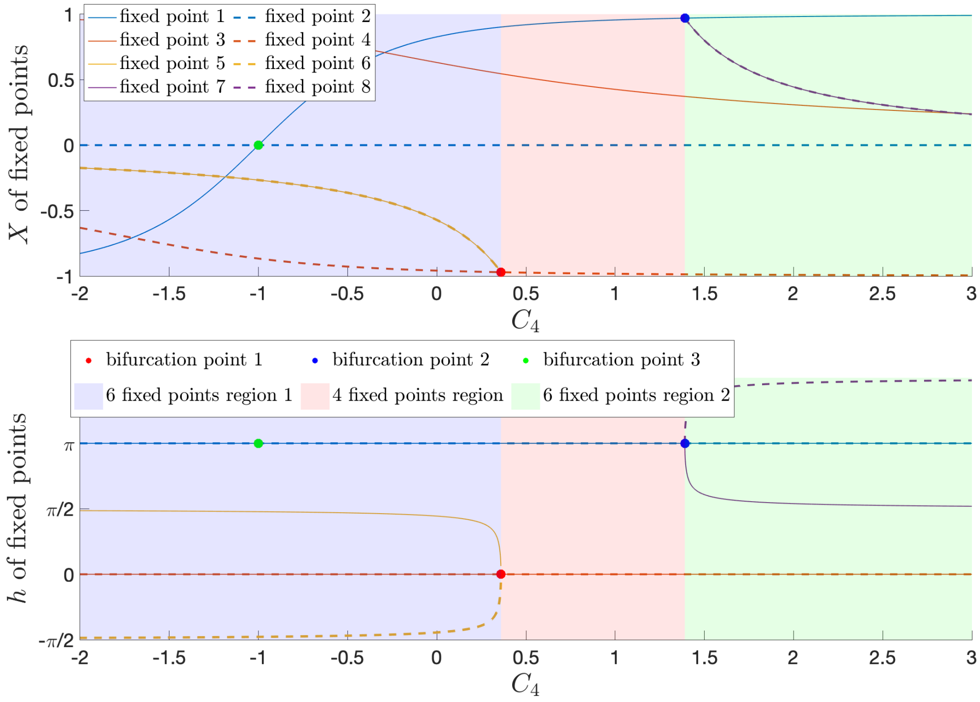
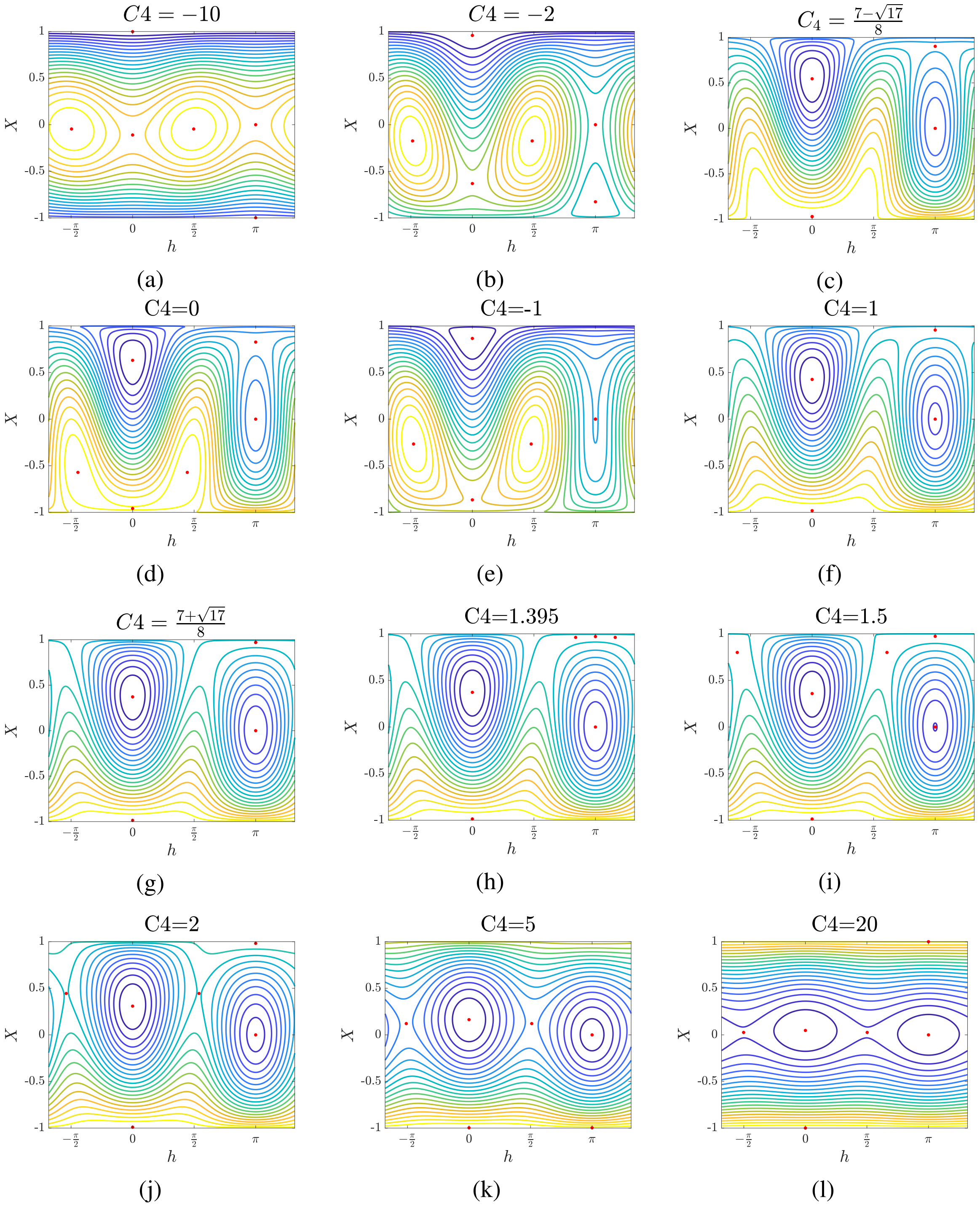


Figure A.1: Bifurcation diagram with the varying parameter being C_4. Families of fixed points are denoted by different colors. Dots indicate bifurcation locations. Fixed point 2,3,5,6,7,8 are centers; fixed point 1 changes from center to saddle at bifurcation point 1; fixed point 4 changes from saddle to center at bifurcation point 2. **Bifurcation point 1** (Hamiltonian pitchfork bifurcation): for $C_4 > {}^{7-\sqrt{17}}/_{8}$, fixed point 4 is a center; for $C_4 < {}^{7-\sqrt{17}}/_{8}$, fixed point 4 is a saddle, there are two more centers (fixed point 5,6). **Bifurcation point 2** (Hamiltonian pitchfork bifurcation): for $C_4 < {}^{7+\sqrt{17}}/_{8}$, fixed point 1 is a center; for $C_4 > {}^{7+\sqrt{17}}/_{8}$, fixed point 1 becomes a saddle, there are two more centers (fixed point 7,8). **Bifurcation point 3** is a saddle-node bifurcation.



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

(j) (k) (l)

Figure A.2: Figs. A.2c, A.2e and A.2g are phase portraits where bifurcations take place.

APPENDIX B

SUPPLEMENTARY MATERIALS OF CHAPTER 3

B.1 Approximation of the Potential Energy

The procedure to approximate $V\left(\boldsymbol{q}_i, \boldsymbol{q}_j, \boldsymbol{R}_i, \boldsymbol{R}_j\right)$ in Eq. (3.9) by Taylor expansion is shown below:

$$
\begin{aligned}
& V\left(\boldsymbol{q}_i, \boldsymbol{q}_j, \boldsymbol{R}_i, \boldsymbol{R}_j\right) \\
={}& \int_{\mathcal{B}_i}\int_{\mathcal{B}_j} -\frac{\mathcal{G}\rho(\boldsymbol{x}_i)\rho(\boldsymbol{x}_j)}{\|(\boldsymbol{q}_i + \boldsymbol{R}_i\boldsymbol{x}_i) - (\boldsymbol{q}_j + \boldsymbol{R}_j\boldsymbol{x}_j)\|}\, d\boldsymbol{x}_i d\boldsymbol{x}_j \\
={}& \int_{\mathcal{B}_i}\int_{\mathcal{B}_j} -\frac{\mathcal{G}\rho(\boldsymbol{x}_i)\rho(\boldsymbol{x}_j)}{\sqrt{\|\boldsymbol{q}_i - \boldsymbol{q}_j\|^2 + \|\boldsymbol{R}_i\boldsymbol{x}_i - \boldsymbol{R}_j\boldsymbol{x}_j\|^2 + 2(\boldsymbol{q}_i - \boldsymbol{q}_j)^T(\boldsymbol{R}_i\boldsymbol{x}_i - \boldsymbol{R}_j\boldsymbol{x}_j)}}\, d\boldsymbol{x}_i d\boldsymbol{x}_j \\
={}& \int_{\mathcal{B}_i}\int_{\mathcal{B}_j} -\frac{\mathcal{G}\rho(\boldsymbol{x}_i)\rho(\boldsymbol{x}_j)}{\|\boldsymbol{q}_i - \boldsymbol{q}_j\|}\left(1 - \frac{1}{2}\left[\frac{\|\boldsymbol{R}_i\boldsymbol{x}_i - \boldsymbol{R}_j\boldsymbol{x}_j\|^2 + 2(\boldsymbol{q}_i - \boldsymbol{q}_j)^T(\boldsymbol{R}_i\boldsymbol{x}_i - \boldsymbol{R}_j\boldsymbol{x}_j)}{\|\boldsymbol{q}_i - \boldsymbol{q}_j\|^2}\right]\right. \\
& \left. + \frac{3}{8}\left[\frac{\|\boldsymbol{R}_i\boldsymbol{x}_i - \boldsymbol{R}_j\boldsymbol{x}_j\|^2 + 2(\boldsymbol{q}_i - \boldsymbol{q}_j)^T(\boldsymbol{R}_i\boldsymbol{x}_i - \boldsymbol{R}_j\boldsymbol{x}_j)}{\|\boldsymbol{q}_i - \boldsymbol{q}_j\|^2}\right]^2\right) d\boldsymbol{x}_i d\boldsymbol{x}_j + \mathcal{O}(\eta^3) \\
={}& -\frac{\mathcal{G}m_i m_j}{\|\boldsymbol{q}_i - \boldsymbol{q}_j\|} + \frac{\mathcal{G}\left(m_i Tr[\boldsymbol{J}_i^{(d)}] + m_j Tr[\boldsymbol{J}_j^{(d)}]\right)}{2\|\boldsymbol{q}_i - \boldsymbol{q}_j\|^3} \\
& - \frac{3\mathcal{G}(\boldsymbol{q}_i - \boldsymbol{q}_j)^T\left(m_j\boldsymbol{R}_i\boldsymbol{J}_i^{(d)}\boldsymbol{R}_i^T + m_i\boldsymbol{R}_j\boldsymbol{J}_j^{(d)}\boldsymbol{R}_j^T\right)(\boldsymbol{q}_i - \boldsymbol{q}_j)}{2\|\boldsymbol{q}_i - \boldsymbol{q}_j\|^5} + \mathcal{O}(\eta^3)
\end{aligned}
\tag{B.1}
$$

where $\eta = \frac{\max(\mathcal{R}_i, \mathcal{R}_j)}{\|\boldsymbol{q}_i - \boldsymbol{q}_j\|}$ ($\mathcal{R}_i$ is the largest distance from the center in the ith body). If we use $\boldsymbol{J}$ instead of $\boldsymbol{J}_d$, we have

$$\begin{aligned} & V\left(\boldsymbol{q}_i, \boldsymbol{q}_j, \boldsymbol{R}_i, \boldsymbol{R}_j\right) \\ = & -\frac{\mathcal{G} m_i m_j}{\|\boldsymbol{q}_i - \boldsymbol{q}_j\|} - \frac{\mathcal{G}\left(m_i Tr[\boldsymbol{J}_i] + m_j Tr[\boldsymbol{J}_j]\right)}{2\left\|\boldsymbol{q}_i - \boldsymbol{q}_j\right\|^3} \\ & + \frac{3\mathcal{G}(\boldsymbol{q}_i - \boldsymbol{q}_j)^T \left(m_j \boldsymbol{R}_i \boldsymbol{J}_i \boldsymbol{R}_i^T + m_i \boldsymbol{R}_j \boldsymbol{J}_j \boldsymbol{R}_j^T\right)(\boldsymbol{q}_i - \boldsymbol{q}_j)}{2\left\|\boldsymbol{q}_i - \boldsymbol{q}_j\right\|^5} + \mathcal{O}(\eta^3) \end{aligned} \tag{B.2}$$

Higher order expansions:

$$
\begin{aligned}
&V\left(\boldsymbol{q}_i, \boldsymbol{q}_j, \boldsymbol{R}_i, \boldsymbol{R}_j\right) \\
=&\mathcal{G} m_i m_j \Bigg\{ -\frac{1}{\left\|\boldsymbol{q}_i-\boldsymbol{q}_j\right\|}+\frac{1}{2\left\|\boldsymbol{q}_i-\boldsymbol{q}_j\right\|^3}\left[\frac{1}{5}\left(a_i^2+b_i^2+c_i^2+a_j^2+b_j^2+c_j^2\right)\right] \\
&-\frac{3}{2\left\|\boldsymbol{q}_i-\boldsymbol{q}_j\right\|^5}\left(\boldsymbol{q}_i-\boldsymbol{q}_j\right)^T\left(\boldsymbol{R}_i \frac{1}{5}\begin{bmatrix} a_i^2 & 0 & 0 \\ 0 & b_i^2 & 0 \\ 0 & 0 & c_i^2 \end{bmatrix} \boldsymbol{R}_i^T+\boldsymbol{R}_j \frac{1}{5}\begin{bmatrix} a_j^2 & 0 & 0 \\ 0 & b_j^2 & 0 \\ 0 & 0 & c_j^2 \end{bmatrix} \boldsymbol{R}_j^T\right)\left(\boldsymbol{q}_i-\boldsymbol{q}_j\right) \\
&-\frac{3}{8\left\|\boldsymbol{q}_i-\boldsymbol{q}_j\right\|^5}\left(\frac{1}{35}(3a_i^4+3b_i^4+3c_i^4+2(a_i^2b_i^2+a_i^2c_i^2+b_i^2c_i^2))+\frac{1}{35}(3a_j^4+3b_j^4+3c_j^4+2(a_j^2b_j^2+a_j^2c_j^2+b_j^2c_j^2))\right) \\
&-\frac{3}{4\left\|\boldsymbol{q}_i-\boldsymbol{q}_j\right\|^5} Tr\left[\boldsymbol{R}_i^T \boldsymbol{R}_j \frac{1}{5}\begin{bmatrix} a_j^2 & 0 & 0 \\ 0 & b_j^2 & 0 \\ 0 & 0 & c_j^2 \end{bmatrix} \boldsymbol{R}_j^T \boldsymbol{R}_i \frac{1}{5}\begin{bmatrix} a_i^2 & 0 & 0 \\ 0 & b_i^2 & 0 \\ 0 & 0 & c_i^2 \end{bmatrix}\right] \\
&+\frac{15}{4\left\|\boldsymbol{q}_i-\boldsymbol{q}_j\right\|^7}\left(\left(\boldsymbol{q}_i-\boldsymbol{q}_j\right)^T \boldsymbol{R}_i \frac{1}{35}\begin{bmatrix} a_i^2\left(3a_i^3+b_i^2+c_i^2\right) & 0 & 0 \\ 0 & b_i^2\left(a_i^2+3b_i^2+c_i^2\right) & 0 \\ 0 & 0 & c_i^2\left(a_i^2+b_i^2+3c_i^2\right) \end{bmatrix} \boldsymbol{R}_i^T\left(\boldsymbol{q}_i-\boldsymbol{q}_j\right)\right) \\
&+\frac{15}{4\left\|\boldsymbol{q}_i-\boldsymbol{q}_j\right\|^7}\left(\left(\boldsymbol{q}_i-\boldsymbol{q}_j\right)^T \boldsymbol{R}_j \frac{1}{35}\begin{bmatrix} a_j^2\left(3a_j^2+b_j^2+c_j^2\right) & 0 & 0 \\ 0 & b_j^2\left(a_j^2+3b_j^2+c_j^2\right) & 0 \\ 0 & 0 & c_j^2\left(a_j^2+b_j^2+3c_j^2\right) \end{bmatrix} \boldsymbol{R}_j^T\left(\boldsymbol{q}_i-\boldsymbol{q}_j\right)\right) \\
&+\frac{15}{4\left\|\boldsymbol{q}_i-\boldsymbol{q}_j\right\|^7}\left(\frac{1}{5}\left(a_i^2+b_i^2+c_i^2\right)\left(\boldsymbol{q}_i-\boldsymbol{q}_j\right)^T \boldsymbol{R}_j \frac{1}{5}\begin{bmatrix} a_j^2 & 0 & 0 \\ 0 & b_j^2 & 0 \\ 0 & 0 & c_j^2 \end{bmatrix} \boldsymbol{R}_j^T\left(\boldsymbol{q}_i-\boldsymbol{q}_j\right)\right) \\
&+\frac{15}{4\left\|\boldsymbol{q}_i-\boldsymbol{q}_j\right\|^7}\left(\frac{1}{5}\left(a_j^2+b_j^2+c_j^2\right)\left(\boldsymbol{q}_i-\boldsymbol{q}_j\right)^T \boldsymbol{R}_i \frac{1}{5}\begin{bmatrix} a_i^2 & 0 & 0 \\ 0 & b_i^2 & 0 \\ 0 & 0 & c_i^2 \end{bmatrix} \boldsymbol{R}_i^T\left(\boldsymbol{q}_i-\boldsymbol{q}_j\right)\right) \\
&+\frac{15}{\left\|\boldsymbol{q}_i-\boldsymbol{q}_j\right\|^7}\left(\left(\boldsymbol{q}_i-\boldsymbol{q}_j\right)^T \boldsymbol{R}_i \frac{1}{5}\begin{bmatrix} a_i^2 & 0 & 0 \\ 0 & b_i^2 & 0 \\ 0 & 0 & c_i^2 \end{bmatrix} \boldsymbol{R}_i^T \boldsymbol{R}_j \frac{1}{5}\begin{bmatrix} a_j^2 & 0 & 0 \\ 0 & b_j^2 & 0 \\ 0 & 0 & c_j^2 \end{bmatrix} \boldsymbol{R}_j^T\left(\boldsymbol{q}_i-\boldsymbol{q}_j\right)\right) \\
&-\frac{35}{8\left\|\boldsymbol{q}_i-\boldsymbol{q}_j\right\|^9}\left(Tr\left[\left(\boldsymbol{q}_i-\boldsymbol{q}_j\right)\left(\boldsymbol{q}_i-\boldsymbol{q}_j\right)^T \boldsymbol{R}_i \frac{3}{35}\begin{bmatrix} a_i^4 & 0 & 0 \\ 0 & b_i^4 & 0 \\ 0 & 0 & c_i^4 \end{bmatrix} \boldsymbol{R}_i^T\left(\boldsymbol{q}_i-\boldsymbol{q}_j\right)\left(\boldsymbol{q}_i-\boldsymbol{q}_j\right)^T\right]\right) \\
&-\frac{35}{8\left\|\boldsymbol{q}_i-\boldsymbol{q}_j\right\|^9}\left(Tr\left[\left(\boldsymbol{q}_i-\boldsymbol{q}_j\right)\left(\boldsymbol{q}_i-\boldsymbol{q}_j\right)^T \boldsymbol{R}_j \frac{3}{35}\begin{bmatrix} a_j^4 & 0 & 0 \\ 0 & b_j^4 & 0 \\ 0 & 0 & c_j^4 \end{bmatrix} \boldsymbol{R}_j^T\left(\boldsymbol{q}_i-\boldsymbol{q}_j\right)\left(\boldsymbol{q}_i-\boldsymbol{q}_j\right)^T\right]\right) \\
&-\frac{105}{4\left\|\boldsymbol{q}_i-\boldsymbol{q}_j\right\|^9}\left(\boldsymbol{q}_i-\boldsymbol{q}_j\right)^T \boldsymbol{R}_i \frac{1}{5}\begin{bmatrix} a_i^2 & 0 & 0 \\ 0 & b_i^2 & 0 \\ 0 & 0 & c_i^2 \end{bmatrix} \boldsymbol{R}_i^T\left(\boldsymbol{q}_i-\boldsymbol{q}_j\right)\left(\boldsymbol{q}_i-\boldsymbol{q}_j\right)^T \boldsymbol{R}_j \frac{1}{5}\begin{bmatrix} a_j^2 & 0 & 0 \\ 0 & b_j^2 & 0 \\ 0 & 0 & c_j^2 \end{bmatrix} \boldsymbol{R}_j^T\left(\boldsymbol{q}_i-\boldsymbol{q}_j\right)\Bigg\}+\mathcal{O}\left(\eta^5\right).
\end{aligned}
\tag{B.3}
$$

B.2 Properties of the *hat-map*

With $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^3$, $\mathbf{D}=\begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{bmatrix}$, we have

- $\hat{\boldsymbol{u}} \boldsymbol{v}=\boldsymbol{u} \times \boldsymbol{v}$.

- $\widehat{\boldsymbol{u} \times \boldsymbol{v}} = \hat{\boldsymbol{u}}\hat{\boldsymbol{v}} - \hat{\boldsymbol{v}}\hat{\boldsymbol{u}}$.

- $\hat{\boldsymbol{u}}\boldsymbol{D} - \boldsymbol{D}\hat{\boldsymbol{u}}^T = Tr\left[\boldsymbol{D}\right]\hat{\boldsymbol{u}} - \widehat{\boldsymbol{D}\boldsymbol{u}}$.

- $\hat{\boldsymbol{u}}^T\hat{\boldsymbol{u}}\boldsymbol{D} - \boldsymbol{D}\hat{\boldsymbol{u}}^T\hat{\boldsymbol{u}} = \widehat{\boldsymbol{u} \times \boldsymbol{D}\boldsymbol{u}}$

B.3 Proof of the Hierarchical Composition Error

Theorem 7. *Given four Hamiltonian flows* $\{\varphi_t^{[i]}\}_{i=1}^4$ *of* H_i *with* $H = H_1 + H_2 + H_3 + H_4$*. Construct an integrator* $\varphi_h := \mathcal{C}_3(\mathcal{C}_1(\varphi_h^{[1]}, \varphi_h^{[2]}), \mathcal{C}_2(\varphi_h^{[3]}, \varphi_h^{[4]}))$ *via composition methods* $\mathcal{C}_i$*,* $i = 1, 2, 3$ *such that*

$$\mathcal{C}_i(\varphi_h^A, \varphi_h^B) = \varphi^A_{a_1^{[i]}h} \circ \varphi^B_{b_1^{[i]}h} \circ \varphi^A_{a_2^{[i]}h} \circ \varphi^B_{b_2^{[i]}h} \circ \cdots \varphi^A_{a_{n_i}^{[i]}h} \circ \varphi^B_{b_{n_i}^{[i]}h}.$$

Then $\mathcal{E}(\varphi_h)$ *equals to the summation of orders of* $\mathcal{C}_i$*,* $i = 1, 2, 3$ *with* $\mathcal{E}(\cdot)$ *being the global error function.*

Proof. Assume the associated Lie operators of $\varphi_t^{[i]}$'s vector fields are $\mathcal{L}_{H_i}$. There exists a Lie operator A_1 such that for $\mathcal{C}_1(\varphi_h^{[1]}, \varphi_h^{[2]})$,

$$e^{a_1^{[1]}\mathcal{L}_{H_1}} e^{b_1^{[1]}\mathcal{L}_{H_2}} e^{a_2^{[1]}\mathcal{L}_{H_1}} e^{b_2^{[1]}\mathcal{L}_{H_2}} \cdots e^{a_{n_1}^{[1]}\mathcal{L}_{H_1}} e^{b_{n_1}^{[1]}\mathcal{L}_{H_2}} = e^{A_1} = e^{\mathcal{L}_{H_1} + \mathcal{L}_{H_2} + E_1}$$

with the order of E_1 equals the order of $\mathcal{C}_1$. Similarly, we have

$$e^{a_1^{[2]}\mathcal{L}_{H_3}} e^{b_1^{[2]}\mathcal{L}_{H_4}} e^{a_2^{[2]}\mathcal{L}_{H_3}} e^{b_2^{[2]}\mathcal{L}_{H_4}} \cdots e^{a_{n_2}^{[2]}\mathcal{L}_{H_3}} e^{b_{n_2}^{[2]}\mathcal{L}_{H_4}} = e^{A_2} = e^{\mathcal{L}_{H_3} + \mathcal{L}_{H_4} + E_2}$$

with the order of E_2 equals the order of $\mathcal{C}_2$. Further for $\mathcal{C}_3$,

$$\begin{aligned} & e^{a_1^{[3]}A_1} e^{b_1^{[3]}A_2} e^{a_2^{[3]}A_1} e^{b_2^{[3]}A_2} \cdots e^{a_{n_3}^{[3]}A_1} e^{b_{n_3}^{[3]}A_2} \\ =& e^{A_3} = e^{A_1 + A_2 + E_3} = e^{\mathcal{L}_{H_1} + \mathcal{L}_{H_2} + \mathcal{L}_{H_3} + \mathcal{L}_{H_4} + E_1 + E_2 + E_3}, \end{aligned}$$

with the order of E_3 equals the order of $\mathcal{C}_3$. Therefore, the global error of φ_h is the summation of the orders of $\mathcal{C}_i,\ i = 1, 2, 3$. $\square$

APPENDIX C

SUPPLEMENTARY MATERIALS OF CHAPTER 4

C.1 Mathematical Proof

C.1.1 Proof of the Generic Prediction Error Bound

Proof of Thm. 5 is shown below.

Lemma 3. *Consider $\dot{x} = f(x)$ and $\dot{y} = f(y)$, where f is L-Lipschitz continuous. Then*

$$\|x(t) - y(t)\| \leq \exp(Lt)\|x(0) - y(0)\|. \tag{C.1}$$

Proof. Since

$$\begin{aligned} x(t) &= x(0) + \int_0^t f(x(\tau))d\tau \\ y(t) &= y(0) + \int_0^t f(y(\tau))d\tau, \end{aligned}$$

triangular inequality and Lipschitz continuity give

$$\begin{aligned} &\|x(t) - y(t)\| \\ &\leq \|x(0) - y(0)\| + \int_0^t \|f(x(\tau)) - f(y(\tau))\|d\tau \\ &\leq \|x(0) - y(0)\| + \int_0^t L\|x(\tau) - y(\tau)\|d\tau. \end{aligned}$$

Gronwall inequality thus gives Eq. (C.1). □

Lemma 4. *Consider $\dot{x} = f(x)$, with $x(0) = x_0$ and L-Lipschitz continuous f. Then*

$$\|x(h) - x_0\| \leq \frac{\exp(Lh) - 1}{L} \|f(x_0)\|. \tag{C.2}$$

(Note $\frac{\exp(Lh)-1}{L} = \mathcal{O}(h)$.)

Proof. Note

$$\begin{aligned} x(h) &= x_0 + \int_0^h f(x(\tau)) d\tau \\ &= x_0 + \int_0^h f(x(\tau)) - f(x_0) + f(x_0) d\tau. \end{aligned}$$

Let $D(t) := x(t) - x_0$. Then triangular inequality and Lipschitz continuity of f give

$$D(h) \leq \int_0^h LD(\tau) + \|f(x_0)\| d\tau$$

Gronwall lemma thus yields

$$D(h) \leq \exp(Lh) D(0) + \frac{\exp(Lh) - 1}{L} \|f(x_0)\|.$$

Since $D(0) = 0$, Eq. (C.2) is proved. □

Proof of Thm. 5. Let $E_n = \|x(nh) - x_n\|$ denote the prediction accuracy, and $\phi^h_{x_0}$ be the h-time flow map of the latent dynamics, i.e., $\phi^h_{x_0} := x(h)$ where $x(\cdot)$ satisfies $\dot{x} = f(x)$ subject to $x(0) = x_0$. Then

$$x((n+1)h) - x_{n+1} = x((n+1)h) - \phi^h_{x_n} + \phi^h_{x_n} - x_{n+1},$$

and therefore

$$E_{n+1} \leq \|x((n+1)h) - \phi^h_{x_n}\| + \|\phi^h_{x_n} - x_{n+1}\|.$$

The first term is exactly $\|\phi^h_{x(nh)} - \phi^h_{x_n}\|$, and by Lemma. 3, it is bounded by

$$\|\phi^h_{x(nh)} - \phi^h_{x_n}\| \leq \exp(Lh)\|x(nh) - x_n\| = \exp(Lh)E_n.$$

For the second term, Taylor expansion gives

$$\phi^h_{x_n} = x_n + hf(x_n) + h^2/2f'(\phi^\xi_{x_n})f(\phi^\xi_{x_n})$$

for some $\xi \in [0, h]$, and therefore

$$\begin{aligned}\|\phi^h_{x_n} - x_{n+1}\| &= \|h(f(x_n) - \tilde{f}(x_n)) + h^2/2f'(\phi^\xi_{x_n})f(\phi^\xi_{x_n})\| \\ &\leq h\delta + h^2/2\|f'(\phi^\xi_{x_n})\|\|f(\phi^\xi_{x_n})\|.\end{aligned}$$

Note $\|f'\| \leq L$ as f is $\mathcal{C}^1$ and L-Lipschitz. For the $f(\phi^\xi_{x_n})$ factor, note Lemma. 4 gives

$$\|\phi^\xi_{x_n} - x_n\| \leq \frac{\exp(L\xi) - 1}{L}\|f(x_n)\|,$$

and therefore

$$\begin{aligned}\|f(\phi^\xi_{x_n})\| &= \|f(x_n) + f(\phi^\xi_{x_n}) - f(x_n)\| \\ &\leq \|f(x_n)\| + \|f(\phi^\xi_{x_n}) - f(x_n)\| \\ &\leq \|f(x_n)\| + L\|\phi^\xi_{x_n} - x_n\| \\ &\leq \exp(L\xi)\|f(x_n)\|\end{aligned}$$

Since $0 \leq \xi \leq h$, $\exp(L\xi)$ is bounded. Moreover, $f(x_n)$ is bounded because f is Lipschitz and therefore continuous and x_n is assumed to be bounded. Therefore, there exists constant C such that

$$\|f'(\phi^\xi_{x_n})\|\|f(\phi^\xi_{x_n})\| \leq C$$

Summarizing both terms, we have

$$E_{n+1} \leq E_n \exp(Lh) + h\delta + Ch^2/2.$$

Mathematical induction thus gives

$$\begin{aligned}
E_N &\leq E_0 \mathrm{exp}(Lh)^N \\
&\quad + \left(\exp(Lh)^{N-1} + \exp(Lh)^{N-2} + \cdots + 1\right)(h\delta + Ch^2/2) \\
&= E_0 \exp(LT) + \frac{\exp(LT) - 1}{\exp(Lh) - 1}(h\delta + Ch^2/2) \\
&\leq E_0 \exp(LT) + \frac{\exp(LT) - 1}{Lh}(h\delta + Ch^2/2) \\
&= \frac{\exp(LT) - 1}{L}(\delta + Ch/2).
\end{aligned}$$

□

C.1.2 Proof of GFNN's Prediction Error Bound

Definition 5 (Diophantine condition). *A frequency vector* $\boldsymbol{\omega} = \{\omega_1, \omega_2, \ldots, \omega_d\}$ *satisfies Diophantine condition if and only there exists positive constants* γ, ν *such that* $\boldsymbol{\omega}$ *satisfies* (γ, ν)*-Diophantine condition.*

Definition 6 ((γ, ν)-Diophantine set). *For a set* $\Omega \subseteq \mathbb{R}^d$*, the corresponding* (γ, ν)*-Diophantine set is defined as*

$$\begin{aligned}
\Omega^* (\gamma, \nu) \overset{def}{=} \big\{ &\boldsymbol{\omega} \in \Omega : \\
&\boldsymbol{\omega} \textit{ satisfy } (\gamma, \nu) \textit{-Diophantine condition} \big\}.
\end{aligned}$$

Definition 7 (Diophantine set). *For a set* $\Omega \subseteq \mathbb{R}^d$*, the corresponding Diophantine set is* *defined as*

$$\Omega^* \overset{def}{=} \bigcup_{\gamma>0,\nu>0} \Omega^*\left(\gamma,\nu\right).$$

Theorem 8. *For any bounded domain $\Omega \subseteq \mathbb{R}^d$, there exists $C > 0$, such that the Lebesgue measure of the complementary of (γ, ν)-Diophantine set with $\nu \geq d$ is bounded from above,*

$$\lambda\left(\Omega \backslash \Omega^*(\gamma,\nu)\right) \leq C \cdot \gamma.$$

Proof. See for instance [89]. ☐

Theorem 9. *For any bounded domain $\Omega \subseteq \mathbb{R}^d$, Diophantine frequencies exist almost everywhere.*

Proof. Since $\lambda\left(\Omega\backslash\Omega^*\right) \leq \lambda\left(\Omega\backslash\Omega^*\left(\gamma,\nu\right)\right)$, $\forall\gamma > 0, \nu > 0$, Thm. 8 gives, $\forall\gamma > 0$,

$$\lambda\left(\Omega\backslash\Omega^*\right) \leq C \cdot \gamma,$$

meaning that Diophantine frequencies exist *almost everywhere* in Ω. ☐

Remark 7. *Even Diophantine frequencies exist almost everywhere in bounded domain $\Omega \subseteq \mathbb{R}^d$, $\Omega\backslash\Omega^*$ is still an open and dense set in $\mathbb{R}^d$ (see for instance [89]).*

Lemma 5 (Cauchy's inequality)**.** *Suppose that f is a holomorphic function on a closed ball $\overline{\mathcal{B}_r\left(\theta^*\right)} \subset \mathbb{C}$ with $r > 0$. If $\left|f\left(\theta\right)\right| \leq M$ for all θ on the boundary of $\mathcal{B}_r\left(\theta^*\right)$, then for all $n \geq 0$,*

$$\left|f^{(n)}\left(\theta^*\right)\right| \leq \frac{n!M}{r^n}.$$

Proof. See for instance [170]. ☐

Definition 8 (average over angles)**.** *Assume $F(\boldsymbol{\theta})$ is periodic in each argument, i.e., $F : \mathbb{T}^d \to \mathbb{R}$, then the (angle) average of F is defined as*

$$\overline{F} = \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} F(\boldsymbol{\theta})\, d\boldsymbol{\theta}. \tag{C.3}$$

Definition 9 (complex extension of $\mathbb{T}^d$)**.** *The complex extension of $\mathbb{T}^d$ of width ρ is defined as*

$$\mathcal{B}_\rho\left(\mathbb{T}^d\right) = \left\{\boldsymbol{\theta} \in \mathbb{T}^d + i\mathbb{R}^d; \|Im\theta\| < \rho\right\}. \tag{C.4}$$

Definition 10. *For an analytic function $\boldsymbol{f}(\cdot) = \begin{bmatrix} f_1(\cdot), & \ldots, & f_d(\cdot) \end{bmatrix} \in \mathbb{C}^d$, we define the following norm*

$$\|\boldsymbol{f}\|_{\infty,S} := \sum_{i=1}^{d} \sup_{x \in S} |f_i(x)|. \tag{C.5}$$

Lemma 6. *Suppose $\boldsymbol{\omega} \in \mathbb{R}^d$ satisfies the (γ,ν)-Diophantine condition and $G(\boldsymbol{\theta}) \in \mathbb{R}$ is a bounded and analytic function on $\mathcal{B}_\rho\left(\mathbb{T}^d\right)$. Then, with $\overline{G}$ being the average of $G(\boldsymbol{\theta})$, the PDE*

$$DF(\boldsymbol{\theta}) \cdot \boldsymbol{\omega} + G(\boldsymbol{\theta}) = \overline{G} \tag{C.6}$$

has a unique real analytic solution $F(\cdot)$ with $\overline{F} = 0$. Moreover, for every positive $\delta < \min(\rho, 1)$, F is bounded on $\mathcal{B}_{\rho-\delta}\left(\mathbb{T}^d\right)$ by

$$\begin{cases} \|F\|_{\infty,\mathcal{B}_{\rho-\delta}(\mathbb{T}^d)} \le \kappa_0 \delta^{-\alpha+1} \|G\|_{\infty,\mathcal{B}_\rho(\mathbb{T}^d)}, \\ \|\partial_{\boldsymbol{\theta}} F\|_{\infty,\mathcal{B}_{\rho-\delta}(\mathbb{T}^d)} \le \kappa_1 \delta^{-\alpha} \|G\|_{\infty,\mathcal{B}_\rho(\mathbb{T}^d)}, \end{cases}$$

with $\alpha = \nu + d + 1$ and $\kappa_0 = \nu^{-1} 8^d 2^\nu \nu!$, $\kappa_1 = \nu^{-1} 8^d 2^{\nu+1}(\nu+1)!$.

Proof. See for instance [89]. □

Lemma 7. *Consider a nearly integrable system with generating function* $S\left(\boldsymbol{I}_0, \boldsymbol{\varphi}_1\right) = S_0\left(\boldsymbol{I}_0\right) + \varepsilon S_1\left(\boldsymbol{I}_0, \boldsymbol{\varphi}_1\right)$*. Suppose* S_0 *and* S_1 *are analytic and bounded in a complex neighborhood of* $\mathcal{D}_1 \subseteq \mathbb{R}^d$ *and* $\mathcal{D} = \mathcal{D}_1 \times \mathbb{T}^d$ *respectively. Then, there exists a real analytic canonical transformation* $(\boldsymbol{J}, \boldsymbol{\theta}) \leftrightarrow (\boldsymbol{I}, \boldsymbol{\varphi})$ *generated by* $\mathcal{T}\left(\boldsymbol{J}, \boldsymbol{\varphi}\right) = \boldsymbol{J} \cdot \boldsymbol{\varphi} + \varepsilon \mathcal{T}_1\left(\boldsymbol{J}, \boldsymbol{\varphi}\right)$*, such that the generating function in* $\boldsymbol{J}, \boldsymbol{\theta}$ *variables takes the form of*

$$\tilde{S}\left(\boldsymbol{J}_0, \boldsymbol{\theta}_1\right) = \tilde{S}_0\left(\boldsymbol{J}_0\right) + \varepsilon^2 \tilde{R}_2\left(\boldsymbol{J}_0, \boldsymbol{\theta}_1, \varepsilon\right). \tag{C.7}$$

with $\varepsilon^2 \tilde{R}_2$ *being a higher-order perturbation to a new integrable system* $\tilde{S}_0$*. Moreover, this result is constructive: the transformation* $\boldsymbol{J}, \boldsymbol{\theta} \leftrightarrow \boldsymbol{I}, \boldsymbol{\varphi}$ *is given by* $\mathcal{T}$ *through*

$$\begin{cases} \boldsymbol{I}_i = \partial_2 \mathcal{T}\left(\boldsymbol{J}_i, \boldsymbol{\varphi}_i\right), \\ \boldsymbol{\theta}_i = \partial_1 \mathcal{T}\left(\boldsymbol{J}_i, \boldsymbol{\varphi}_i\right). \end{cases} \quad \forall i = 0, 1 \tag{C.8}$$

and $\mathcal{T}_1$ *is the solution to the PDE*

$$\partial_2 \mathcal{T}_1\left(\boldsymbol{J}, \boldsymbol{\varphi}\right) \cdot \boldsymbol{\omega}\left(\boldsymbol{J}\right) + G_1\left(\boldsymbol{J}, \boldsymbol{\varphi}\right) = \overline{G_1}(\boldsymbol{J}), \tag{C.9}$$

where $\boldsymbol{\omega}(\cdot) = \nabla S_0\left(\cdot\right)$*,* $G_1(\boldsymbol{J}, \boldsymbol{\varphi}) = S_1(\boldsymbol{J}, \boldsymbol{\varphi} + h\nabla S_0(\boldsymbol{J}))$*, and* $\overline{G_1}(\boldsymbol{J})$ *is its angle average.*

Proof. The generating function S in $\boldsymbol{J}, \boldsymbol{\theta}$ variables ($\tilde{S}$) can be converted in the following using Eq. (C.8),

$$\begin{aligned}
&\tilde{S}\left(\boldsymbol{J}_0, \boldsymbol{\theta}_1\right) = S\left(\boldsymbol{I}_0, \boldsymbol{\varphi}_1\right) \\
=& S\left(\boldsymbol{J}_0 + \varepsilon\partial_2\mathcal{T}_1\left(\boldsymbol{J}_0, \boldsymbol{\varphi}_0\right), \boldsymbol{\varphi}_1\right) \\
=& S_0\left(\boldsymbol{J}_0 + \varepsilon\partial_2\mathcal{T}_1\left(\boldsymbol{J}_0, \boldsymbol{\varphi}_0\right)\right) \\
&+ \varepsilon S_1\left(\boldsymbol{J}_0 + \varepsilon\partial_2\mathcal{T}_1\left(\boldsymbol{J}_0, \boldsymbol{\varphi}_0\right), \boldsymbol{\varphi}_0 + h\partial_1 S(\boldsymbol{I}_0, \boldsymbol{\varphi}_1)\right) \\
&+ \mathcal{O}\left(\varepsilon^2\right) \qquad \text{(C.10)} \\
=& S_0\left(\boldsymbol{J}_0\right) + \varepsilon\partial_2\mathcal{T}_1\left(\boldsymbol{J}_0, \boldsymbol{\varphi}_0\right) \cdot \nabla S_0\left(\boldsymbol{J}_0\right) \\
&+ \varepsilon S_1\left(\boldsymbol{J}_0, \boldsymbol{\varphi}_0 + h\nabla S(\boldsymbol{J}_0)\right) + \mathcal{O}\left(\varepsilon^2\right) \\
=& S_0\left(\boldsymbol{J}_0\right) + \varepsilon\big\{\underline{\partial_2\mathcal{T}_1\left(\boldsymbol{J}_0, \boldsymbol{\varphi}_0\right) \cdot \nabla S_0\left(\boldsymbol{J}_0\right)} \\
&\underline{+ S_1\left(\boldsymbol{J}_0, \boldsymbol{\varphi}_0 + h\nabla S(\boldsymbol{J}_0)\right)}\big\} + \mathcal{O}\left(\varepsilon^2\right)
\end{aligned}$$

with h the constant in Eq. (4.4). Collecting all $\mathcal{O}(\varepsilon^2)$ terms, denoting them by a remainder term $\tilde{R}_2$, and converting all angles in $\tilde{R}_2$ to $\boldsymbol{\theta}_1$, we have

$$\begin{aligned}
&\tilde{S}\left(\boldsymbol{J}_0, \boldsymbol{\theta}_1\right) \\
=& S_0\left(\boldsymbol{J}_0\right) + \varepsilon\big\{\underline{\partial_2\mathcal{T}_1\left(\boldsymbol{J}_0, \boldsymbol{\varphi}_0\right) \cdot \nabla S_0\left(\boldsymbol{J}_0\right)} \qquad \text{(C.11)} \\
&\underline{+ S_1\left(\boldsymbol{J}_0, \boldsymbol{\varphi}_0 + h\nabla S_0(\boldsymbol{J}_0)\right)}\big\} + \varepsilon^2\tilde{R}\left(\boldsymbol{J}_0, \boldsymbol{\theta}_1, \varepsilon\right).
\end{aligned}$$

As long as the terms underlined in Eq. (C.11) add up to a function of $\boldsymbol{J}_0$ only, $\tilde{S}\left(\boldsymbol{J}_0, \boldsymbol{\theta}_1\right)$ won't have angle dependence till the $\mathcal{O}\left(\varepsilon^2\right)$ term. This leads to a solvability requirement. More precisely, let $G_1(\boldsymbol{J}_0, \boldsymbol{\varphi}_0) = S_1(\boldsymbol{J}_0, \boldsymbol{\varphi}_0 + h\nabla S_0(\boldsymbol{J}_0))$ and $\overline{G_1}(\boldsymbol{J}_0)$ be its angle average. Then the PDE

$$\partial_2\mathcal{T}_1\left(\boldsymbol{J}_0, \boldsymbol{\varphi}_0\right) \cdot \nabla S_0\left(\boldsymbol{J}_0\right) + G_1\left(\boldsymbol{J}_0, \boldsymbol{\varphi}_0\right) = \overline{G_1}(\boldsymbol{J}_0) \qquad \text{(C.12)}$$

has a solution $\mathcal{T}_1$, and it makes the underlined terms $\overline{G_1}(\boldsymbol{J}_0)$. Therefore, $\mathcal{T}_1$ and hence the generating function $\mathcal{T}$ can be solved for from Eq. (C.12). The generating function $\tilde{S}$ in

$\boldsymbol{J}, \boldsymbol{\theta}$ variables takes the form

$$\tilde{S}\left(\boldsymbol{J}_0, \boldsymbol{\theta}_1\right) = \tilde{S}_0\left(\boldsymbol{J}_0\right) + \varepsilon^2 \tilde{R}_2\left(\boldsymbol{J}_0, \boldsymbol{\theta}_1, \varepsilon\right),$$

with $\tilde{S}_0(\boldsymbol{J}_0) = S_0(\boldsymbol{J}_0) + \varepsilon\overline{G}(\boldsymbol{J}_0)$. □

Remark 8. *Note that boundedness of $\tilde{R}_2$ requires some extra condition on ∇S_0 (being Diophantine at some point; see Lemma. 9 for details). With bounded $\tilde{R}_2$, the generating function $S\left(\cdot,\cdot\right)$ in $\boldsymbol{I}, \boldsymbol{\varphi}$ variables is near integrable of order $\mathcal{O}\left(\varepsilon\right)$ while $\tilde{S}\left(\cdot,\cdot\right)$ in $\boldsymbol{J}, \boldsymbol{\theta}$ variables is near integrable of order $\mathcal{O}\left(\varepsilon^2\right)$. Therefore, under the transformation $\mathcal{T}$, we get a 'better' set of variables $\boldsymbol{J}, \boldsymbol{\theta}$ instead of $\boldsymbol{I}, \boldsymbol{\varphi}$, as the $\boldsymbol{J}, \boldsymbol{\theta}$ dynamics is closer to being integrable, hence the dynamics of $\boldsymbol{J}, \boldsymbol{\theta}$ can be estimated for longer time.*

Remark 9. *As angles satisfy periodic boundary conditions, $\mathcal{T}_1$ and S_1 can be expanded in Fourier series*

$$\begin{cases} \mathcal{T}_1\left(\boldsymbol{J}, \boldsymbol{\theta}\right) = \sum\limits_{\boldsymbol{k}\in\mathbb{Z}^d} t_{\boldsymbol{k}}\left(\boldsymbol{J}\right) \cdot e^{i(\boldsymbol{k}\cdot\boldsymbol{\theta})}, \\ S_1\left(\boldsymbol{J}, \boldsymbol{\theta}\right) = \sum\limits_{\boldsymbol{k}\in\mathbb{Z}^d} s_{\boldsymbol{k}}\left(\boldsymbol{J}\right) \cdot e^{i(\boldsymbol{k}\cdot\boldsymbol{\theta})}. \end{cases} \tag{C.13}$$

Plugging Eq. (C.13) *into Eq.* (C.9)*, we have*

$$t_{\boldsymbol{k}} \cdot \left(\boldsymbol{k} \cdot \boldsymbol{\omega}\left(\boldsymbol{J}\right)\right) + s_{\boldsymbol{k}} = 0.$$

Noting that if $\boldsymbol{\omega}$ doesn't satisfy Diophantine condition, $\boldsymbol{k}\cdot\boldsymbol{\omega}$ can be small and may even vanish for some $\boldsymbol{k} \in \mathbb{Z}^d$, meaning that under some circumstances, the transformation constructed by $\mathcal{T}\left(\boldsymbol{J}, \boldsymbol{\varphi}\right) = \boldsymbol{J}\cdot\boldsymbol{\varphi} + \varepsilon\mathcal{T}_1\left(\boldsymbol{J}, \boldsymbol{\varphi}\right)$ is no longer of near identity ($Id + \mathcal{O}(\varepsilon)$) as $\mathcal{T}_1$ is not of order $\mathcal{O}\left(1\right)$ any more.

Lemma 8. *Consider a nearly integrable system with generating function $S\left(\boldsymbol{I}_0, \boldsymbol{\varphi}_1\right) = S_0\left(\boldsymbol{I}_0\right) + \varepsilon S_1\left(\boldsymbol{I}_0, \boldsymbol{\varphi}_1\right)$. Suppose S_0 and S_1 are analytic and bounded in a complex neighbor-*

hood of $\mathcal{D}_1 \subseteq \mathbb{R}^d$ *and* $\mathcal{D} = \mathcal{D}_1 \times \mathbb{T}^d$ *respectively. Then, there exists a real analytic canonical transformation* $(\boldsymbol{J}, \boldsymbol{\theta}) \leftrightarrow (\boldsymbol{I}, \boldsymbol{\varphi})$ *generated by* $\mathcal{T}(\boldsymbol{J}, \boldsymbol{\varphi}) = \boldsymbol{J} \cdot \boldsymbol{\varphi} + \sum_{k=1}^{N-1} \varepsilon^k \cdot \mathcal{T}_k(\boldsymbol{J}, \boldsymbol{\varphi})$, *such that the dynamics produced by the original generating function* S *rewritten in* $\boldsymbol{J}, \boldsymbol{\theta}$ *variables corresponds to a transformed generating function* $\tilde{S}$ *given by*

$$\tilde{S}(\boldsymbol{J}_0, \boldsymbol{\theta}_1) = \tilde{S}_0(\boldsymbol{J}_0) + \varepsilon^N \tilde{R}_N(\boldsymbol{J}_0, \boldsymbol{\theta}_1, \varepsilon), \tag{C.14}$$

where $\varepsilon^N \tilde{R}_N$ *is a high-order perturbation to a new integrable system* $\tilde{S}_0(\boldsymbol{J}_0)$*. Here,* $\boldsymbol{J}, \boldsymbol{\theta} \leftrightarrow \boldsymbol{I}, \boldsymbol{\varphi}$ *is defined by* $\mathcal{T}$ *through*

$$\begin{cases} \boldsymbol{I}_i = \partial_2 \mathcal{T}(\boldsymbol{J}_i, \boldsymbol{\varphi}_i), \\ \boldsymbol{\theta}_i = \partial_1 \mathcal{T}(\boldsymbol{J}_i, \boldsymbol{\varphi}_i), \end{cases} \quad \forall i = 0, 1. \tag{C.15}$$

Proof. Apply $\mathcal{T}(\boldsymbol{J}, \boldsymbol{\varphi}) = \boldsymbol{J} \cdot \boldsymbol{\varphi} + \varepsilon \mathcal{T}_1 + \varepsilon \mathcal{T}_2 + \ldots + \varepsilon^{N-1} \mathcal{T}_{N-1}$ to $\tilde{S}(\boldsymbol{J}_0, \boldsymbol{\theta}_1) = S(\boldsymbol{I}_0, \boldsymbol{\varphi}_1)$ like in the proof of Lemma. 7, Taylor expand, and put all $\mathcal{O}(\varepsilon^N)$ terms into $\tilde{R}_N$. Then we have

$$\begin{aligned} \tilde{S}(\boldsymbol{J}_0, \boldsymbol{\theta}_1) = {} & S_0(\boldsymbol{J}_0) \\ & + \varepsilon \left(\partial_2 \mathcal{T}_1(\boldsymbol{J}_0, \boldsymbol{\varphi}_0) \cdot \boldsymbol{\omega}(\boldsymbol{J}_0) + G_1(\boldsymbol{J}_0, \boldsymbol{\varphi}_0) \right) \\ & + \varepsilon^2 \left(\partial_2 \mathcal{T}_2(\boldsymbol{J}_0, \boldsymbol{\varphi}_0) \cdot \boldsymbol{\omega}(\boldsymbol{J}_0) + G_2(\boldsymbol{J}_0, \boldsymbol{\varphi}_0) \right) \\ & + \ldots \\ & + \varepsilon^N \tilde{R}_N(\boldsymbol{J}_0, \boldsymbol{\theta}_1, \varepsilon), \end{aligned} \tag{C.16}$$

for some functions $G_1, \cdots, G_{N-1}$ periodic in the angles, and $\boldsymbol{\omega}(\cdot) := \nabla S(\cdot)$. Similar to G_1 in the proof of Lemma. 7, $G_2, G_3, \cdots$ can be explicitly computed from Taylor expansions, but our proof does not require their specific expressions. Lemma. 7 solved for $\mathcal{T}_1$ (periodic in angles) by making the underlined expression independent of the angles. Repeating a similar procedure at different orders of ε, $\mathcal{T}_i$ (periodic) can be obtained for $i = 1, 2, \cdots$.

Specifically, $\mathcal{T}_i$ satisfies the PDE

$$\partial_2 \mathcal{T}_i(\boldsymbol{J}_0, \boldsymbol{\varphi}_0) \cdot \boldsymbol{\omega}(\boldsymbol{J}_0) + G_i(\boldsymbol{J}_0, \boldsymbol{\varphi}_0) = \overline{G_i}(\boldsymbol{J}_0) \tag{C.17}$$

(The existence of the solution to this is proved in Lemma. 6). In general,

$$\begin{aligned} \tilde{S}\left(\boldsymbol{J}_0, \boldsymbol{\theta}_1\right) &= S_0\left(\boldsymbol{J}_0\right) + \sum_{k=1}^{N-1} \varepsilon^k \overline{G_k}\left(\boldsymbol{J}_0\right) + \varepsilon^N \tilde{R}_N\left(\boldsymbol{J}_0, \boldsymbol{\theta}_1\right) \\ &= \tilde{S}_0\left(\boldsymbol{J}_0\right) + \varepsilon^N \tilde{R}_N\left(\boldsymbol{J}_0, \boldsymbol{\theta}_1\right), \end{aligned}$$

with $\tilde{S}_0\left(\boldsymbol{J}_0\right) = S_0\left(\boldsymbol{J}_0\right) + \sum_{k=1}^{N-1} \varepsilon^k \overline{G_k}\left(\boldsymbol{J}_0\right)$. □

Note that $\tilde{R}_N$ is not necessarily uniformly bounded in the whole data domain in Lemmas. 7 and 8, but in most cases, the uniform boundedness can be established (Lemma. 9) and under that circumstance, we will be able to quantitatively estimate the $\boldsymbol{J}, \boldsymbol{\theta}$ dynamics (Lemma. 10).

Lemma 9. *Consider a nearly integrable system with generating function* $S\left(\boldsymbol{I}_0, \boldsymbol{\varphi}_1\right) = S_0\left(\boldsymbol{I}_0\right) + \varepsilon S_1\left(\boldsymbol{I}_0, \boldsymbol{\varphi}_1\right)$*. Suppose* S_0 *and* S_1 *are analytic and bounded in a complex neighborhood of* $\mathcal{D}_1$ *and* $\mathcal{D} = \mathcal{D}_1 \times \mathbb{T}^d \subseteq \mathbb{R}^d \times \mathbb{T}^d$ *respectively. There exists a real analytic symplectic change of coordinates of order* $\mathcal{O}(\varepsilon)$*:* $(\boldsymbol{I}, \boldsymbol{\varphi}) \leftrightarrow (\boldsymbol{J}, \boldsymbol{\theta})$ *and under this transformation, the generating function in* $\boldsymbol{J}, \boldsymbol{\theta}$ *takes the form*

$$\tilde{S}\left(\boldsymbol{J}_0, \boldsymbol{\theta}_1\right) = \tilde{S}_0\left(\boldsymbol{J}_0\right) + \varepsilon^N \tilde{R}_N\left(\boldsymbol{J}_0, \boldsymbol{\theta}_1, \varepsilon\right).$$

Suppose that $\omega(\boldsymbol{J}^*)$ *satisfies the* (γ, ν)*-Diophantine condition for some* $\boldsymbol{J}^* \in \mathcal{D}_1$*. Then, for any fixed* $N \geq 2$*, there exist positive constants* $\varepsilon_0, c, C, \rho$ *such that if* $\varepsilon \leq \varepsilon_0$*, then*

$$\left\|\tilde{R}_N(\cdot, \cdot)\right\|_{\infty, \overline{\mathcal{B}_{2\delta}(\boldsymbol{J}^*)} \times \overline{\mathcal{B}_\rho\left(\mathbb{T}^d\right)}} \leq C$$

with $\delta = c(N^2 \left|\log \varepsilon\right|)^{-\nu-1}$.

Proof. Applying the canonical transformation $\mathcal{T}$ constructed in Lemma. 8, $\exists \rho', C' > 0$ such that

$$\left\| \tilde{R}_N \left(\boldsymbol{J}^*, \cdot\right) \right\|_{\infty, \mathcal{B}_{\rho'}(\mathbb{T}^d)} \leq C'(N, d, \gamma, \nu) \tag{C.18}$$

Approximate S with respect to angle variables using Fourier series $\hat{S}_m$ till term $m \propto |\log \varepsilon|$ such that the error is of order $\mathcal{O}(\varepsilon^N)$ in a complex neighborhood of the torus $\left\{\boldsymbol{J} = \boldsymbol{J}^*, \boldsymbol{\varphi} \in \mathbb{T}^d\right\}$. Since $|\boldsymbol{k} \cdot \boldsymbol{\omega}\left(\boldsymbol{J}^*\right)| \geq \gamma \left\|\boldsymbol{k}\right\|_1^{-\nu}$, $\forall \boldsymbol{k} \in \mathbb{Z}^d$, then $\exists$ sufficiently small $c > 0$ such that

$$|\boldsymbol{k} \cdot \boldsymbol{\omega}\left(\boldsymbol{J}\right)| \geq \frac{1}{2}\gamma \left\|\boldsymbol{k}\right\|_1^{-\nu}, \ \left\|\boldsymbol{k}\right\|_1 \leq Nm \tag{C.19}$$

for all $\boldsymbol{J} \in \mathcal{B}_{2\delta}(\boldsymbol{J}^*)$ with $\delta = c(N^2 \left|\log \varepsilon\right|)^{-\nu-1}$. As the Fourier coefficients of $\hat{S}_m$ vanishes for $\left\|\boldsymbol{k}\right\|_1 > Nm$, thus according to condition Eq. (C.19) and combining $S = \hat{S}_m + \mathcal{O}(\varepsilon^N)$, $\exists\ \rho'' > 0$ and $C'' > 0$, such that

$$\left\| \tilde{R}_N \left(\boldsymbol{J}, \cdot\right) \right\|_{\infty, \mathcal{B}_{\rho''}(\mathbb{T}^d)} \leq C''(N, d, \gamma, \nu) \tag{C.20}$$

for all $\left\|\boldsymbol{J} - \boldsymbol{J}^*\right\| \leq 2\delta$.

In general, $\exists C$ and ε independent ρ such that

$$\left\| \tilde{R}_N \left(\cdot, \cdot\right) \right\|_{\infty, \overline{\mathcal{B}_{2\delta}(\boldsymbol{J}^*)} \times \overline{\mathcal{B}_{\rho}(\mathbb{T}^d)}} \leq C(N, d, \gamma, \nu).$$

(for the specific forms of C', C'', which are lengthy but obtainable using tools of Fourier series and Cauchy's inequality, see for instance [89]). ☐

Lemma 10. *Consider a nearly integrable system with generating function* $S\left(\boldsymbol{I}_0, \boldsymbol{\varphi}_1\right) = S_0\left(\boldsymbol{I}_0\right) + \varepsilon S_1\left(\boldsymbol{I}_0, \boldsymbol{\varphi}_1\right)$. *Suppose* S_0 *and* S_1 *are analytic and bounded in a complex neighborhood of* $\mathcal{D}_1$ *and* $\mathcal{D} = \mathcal{D}_1 \times \mathbb{T}^d \subseteq \mathbb{R}^d \times \mathbb{T}^d$ *respectively. Then there exists a real analytic near identity symplectic change of coordinates* $\left(\boldsymbol{I}, \boldsymbol{\varphi}\right) \mapsto \left(\boldsymbol{J}, \boldsymbol{\theta}\right)$ *of order* $\mathcal{O}(\varepsilon)$ *and under* *this transformation, the generating function $\tilde{S}$ in $\boldsymbol{J}, \boldsymbol{\theta}$ variables takes the form*

$$\tilde{S}\left(\boldsymbol{J}, \boldsymbol{\theta}\right) = \tilde{S}_0\left(\boldsymbol{J}\right) + \varepsilon^N \tilde{R}_N\left(\boldsymbol{J}, \boldsymbol{\theta}, \varepsilon\right).$$

where $\tilde{S}_0$ only depends on actions. Suppose that $\boldsymbol{\omega}\left(\boldsymbol{J}^*\right)$ satisfies the (γ, ν)-Diophantine condition for some $\boldsymbol{J}^* \in \mathcal{D}_1$. Then, for any fixed $N \geq 2$, $\exists$ positive constants $\varepsilon_0, c, C, \rho$ such that if $\varepsilon \leq \varepsilon_0$, the dynamics of $\boldsymbol{J}, \boldsymbol{\theta}$ (generated by $\tilde{S}$) with $\left\|\boldsymbol{J}_0 - \boldsymbol{J}^*\right\|_2 \leq c\left|\log\varepsilon\right|^{-\nu-1}$ satisfies

$$\begin{cases} \left\|\boldsymbol{J}_n - \boldsymbol{J}_0\right\|_2 \leq Cnh\varepsilon^N, \\ \left\|\boldsymbol{\theta}_n - \tilde{\boldsymbol{\omega}}\left(\boldsymbol{J}_0\right)nh - \boldsymbol{\theta}_0\right\|_2 \\ \qquad \leq C\left(n^2h^2 + nh\left|\log\varepsilon\right|^{\nu+1}\right)\varepsilon^N. \end{cases} \tag{C.21}$$

Here $\boldsymbol{\omega}\left(\cdot\right) = \nabla S_0\left(\cdot\right)$ and $\tilde{\boldsymbol{\omega}}\left(\cdot\right) = \nabla \tilde{S}_0\left(\cdot\right)$.

Proof. According to Lemma. 9, $\exists c > 0, \rho > 0, C' > 0$ such that for $\delta = c(N^2\left|\log\varepsilon\right|)^{-\nu-1}$, $\boldsymbol{J} \in \overline{\mathcal{B}_\delta\left(\boldsymbol{J}^*\right)}$ and $\boldsymbol{\theta} \in \mathcal{B}_\rho\left(\mathbb{T}^d\right)$, $\left|\tilde{R}_N\left(\boldsymbol{J}, \boldsymbol{\theta}\right)\right| \leq C'$. As $\forall \boldsymbol{J} \in \overline{\mathcal{B}_\delta\left(\boldsymbol{J}^*\right)}$, $\overline{\mathcal{B}_\delta\left(\boldsymbol{J}\right)} \subset \overline{\mathcal{B}_{2\delta}\left(\boldsymbol{J}^*\right)}$, $\left|\tilde{R}_N\left(\tilde{\boldsymbol{J}}, \boldsymbol{\theta}\right)\right| \leq C'$ for all $\tilde{\boldsymbol{J}} \in \overline{\mathcal{B}_\delta\left(\boldsymbol{J}\right)}$ and $\boldsymbol{\theta} \in \mathcal{B}_\rho\left(\mathbb{T}^d\right)$. Using Cauchy's inequality (Lemma. 5), we have

$$\left\|\partial_2 \tilde{R}_N\right\|_{\infty, \overline{\mathcal{B}_\delta(\boldsymbol{J}^*)} \times \overline{\mathcal{B}_\rho(\mathbb{T}^d)}} \leq C' \tag{C.22}$$

and

$$\left\|\partial_1 \tilde{R}_N\right\|_{\infty, \overline{\mathcal{B}_\delta(\boldsymbol{J}^*)} \times \overline{\mathcal{B}_\rho(\mathbb{T}^d)}} \leq \frac{C'}{\delta}. \tag{C.23}$$

Plug Eq. (C.22) in the dynamics of $\boldsymbol{J}, \boldsymbol{\theta}$

$$\begin{cases} \boldsymbol{J}_i = \boldsymbol{J}_{i+1} + h\partial_2 \tilde{S}\left(\boldsymbol{J}_i, \boldsymbol{\theta}_{i+1}\right), \\ \boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_i + h\partial_1 \tilde{S}\left(\boldsymbol{J}_i, \boldsymbol{\theta}_{i+1}\right), \end{cases} \tag{C.24}$$

we have

$$\begin{aligned} &\left\|\boldsymbol{J}_{i+1} - \boldsymbol{J}_i\right\|_2 \leq C' h\varepsilon^N, \, \forall i \in \mathbb{N} \\ \implies &\left\|\boldsymbol{J}_n - \boldsymbol{J}_0\right\|_2 \leq C' nh\varepsilon^N. \end{aligned}$$

for the $\boldsymbol{J}$ sequence. On the other hand, for $\boldsymbol{\theta}$ sequence, plug Eq. (C.23) in Eq. (C.24), we have

$$\left\|\boldsymbol{\theta}_{i+1} - \left(\boldsymbol{\theta}_i + h\tilde{\omega}\left(\boldsymbol{J}_i\right)\right)\right\|_2 \leq \frac{C'}{\delta} h\varepsilon^N. \tag{C.25}$$

Since $\tilde{\omega}$ is analytic on a bounded domain, $\tilde{\omega}$ is Lipschitz. Thus, changing $\boldsymbol{J}_i$ in Eq. (C.25) to $\boldsymbol{J}_0$, $\exists C''$ such that

$$\left\|\boldsymbol{\theta}_{i+1} - \left(\boldsymbol{\theta}_i + h\tilde{\omega}\left(\boldsymbol{J}_0\right)\right)\right\|_2 \leq C'' nh^2\varepsilon^N + \frac{C'}{\delta} h\varepsilon^N.$$

Therefore, letting $C = \max(C', C'')$, we have

$$\begin{aligned} &\left\|\boldsymbol{\theta}_n - \left(\boldsymbol{\theta}_0 + nh\tilde{\omega}\left(\boldsymbol{J}_0\right)\right)\right\|_2 \\ &\leq \sum_{i=0}^{n-1} \left\|\boldsymbol{\theta}_{i+1} - \left(\boldsymbol{\theta}_i + h\tilde{\omega}\left(\boldsymbol{J}_0\right)\right)\right\|_2 \\ &\leq Cnh\left(nh + \frac{1}{\delta}\right)\varepsilon^N \\ &\leq Cnh\left(nh + |\log\varepsilon|^{\nu+1}\right)\varepsilon^N, \end{aligned}$$

and Eq. (C.21) is proved. □

Proof of Thm. 3. Since it is assumed that analytic S_h and S_h^θ satisfy

$$\sum_{i=1,2} \left\| \partial_i S_h^\theta\left(\cdot,\cdot\right) - \partial_i S_h\left(\cdot,\cdot\right) \right\|_\infty \leq C_1 \varepsilon$$

on a bounded domain $\mathcal{D}$, S_h^θ is an $\mathcal{O}(\varepsilon)$ perturbation of S_h (note they can also be different by an $\mathcal{O}(1)$ constant, but adding a constant to a generating function does not change its induced dynamics, and we thus assume without loss of generality that there is no such constant difference). Therefore, as S_h is integrable, S_h^θ can be written as $S_h^\theta(\boldsymbol{I}_0, \boldsymbol{\varphi}_1) = S_h(\boldsymbol{I}_0) + \varepsilon S_1(\boldsymbol{I}_0, \boldsymbol{\varphi}_1)$ for some function S_1 modeling the (normalized) perturbation, and is thus nearly integrable. The latent dynamics, i.e., the exact solution of the integrable $S_h(\boldsymbol{I}_0)$ with initial condition $\boldsymbol{I}_0, \boldsymbol{\varphi}_0$ is

$$\begin{cases} \boldsymbol{I}\left(t\right) = \boldsymbol{I}_0, \\ \boldsymbol{\varphi}\left(t\right) = \left(\boldsymbol{\varphi}_0 + \boldsymbol{\omega}\left(\boldsymbol{I}_0\right) t\right) \mod 2\pi. \end{cases}$$

Applying Lemma. 10 with $N \geq 3$ (so that the $(nh)^2 \varepsilon^N$ term is of order $\mathcal{O}(\varepsilon)$ when $nh\varepsilon = \mathcal{O}(1)$) , there exists a near identity canonical transformation $(\boldsymbol{I}, \boldsymbol{\varphi}) \mapsto (\boldsymbol{J}, \boldsymbol{\theta})$ of order ε such that the solution of S_h^θ in $\boldsymbol{J}, \boldsymbol{\theta}$ variable satisfies

$$\begin{cases} \left\|\boldsymbol{J}_n - \boldsymbol{J}_0\right\|_2 \leq C'\varepsilon, \\ \left\|\boldsymbol{\theta}_n - \left(\boldsymbol{\theta}_0 + \tilde{\boldsymbol{\omega}}\left(\boldsymbol{J}_0\right) nh\right)\right\|_2 \leq C'\varepsilon nh, \end{cases} \tag{C.26}$$

for some constant C' (ε independent) $\forall n \leq h^{-1}\varepsilon^{-1}$ (so that $nh\varepsilon$ is of order $\mathcal{O}(1)$) with $\tilde{\boldsymbol{\omega}}\left(\cdot\right)$ defined in Lemma. 10. Note that the canonical transformation holds for all $(\boldsymbol{I}_i, \boldsymbol{\varphi}_i) \leftrightarrow (\boldsymbol{J}_i, \boldsymbol{\theta}_i)$, we have $\left\|\boldsymbol{I}_i - \boldsymbol{J}_i\right\|_2 \leq k\varepsilon$, $\left\|\boldsymbol{\varphi}_i - \boldsymbol{\theta}_i\right\|_2 \leq k\varepsilon$, $\forall i \in \mathbb{N}$ for some constant $k > 0$ and $\left\|\tilde{\boldsymbol{\omega}}\left(\boldsymbol{J}_0\right) - \boldsymbol{\omega}\left(\boldsymbol{I}_0\right)\right\|_2 \leq k'\varepsilon$ for some positive constants k'. Applying triangular inequality,

$\exists\, C > 0$, such that

$$\begin{cases} \left\|\boldsymbol{I}_n - \boldsymbol{I}_0\right\|_2 \leq C\varepsilon, \\ \left\|\boldsymbol{\varphi}_n - \left(\boldsymbol{\varphi}_0 + \boldsymbol{\omega}\left(\boldsymbol{I}_0\right) nh\right)\right\|_2 \leq C\varepsilon nh, \end{cases}$$

for $n \leq h^{-1}\varepsilon^{-1}$. $\square$

Remark 10. *$h^{-1}\varepsilon^{-1}$ is actually a conservative bound for n and one can extend the bound to be $h^{-1}\varepsilon^{(1-N)/2}$, $\forall N \geq 3$. Since N can be arbitrary, even if ε cannot be made infinitesimal, as long as it is below a threshold, the time of validity of the error bound in Thm. 3 can be extended to arbitrarily long.*

C.2 Experimental Details

Like most neural-network based algorithms for learning dynamics, the full potential of GFNN is achieved in the data rich regime. When preparing training data, we not only prepared in an unbiased way, but also emphasized on fair comparisons so that each of the existing methods is given the same or more training data.

More specifically, for each experiment, the training set contains a number of sequences starting with different initial conditions. When training for predicting Hamiltonian dynamics (continuous), each sequence in the training set is stroboscopically sampled from simulated ground truth, which is obtained using high-order numerical integrator with sufficiently small timestep $\tau \ll h$. For each experiment, the data set with sequences of length 2 will be denoted as $\mathcal{D}_2$, and the data set with sequences of length 5 will be denoted as $\mathcal{D}_5$. VFNN, HNN, SRNN (seq_len=2), and GFNN are trained with the same data set $\mathcal{D}_2$, while SRNN (seq_len=5) is trained with $\mathcal{D}_5$. Note the number of flow maps (ϕ) needed for each sequence in $\mathcal{D}_5$ is 4, while the number of maps for each sequence of $\mathcal{D}_2$ is 1. Therefore, for fairness, the number of sequences in $\mathcal{D}_2$, $n_{train}(\mathcal{D}_2)$, is set to be four times $n_{train}(\mathcal{D}_5)$ in most examples (exceptions will be explained).

All experiments are performed with PyTorch (CUDA) on a machine with GeForce RTX 3070 graphic card, AMD Ryzen 7 3700X 8-Core Processor, 16 GB memory and the Linux distribution of openSUSE Leap 15.2.

Codes are provided.

C.2.1 2-Body Problem

The step size of each data sequence in $\mathcal{D}_2$ and $\mathcal{D}_5$ is $h = 0.1$. The ground truth trajectory is simulated using a $4th$ order symplectic integrator with step size 10^{-3}. The initial condition of each data sequence is uniformly drawn from the orbits with semi-major axis $a \in (0.8, 1.2)$, eccentricity $e \in (0, 0.05)$. In terms of the number of samples, $n_{train}(\mathcal{D}_2) = 100,000$, $n_{train}(\mathcal{D}_5) = 100,000$. Note SRNN (seq_len=5) is provided more training data $n_{train}(\mathcal{D}_5)$ than described above, which would be $25,000$ instead, because less training data didn't provide good performance. The time derivative data of the vector field based methods (VFNN, HNN) are generated using (1st-order) finite difference.

S_h^θ is represented using multilayer perceptron (MLP), with 5 layers and 200, 100, 50, 20 neurons in hidden layers. The Adam optimizer is utilized with batch size 200. The model is trained for more than 20 epochs with initial learning rate 0.01. HNN, SRNN, SympNets are trained by their provided codes. HNN, SRNN are trained under default training setups and SympNets is trained using LA-SympNets with 30 layers and 10 sublayers (deeper than their default setups for improved performance).

C.2.2 Hénon-Heiles System

The step size of each data sequence in $\mathcal{D}_2$ and $\mathcal{D}_5$ is $h = 0.5$. The ground truth trajectory is simulated using a $4th$ order symplectic integrator with step size 10^{-3}. The initial condition of each data sequence is drawn from a centered Gaussian perturbation of states along one orbit randomly with variance 0.01^2. In terms of the number of samples, $n_{train}(\mathcal{D}_2) = 100,000$, $n_{train}(\mathcal{D}_5) = 25,000$. The data sets for the regular motion experiment and for the chaotic dynamics experiment are generated separably around a trajectory with energy level $\frac{1}{12}$ and $\frac{1}{6}$ respectively. The time derivative data of the vector field based methods (VFNN, HNN) are generated using (1st-order) finite difference.

The MLP that represents S_h^θ has 5 layers and 200, 100, 50, 20 neurons in hidden layers. The Adam optimizer is utilized with batch size 200. The model is trained for more than 20 epochs with initial learning rate 0.01. HNN, SRNN are trained by their provided codes under default training setups.

C.2.3 PCR3BP

The step size of each data sequence in $\mathcal{D}_2$ and $\mathcal{D}_5$ is $h = 0.1$. The ground truth trajectory is simulated using RK4 with step size 10^{-3}. The initial condition of each data sequence is drawn from a centered Gaussian perturbation of states along one orbit randomly with variance 0.05^2. In terms of the number of samples, $n_{train}(\mathcal{D}_2) = 100,000$, $n_{train}(\mathcal{D}_5) = 25,000$. The time derivative data of the vector field based methods (VFNN, HNN) are generated using (1st-order) finite difference.

The MLP that represents S_h^θ has 5 layers and 200, 100, 50, 20 neurons in hidden layers. The Adam optimizer is utilized with batch size 200. The model is trained for more than 20 epochs with initial learning rate 0.01. HNN, SRNN are trained by their provided codes under default training setups.

C.2.4 Standard Map

The step size of each data sequence in $\mathcal{D}_2$ and $\mathcal{D}_5$ is $h = 1$. The ground truth map is directly evolved from the discrete-time evolution map Eq. (4.9). The initial condition of each data sequence is drawn from a Gaussian perturbation of states along one orbitrandomly with variance 0.5^2. In terms of the number of samples in training / testing data, $n_{train}(\mathcal{D}_2) = 1,000,000$, $n_{train}(\mathcal{D}_5) = 250,000$. The data sets of the regular motion experiment and for the chaotic dynamics experiment are generated separably with $K = 0.6$ and $K = 1.2$ and correspondingly different initial conditions respectively. The time derivative data of the vector field based methods (VFNN, HNN) are generated using (1st-order) finite difference (with $\Delta t = 1$).

The MLP that represents S_h^θ has 5 layers and 500, 500, 200, 20 neurons in hidden layers. The Adam optimizer is utilized with batch size 1000. The model is trained for more than 20 epochs with initial learning rate 0.001. HNN, SRNN are trained by their provided codes under default training setups.

REFERENCES

[1] P. Collaboration, N. Aghanim, Y. Akrami, M. Alves, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. Banday, R. Barreiro, *et al.*, "Planck 2018 results," *A&A*, vol. 641, A12, 2020.

[2] A. Wegener, "Die entstehung der kontinente," *Geologische Rundschau*, vol. 3, no. 4, pp. 276–292, 1912.

[3] M. K. MILANKOVITCH, "Kanon der erdbestrahlung und seine anwendung auf das eiszeitenproblem," *Royal Serbian Academy Special Publication*, vol. 133, pp. 1–633, 1941.

[4] L. Augustin, C. Barbante, P. R. Barnes, J. M. Barnola, M. Bigler, E. Castellano, O. Cattani, J. Chappellaz, D. Dahl-Jensen, B. Delmonte, *et al.*, "Eight glacial cycles from an antarctic ice core," *Nature*, vol. 429, pp. 623–628, 2004.

[5] J. Laskar and P. Robutel, "The chaotic obliquity of the planets," *Nature*, vol. 361, no. 6413, pp. 608–612, Feb. 1993.

[6] J. J. Lissauer, J. W. Barnes, and J. E. Chambers, "Obliquity variations of a moonless Earth," *Icarus*, vol. 217, no. 1, pp. 77–87, Jan. 2012.

[7] J. Touma and J. Wisdom, "The Chaotic Obliquity of Mars," *Science*, vol. 259, no. 5099, pp. 1294–1297, Feb. 1993.

[8] J. A. Sanders, F. Verhulst, and J. Murdock, *Averaging methods in nonlinear dynamical systems*. Springer, 2007, vol. 59.

[9] J. D. Hays, J. Imbrie, and N. J. Shackleton, "Variations in the Earth's Orbit: Pacemaker of the Ice Ages," *Science*, vol. 194, no. 4270, pp. 1121–1132, Dec. 1976.

[10] J. Imbrie and J. Z. Imbrie, "Modeling the Climatic Response to Orbital Variations," *Science*, vol. 207, no. 4434, pp. 943–953, Feb. 1980.

[11] M. E. Raymo, "The timing of major climate terminations," *Paleoceanography*, vol. 12, no. 4, pp. 577–585, Aug. 1997.

[12] W. R. Ward, "Large-Scale Variations in the Obliquity of Mars," *Science*, vol. 181, no. 4096, pp. 260–262, Jul. 1973.

[13] J. W. Head, G. Neukum, R. Jaumann, H. Hiesinger, E. Hauber, M. Carr, P. Masson, B. Foing, H. Hoffmann, M. Kreslavsky, S. Werner, S. Milkovich, S. van Gasselt,

and HRSC Co-Investigator Team, “Tropical to mid-latitude snow and ice accumulation, flow and glaciation on Mars,” *Nature*, vol. 434, no. 7031, pp. 346–351, Mar. 2005.

[14] F. Forget, R. M. Haberle, F. Montmessin, B. Levrard, and J. W. Head, “Formation of Glaciers on Mars by Atmospheric Precipitation at High Obliquity,” *Science*, vol. 311, no. 5759, pp. 368–371, Jan. 2006.

[15] J. Laskar, F. Joutel, and P. Robutel, “Stabilization of the Earth’s obliquity by the Moon,” *Nature*, vol. 361, no. 6413, pp. 615–617, Feb. 1993.

[16] G. Li and K. Batygin, “On the Spin-axis Dynamics of a Moonless Earth,” *Astrophysical Journal*, vol. 790, no. 1, 69, p. 69, Jul. 2014. arXiv: 1404.7505 `[astro-ph.EP]`.

[17] D. M. Williams and J. F. Kasting, “Habitable Planets with High Obliquities,” *Icarus*, vol. 129, no. 1, pp. 254–267, Sep. 1997.

[18] D. S. Spiegel, K. Menou, and C. A. Scharf, “Habitable Climates: The Influence of Obliquity,” *Astrophysical Journal*, vol. 691, no. 1, pp. 596–610, Jan. 2009. arXiv: 0807.4180 `[astro-ph]`.

[19] R. Deitrick, R. Barnes, C. Bitz, D. Fleming, B. Charnay, V. Meadows, C. Wilhelm, J. Armstrong, and T. R. Quinn, “Exo-Milankovitch Cycles. II. Climates of G-dwarf Planets in Dynamically Hot Systems,” *Astronomical Journal*, vol. 155, no. 6, 266, p. 266, Jun. 2018. arXiv: 1805.00283 `[astro-ph.EP]`.

[20] Y. Shan and G. Li, “Obliquity Variations of Habitable Zone Planets Kepler-62f and Kepler-186f,” *Astronomical Journal*, vol. 155, no. 6, 237, p. 237, Jun. 2018. arXiv: 1710.07303 `[astro-ph.EP]`.

[21] M. Saillenfest, J. Laskar, and G. Boué, “Secular spin-axis dynamics of exoplanets,” *Astronomy and Astrophysics*, vol. 623, A4, A4, Mar. 2019. arXiv: `1901.02831` `[astro-ph.EP]`.

[22] A. Duquennoy and M. Mayor, “Multiplicity among solar-type stars in the solar neighbourhood. II - Distribution of the orbital elements in an unbiased sample.,” *Astronomy and Astrophysics*, vol. 500, pp. 337–376, Aug. 1991.

[23] D. Raghavan, H. A. McAlister, T. J. Henry, D. W. Latham, G. W. Marcy, B. D. Mason, D. R. Gies, R. J. White, and T. A. ten Brummelaar, “A Survey of Stellar Families: Multiplicity of Solar-type Stars,” *Astrophysical Journal, Supplement*, vol. 190, no. 1, pp. 1–42, Sep. 2010. arXiv: 1007.0414 `[astro-ph.SR]`.

[24] B. Quarles, G. Li, and J. J. Lissauer, “Obliquity Evolution of Circumstellar Planets in Sun-like Stellar Binaries,” *Astrophysical Journal*, vol. 886, no. 1, 56, p. 56, Nov. 2019. arXiv: 1911.08431 `[astro-ph.EP]`.

[25] L. R. Doyle, J. A. Carter, D. C. Fabrycky, R. W. Slawson, S. B. Howell, J. N. Winn, J. A. Orosz, A. Přsa, W. F. Welsh, S. N. Quinn, D. Latham, G. Torres, L. A. Buchhave, G. W. Marcy, J. J. Fortney, A. Shporer, E. B. Ford, J. J. Lissauer, D. Ragozzine, M. Rucker, N. Batalha, J. M. Jenkins, W. J. Borucki, D. Koch, C. K. Middour, J. R. Hall, S. McCauliff, M. N. Fanelli, E. V. Quintana, M. J. Holman, D. A. Caldwell, M. Still, R. P. Stefanik, W. R. Brown, G. A. Esquerdo, S. Tang, G. Furesz, J. C. Geary, P. Berlind, M. L. Calkins, D. R. Short, J. H. Steffen, D. Sasselov, E. W. Dunham, W. D. Cochran, A. Boss, M. R. Haas, D. Buzasi, and D. Fischer, “Kepler-16: A Transiting Circumbinary Planet,” *Science*, vol. 333, no. 6049, p. 1602, Sep. 2011. arXiv: 1109.3432 `[astro-ph.EP]`.

[26] V. B. Kostov, J. A. Orosz, W. F. Welsh, L. R. Doyle, D. C. Fabrycky, N. Haghighipour, B. Quarles, D. R. Short, W. D. Cochran, M. Endl, E. B. Ford, J. Gregorio, T. C. Hinse, H. Isaacson, J. M. Jenkins, E. L. N. Jensen, S. Kane, I. Kull, D. W. Latham, J. J. Lissauer, G. W. Marcy, T. Mazeh, T. W. A. Müller, J. Pepper, S. N. Quinn, D. Ragozzine, A. Shporer, J. H. Steffen, G. Torres, G. Windmiller, and W. J. Borucki, “Kepler-1647b: The Largest and Longest-period Kepler Transiting Circumbinary Planet,” *Astrophysical Journal*, vol. 827, no. 1, 86, p. 86, Aug. 2016. arXiv: 1512.00189 `[astro-ph.EP]`.

[27] D. J. Armstrong, H. P. Osborn, D. J. A. Brown, F. Faedi, Y. Gómez Maqueo Chew, D. V. Martin, D. Pollacco, and S. Udry, “On the abundance of circumbinary planets,” *Monthly Notices of the RAS*, vol. 444, no. 2, pp. 1873–1883, Oct. 2014. arXiv: 1404.5617 `[astro-ph.EP]`.

[28] G. Li, M. J. Holman, and M. Tao, “Uncovering Circumbinary Planetary Architectural Properties from Selection Biases,” *Astrophysical Journal*, vol. 831, no. 1, 96, p. 96, Nov. 2016. arXiv: 1608.01768 `[astro-ph.EP]`.

[29] R. Chen, G. Li, and M. Tao, “Grit: A package for structure-preserving simulations of gravitationally interacting rigid-bodies,” *arXiv preprint arXiv:2103.12767*, 2021.

[30] G. Li, M. J. Holman, and M. Tao, “Uncovering circumbinary planetary architectural properties from selection biases,” *The Astrophysical Journal*, vol. 831, no. 1, p. 96, 2016.

[31] R. Miranda and D. Lai, “Tidal truncation of inclined circumstellar and circumbinary discs in young stellar binaries,” *Monthly Notices of the RAS*, vol. 452, no. 3, pp. 2396–2409, Sep. 2015. arXiv: 1504.02917 `[astro-ph.EP]`.

[32] D. V. Martin, T. Mazeh, and D. C. Fabrycky, "No circumbinary planets transiting the tightest Kepler binaries - a possible fingerprint of a third star," *Monthly Notices of the RAS*, vol. 453, no. 4, pp. 3554–3567, Nov. 2015. arXiv: `1505.05749 [astro-ph.EP]`.

[33] F. Farago, J. Laskar, and J. Couetdic, "Averaging on the motion of a fast revolving body. Application to the stability study of a planetary system," *Celestial Mechanics and Dynamical Astronomy*, vol. 104, no. 3, pp. 291–306, Jul. 2009.

[34] P. Lochak and C. Meunier, *Multiphase averaging for classical systems: with applications to adiabatic theorems*. Springer Science & Business Media, 2012, vol. 72.

[35] V. Szebehely, *Theory of orbit: The restricted problem of three Bodies*. Elsevier, 2012.

[36] J. Llibre and C. Pinol, "On the elliptic restricted three-body problem," *Celestial Mechanics and Dynamical Astronomy*, vol. 48, no. 4, pp. 319–345, 1990.

[37] W. S. Koon, M. W. Lo, J. E. Marsden, and S. D. Ross, "Heteroclinic connections between periodic orbits and resonance transitions in celestial mechanics," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 10, no. 2, pp. 427–469, 2000.

[38] B. Kumar, R. L. Anderson, and R. de la Llave, "High-order resonant orbit manifold expansions for mission design in the planar circular restricted 3-body problem," *Communications in Nonlinear Science and Numerical Simulation*, vol. 97, p. 105 691, 2021.

[39] O. Neron de Surgy and J. Laskar, "On the long term evolution of the spin of the earth.," *Astronomy and Astrophysics*, vol. 318, pp. 975–989, 1997.

[40] J. A. Orosz, W. F. Welsh, J. A. Carter, D. C. Fabrycky, W. D. Cochran, M. Endl, E. B. Ford, N. Haghighipour, P. J. MacQueen, T. Mazeh, *et al.*, "Kepler-47: A transiting circumbinary multiplanet system," *Science*, vol. 337, no. 6101, pp. 1511–1514, 2012.

[41] D. V. Martin, "The binary mass ratios of circumbinary planet hosts," *Monthly Notices of the RAS*, vol. 488, no. 3, pp. 3482–3491, Sep. 2019. arXiv: `1904.04832 [astro-ph.EP]`.

[42] M. J. Holman and P. A. Wiegert, "Long-Term Stability of Planets in Binary Systems," *Astronomical Journal*, vol. 117, no. 1, pp. 621–628, Jan. 1999. arXiv: astro-ph/9809315 `[astro-ph]`.

[43] G. Li and K. Batygin, “On the spin-axis dynamics of a moonless earth,” *The Astrophysical Journal*, vol. 790, no. 1, p. 69, 2014.

[44] C. D. Murray and S. F. Dermott, *Solar system dynamics*. Cambridge university press, 1999.

[45] J. Schneider, “On the occultations of a binary star by a circum-orbiting dark companion,” *Planetary Space Science*, vol. 42, no. 7, pp. 539–544, Jul. 1994.

[46] J. A. Serret, *Mémoire sur l’emploi de la méthode de la variation des arbitraires dans la théorie des mouvements de rotation*. Firmin Didot, 1866.

[47] F. Tisserand, *Traité de mécanique céleste*. Gauthier-Villars, 1891.

[48] H. Andoyer, “Cours de mécanique céleste,” *Paris, Gauthier-Villars et cie, 1923-26.*, 1923.

[49] P. Gurfil, A. Elipe, W. Tangren, and M. Efroimsky, “The serret-andoyer formalism in rigid-body dynamics: I. symmetries and perturbations,” *Regular and Chaotic Dynamics*, vol. 12, no. 4, pp. 389–425, Aug. 2007.

[50] A. Morbidelli, *Modern celestial mechanics: aspects of solar system dynamics*. 2002.

[51] H. Kinoshita, “First-order perturbations of the two finite body problem,” *Publications of the Astronomical Society of Japan*, vol. 24, p. 423, 1972.

[52] G. Li and K. Batygin, “Pre-late Heavy Bombardment Evolution of the Earth’s Obliquity,” *Astrophysical Journal*, vol. 795, no. 1, 67, p. 67, Nov. 2014. arXiv: 1409.2881 `[astro-ph.EP]`.

[53] B. Quarles, S. Satyal, V. Kostov, N. Kaib, and N. Haghighipour, “Stability Limits of Circumbinary Planets: Is There a Pile-up in the Kepler CBPs?” *Astrophysical Journal*, vol. 856, no. 2, 150, p. 150, Apr. 2018. arXiv: 1802.08868 `[astro-ph.EP]`.

[54] A. C. M. Correia, A. Leleu, N. Rambaux, and P. Robutel, “Spin-orbit coupling and chaotic rotation for circumbinary bodies. Application to the small satellites of the Pluto-Charon system,” *Astronomy and Astrophysics*, vol. 580, L14, p. L14, Aug. 2015. arXiv: 1506.06733 `[astro-ph.EP]`.

[55] A. C. M. Correia, G. Boué, and J. Laskar, “Secular and tidal evolution of circumbinary systems,” *Celestial Mechanics and Dynamical Astronomy*, vol. 126, no. 1-3, pp. 189–225, Nov. 2016. arXiv: 1608.03484 `[astro-ph.EP]`.

[56] E. Gaidos and D. M. Williams, "Seasonality on terrestrial extrasolar planets: inferring obliquity and surface conditions from infrared light curves," *New Astronomy*, vol. 10, no. 1, pp. 67–77, Nov. 2004.

[57] J. A. Carter and J. N. Winn, "The Detectability of Transit Depth Variations Due to Exoplanetary Oblateness and Spin Precession," *Astrophysical Journal*, vol. 716, no. 1, pp. 850–856, Jun. 2010. arXiv: 1005.1663 `[astro-ph.EP]`.

[58] J. C. Schwartz, C. Sekowski, H. M. Haggard, E. Pallé, and N. B. Cowan, "Inferring planetary obliquity using rotational and orbital photometry," *Monthly Notices of the RAS*, vol. 457, no. 1, pp. 926–938, Mar. 2016. arXiv: 1511.05152 `[astro-ph.EP]`.

[59] M. L. Bryan, E. Chiang, B. P. Bowler, C. V. Morley, S. Millholland, S. Blunt, K. B. Ashok, E. Nielsen, H. Ngo, D. Mawet, and H. A. Knutson, "Obliquity Constraints on an Extrasolar Planetary-mass Companion," *Astronomical Journal*, vol. 159, no. 4, 181, p. 181, Apr. 2020. arXiv: 2002.11131 `[astro-ph.EP]`.

[60] M. L. Bryan, B. Benneke, H. A. Knutson, K. Batygin, and B. P. Bowler, "Constraints on the spin evolution of young planetary-mass companions," *Nature Astronomy*, vol. 2, pp. 138–144, Dec. 2018. arXiv: 1712.00457 `[astro-ph.EP]`.

[61] P. S. Muirhead, A. W. Mann, A. Vanderburg, T. D. Morton, A. Kraus, M. Ireland, J. J. Swift, G. A. Feiden, E. Gaidos, and J. Z. Gazak, "Kepler-445, Kepler-446 and the Occurrence of Compact Multiples Orbiting Mid-M Dwarf Stars," *Astrophysical Journal*, vol. 801, no. 1, 18, p. 18, Mar. 2015. arXiv: 1501.01305 `[astro-ph.SR]`.

[62] W. Zhu, C. Petrovich, Y. Wu, S. Dong, and J. Xie, "About 30% of Sun-like Stars Have Kepler-like Planetary Systems: A Study of Their Intrinsic Architecture," *Astrophysical Journal*, vol. 860, no. 2, 101, p. 101, Jun. 2018. arXiv: 1802.09526 `[astro-ph.EP]`.

[63] E. Bolmont, B. .-O. Demory, S. Blanco-Cuaresma, E. Agol, S. L. Grimm, P. Auclair-Desrotour, F. Selsis, and A. Leleu, "Impact of tides on the transit-timing fits to the TRAPPIST-1 system," *Astronomy and Astrophysics*, vol. 635, A117, A117, Mar. 2020. arXiv: 2002.02015 `[astro-ph.EP]`.

[64] A. M. Vinson, D. Tamayo, and B. M. S. Hansen, "The chaotic nature of TRAPPIST-1 planetary spin states," *Monthly Notices of the RAS*, vol. 488, no. 4, pp. 5739–5747, Oct. 2019. arXiv: 1905.11419 `[astro-ph.EP]`.

[65] S. Millholland and G. Laughlin, "Obliquity-driven sculpting of exoplanetary systems," *Nature Astronomy*, vol. 3, pp. 424–433, Mar. 2019. arXiv: 1903.01386 `[astro-ph.EP]`.

[66] P. Hut, “Tidal evolution in close binary systems.,” *Astronomy and Astrophysics*, vol. 99, pp. 126–140, Jun. 1981.

[67] P. P. Eggleton, L. G. Kiseleva, and P. Hut, “The equilibrium tide model for tidal friction,” *The Astrophysical Journal*, vol. 499, no. 2, p. 853, 1998.

[68] R. A. Mardling and D. N. C. Lin, “Calculating the Tidal, Spin, and Dynamical Evolution of Extrasolar Planetary Systems,” *Astrophysical Journal*, vol. 573, no. 2, pp. 829–844, Jul. 2002.

[69] E. Bolmont, S. N. Raymond, J. Leconte, F. Hersant, and A. C. M. Correia, “Mercury-T: A new code to study tidally evolving multi-planet systems. Applications to Kepler-62,” *Astronomy and Astrophysics*, vol. 583, A116, A116, Nov. 2015. arXiv: 1507.04751 `[astro-ph.EP]`.

[70] S. Blanco-Cuaresma and E. Bolmont, “Studying Tidal Effects In Planetary Systems With Posidonius. A N-Body Simulator Written In Rust.,” in *EWASS Special Session 4 (2017): Star-planet interactions (EWASS-SS4-2017)*, Oct. 2017. arXiv: 1712.01281 `[astro-ph.EP]`.

[71] J. Touma and J. Wisdom, “Lie-Poisson Integrators for Rigid Body Dynamics in the Solar System,” *Astronomical Journal*, vol. 107, p. 1189, Mar. 1994.

[72] S. Breiter, D. Nesvorný, and D. Vokrouhlický, “Efficient Lie-Poisson Integrator for Secular Spin Dynamics of Rigid Bodies,” *Astronomical Journal*, vol. 130, no. 3, pp. 1267–1277, Sep. 2005.

[73] A. Iserles, H. Z. Munthe-Kaas, S. P. Nørsett, and A. Zanna, “Lie-group methods,” *Acta numerica*, vol. 9, pp. 215–365, 2000.

[74] N. Bou-Rabee and J. E. Marsden, “Hamilton–pontryagin integrators on lie groups part i: Introduction and structure-preserving properties,” *Foundations of Computational Mathematics*, vol. 9, no. 2, pp. 197–219, 2009.

[75] E. Celledoni, H. Marthinsen, and B. Owren, “An introduction to lie group integrators–basics, new developments and applications,” *Journal of Computational Physics*, vol. 257, pp. 1040–1061, 2014.

[76] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*, Second. Berlin Heidelberg New York: Springer, 2006.

[77] B. Leimkuhler and S. Reich, *Simulating Hamiltonian dynamics*, ser. Cambridge Monographs on Applied and Computational Mathematics. Cambridge: Cambridge University Press, 2004, vol. 14, pp. xvi+379, ISBN: 0-521-77290-7.

[78] S. Blanes and F. Casas, *A concise introduction to geometric numerical integration*. CRC press, 2017.

[79] J. Sanz-Serna and M. Calvo, *Numerical Hamiltonian problems*, 1st. Chapman and Hall/CRC, 1994.

[80] A. Dullweber, B. Leimkuhler, and R. McLachlan, “Symplectic splitting methods for rigid body molecular dynamics,” *The Journal of chemical physics*, vol. 107, no. 15, pp. 5840–5851, 1997.

[81] F. Fassò, “Comparison of splitting algorithms for the rigid body,” *Journal of computational physics*, vol. 189, no. 2, pp. 527–538, 2003.

[82] R. van Zon and J. Schofield, “Symplectic algorithms for simulations of rigid-body systems using the exact solution of free motion,” *Physical Review E*, vol. 75, no. 5, p. 056 701, 2007.

[83] E. Celledoni, F. Fassò, N. Säfström, and A. Zanna, “The exact computation of the free rigid body motion and its use in splitting methods,” *SIAM Journal on Scientific Computing*, vol. 30, no. 4, pp. 2084–2112, 2008.

[84] G. Vilmart, “Reducing round-off errors in rigid body dynamics,” *Journal of computational physics*, vol. 227, no. 15, pp. 7083–7088, 2008.

[85] T. Lee, M. Leok, and N. H. McClamroch, “Lie group variational integrators for the full body problem in orbital mechanics,” *Celestial Mechanics and Dynamical Astronomy*, vol. 98, no. 2, pp. 121–144, Jun. 2007.

[86] R. I. McLachlan and G. R. W. Quispel, “Splitting methods,” *Acta Numerica*, vol. 11, pp. 341–434, 2002.

[87] S. Blanes, F. Casas, and A. Murua, “Splitting and composition methods in the numerical integration of differential equations,” *arXiv preprint arXiv:0812.0377*, 2008.

[88] S. Blanes, F. Casas, A. Farres, J. Laskar, J. Makazaga, and A. Murua, “New families of symplectic splitting methods for numerical integration in dynamical astronomy,” *Applied Numerical Mathematics*, vol. 68, pp. 58–72, 2013.

[89] E. Hairer, C. Lubich, and G. Wanner, *Geometric numerical integration: structure-preserving algorithms for ordinary differential equations*. Springer Science & Business Media, 2006, vol. 31.

[90] S. Reich and K.-Z. Zentrum, “Symplectic integrators for systems of rigid bodies,” *Fields Inst. Commun*, vol. 10, pp. 181–191, 1996.

[91] J. E. Marsden and T. S. Ratiu, *Introduction to mechanics and symmetry: a basic exposition of classical mechanical systems*. Springer, 1994.

[92] D. Holm, T. Schmah, and C. Stoica, *Geometric mechanics and symmetry: from finite to infinite dimensions*, ser. Oxford texts in applied and engineering mathematics. Oxford University Press, 2009.

[93] T. Lee, N. H. McClamroch, and M. Leok, "A lie group variational integrator for the attitude dynamics of a rigid body with applications to the 3d pendulum," in *Proceedings of 2005 IEEE Conference on Control Applications, 2005. CCA 2005.*, IEEE, 2005, pp. 962–967.

[94] M. Suzuki, "Fractal decomposition of exponential operators with applications to many-body theories and monte carlo simulations," *Physics Letters A*, vol. 146, no. 6, pp. 319–323, 1990.

[95] H. Yoshida, "Construction of higher order symplectic integrators," *Physics letters A*, vol. 150, no. 5-7, pp. 262–268, 1990.

[96] J. Laskar and P. Robutel, "High order symplectic integrators for perturbed hamiltonian systems," *Celestial Mechanics and Dynamical Astronomy*, vol. 80, no. 1, pp. 39–62, 2001.

[97] R. I. McLachlan, "Composition methods in the presence of small parameters," *BIT numerical mathematics*, vol. 35, no. 2, pp. 258–268, 1995.

[98] J. Wisdom and M. Holman, "Symplectic maps for the n-body problem," *The Astronomical Journal*, vol. 102, pp. 1528–1538, 1991.

[99] M. J. Duncan, H. F. Levison, and M. H. Lee, "A multiple time step symplectic algorithm for integrating close encounters," *The Astronomical Journal*, vol. 116, no. 4, p. 2067, 1998.

[100] D. Tamayo, H. Rein, P. Shi, and D. M. Hernandez, "REBOUNDx: a library for adding conservative and dissipative forces to otherwise symplectic N-body integrations," *Monthly Notices of the RAS*, vol. 491, no. 2, pp. 2885–2901, Jan. 2020. arXiv: 1908.05634 `[astro-ph.EP]`.

[101] L. Blanchet, "Gravitational Radiation from Post-Newtonian Sources and Inspiralling Compact Binaries," *Living Reviews in Relativity*, vol. 9, no. 1, 4, p. 4, Jun. 2006.

[102] J. D. Anderson, P. B. Esposito, W. Martin, C. L. Thornton, and D. O. Muhleman, "Experimental test of general relativity using time-delay data from Mariner 6 and Mariner 7.," *Astrophysical Journal*, vol. 200, pp. 221–233, Aug. 1975.

[103] F. Benitez and T. Gallardo, “The relativistic factor in the orbital dynamics of point masses,” *Celestial Mechanics and Dynamical Astronomy*, vol. 101, no. 3, pp. 289–307, Jul. 2008. arXiv: 0709.1160 `[astro-ph]`.

[104] M. Tao and T. Ohsawa, “Variational optimization on lie groups, with examples of leading (generalized) eigenvalue problems,” in *International Conference on Artificial Intelligence and Statistics*, PMLR, 2020, pp. 4269–4280.

[105] S. Kreyche, J. Barnes, B. Quarles, and J. Chambers, “Exploring tidal obliquity variations with SMERCURY-T,” *Planetary Science Journal*, Submitted.

[106] J. E. Chambers, “A hybrid symplectic integrator that permits close encounters between massive bodies,” *Monthly Notices of the RAS*, vol. 304, no. 4, pp. 793–799, Apr. 1999.

[107] J. Giorgini, D. Yeomans, A. Chamberlin, P. Chodas, R. Jacobson, M. Keesey, J. Lieske, S. Ostro, E. Standish, and R. Wimberly, “Jpl’s on-line solar system data service,” in *AAS/Division for Planetary Sciences Meeting Abstracts# 28*, vol. 28, 1996, pp. 25–04.

[108] P. P. Eggleton and L. Kiseleva-Eggleton, “Orbital Evolution in Binary and Triple Stars, with an Application to SS Lacertae,” *Astrophysical Journal*, vol. 562, no. 2, pp. 1012–1030, Dec. 2001. arXiv: astro-ph/0104126 `[astro-ph]`.

[109] D. Fabrycky and S. Tremaine, “Shrinking Binary and Planetary Orbits by Kozai Cycles with Tidal Friction,” *Astrophysical Journal*, vol. 669, no. 2, pp. 1298–1315, Nov. 2007. arXiv: 0705.4285 `[astro-ph]`.

[110] T. Van Hoolst, N. Rambaux, Ö. Karatekin, V. Dehant, and A. Rivoldini, “The librations, shape, and icy shell of Europa,” *Icarus*, vol. 195, no. 1, pp. 386–399, May 2008.

[111] E. Agol and D. C. Fabrycky, “Transit-Timing and Duration Variations for the Discovery and Characterization of Exoplanets,” in *Handbook of Exoplanets*, H. J. Deeg and J. A. Belmonte, Eds. 2018.

[112] J. Miralda-Escudé, “Orbital Perturbations of Transiting Planets: A Possible Method to Measure Stellar Quadrupoles and to Detect Earth-Mass Planets,” *Astrophysical Journal*, vol. 564, no. 2, pp. 1019–1023, Jan. 2002. arXiv: astro-ph/0104034 `[astro-ph]`.

[113] J. S. Heyl and B. J. Gladman, “Using long-term transit timing to detect terrestrial planets,” *Monthly Notices of the RAS*, vol. 377, no. 4, pp. 1511–1519, Jun. 2007. arXiv: astro-ph/0610267 `[astro-ph]`.

[114] D. Ragozzine and A. S. Wolf, "Probing the Interiors of very Hot Jupiters Using Transit Light Curves," *Astrophysical Journal*, vol. 698, no. 2, pp. 1778–1794, Jun. 2009. arXiv: 0807.2856 `[astro-ph]`.

[115] G. Maciejewski, M. Fernández, F. Aceituno, S. Martıén-Ruiz, J. Ohlert, D. Dimitrov, K. Szyszka, C. von Essen, M. Mugrauer, R. Bischoff, K. .-U. Michel, M. Mallonn, M. Stangret, and D. Moździerski, "Planet-Star Interactions with Precise Transit Timing. I. The Refined Orbital Decay Rate for WASP-12 b and Initial Constraints for HAT-P-23 b, KELT-1 b, KELT-16 b, WASP-33 b and WASP-103 b," *ACTA ASTRONOMICA*, vol. 68, no. 4, pp. 371–401, Dec. 2018. arXiv: 1812.02438 `[astro-ph.EP]`.

[116] S. L. Grimm, B.-O. Demory, M. Gillon, C. Dorn, E. Agol, A. Burdanov, L. Delrez, M. Sestovic, A. H. M. J. Triaud, M. Turbet, É. Bolmont, A. Caldas, J. de Wit, E. Jehin, J. Leconte, S. N. Raymond, V. Van Grootel, A. J. Burgasser, S. Carey, D. Fabrycky, K. Heng, D. M. Hernandez, J. G. Ingalls, S. Lederer, F. Selsis, and D. Queloz, "The nature of the TRAPPIST-1 exoplanets," *Astronomy and Astrophysics*, vol. 613, A68, A68, May 2018. arXiv: 1802.01377 `[astro-ph.EP]`.

[117] E. Agol, C. Dorn, S. L. Grimm, M. Turbet, E. Ducrot, L. Delrez, M. Gillon, B.-O. Demory, A. Burdanov, K. Barkaoui, Z. Benkhaldoun, E. Bolmont, A. Burgasser, S. Carey, J. de Wit, D. Fabrycky, D. Foreman-Mackey, J. Haldemann, D. M. Hernandez, J. Ingalls, E. Jehin, Z. Langford, J. Leconte, S. M. Lederer, R. Luger, R. Malhotra, V. S. Meadows, B. M. Morris, F. J. Pozuelos, D. Queloz, S. M. Raymond, F. Selsis, M. Sestovic, A. H. M. J. Triaud, and V. Van Grootel, "Refining the transit timing and photometric analysis of TRAPPIST-1: Masses, radii, densities, dynamics, and ephemerides," *arXiv e-prints*, arXiv:2010.01074, arXiv:2010.01074, Oct. 2020. arXiv: 2010.01074 `[astro-ph.EP]`.

[118] J. F. Kasting, D. P. Whitmire, and R. T. Reynolds, "Habitable Zones around Main Sequence Stars," *Icarus*, vol. 101, no. 1, pp. 108–128, Jan. 1993.

[119] M. M. Joshi, R. M. Haberle, and R. T. Reynolds, "Simulations of the Atmospheres of Synchronously Rotating Terrestrial Planets Orbiting M Dwarfs: Conditions for Atmospheric Collapse and the Implications for Habitability," *Icarus*, vol. 129, no. 2, pp. 450–465, Oct. 1997.

[120] R. Wordsworth, "Atmospheric Heat Redistribution and Collapse on Tidally Locked Rocky Planets," *Astrophysical Journal*, vol. 806, no. 2, 180, p. 180, Jun. 2015. arXiv: 1412.5575 `[astro-ph.EP]`.

[121] D. Tamayo, H. Rein, C. Petrovich, and N. Murray, "Convergent Migration Renders TRAPPIST-1 Long-lived," *Astrophysical Journal, Letters*, vol. 840, no. 2, L19, p. L19, May 2017. arXiv: 1704.02957 `[astro-ph.EP]`.

[122] J. Li, D. Lai, K. R. Anderson, and B. Pu, "Giant planet scatterings and collisions: hydrodynamics, merger-ejection branching ratio, and properties of the remnants," *Monthly Notices of the RAS*, vol. 501, no. 2, pp. 1621–1632, Feb. 2021. arXiv: 2006.10067 `[astro-ph.EP]`.

[123] H. Goldstein, *Classical mechanics*, 1980.

[124] G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.

[125] H. Abarbanel, *Analysis of observed chaotic data*. Springer Science & Business Media, 2012.

[126] H. Kantz and T. Schreiber, *Nonlinear time series analysis*. Cambridge university press, 2004, vol. 7.

[127] E. Bradley and H. Kantz, "Nonlinear time-series analysis revisited," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 25, no. 9, p. 097 610, 2015.

[128] E. Baake, M. Baake, H. Bock, and K. Briggs, "Fitting ordinary differential equations to chaotic data," *Physical Review A*, vol. 45, no. 8, p. 5524, 1992.

[129] J. Bongard and H. Lipson, "Automated reverse engineering of nonlinear dynamical systems," *Proceedings of the National Academy of Sciences*, vol. 104, no. 24, pp. 9943–9948, 2007.

[130] M. Schmidt and H. Lipson, "Distilling free-form natural laws from experimental data," *science*, vol. 324, no. 5923, pp. 81–85, 2009.

[131] S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Discovering governing equations from data by sparse identification of nonlinear dynamical systems," *Proceedings of the national academy of sciences*, vol. 113, no. 15, pp. 3932–3937, 2016.

[132] G. Tran and R. Ward, "Exact recovery of chaotic systems from highly corrupted data," *Multiscale Modeling & Simulation*, vol. 15, no. 3, pp. 1108–1129, 2017.

[133] H. Schaeffer, G. Tran, and R. Ward, "Extracting sparse high-dimensional dynamics from limited data," *SIAM Journal on Applied Mathematics*, vol. 78, no. 6, pp. 3279–3295, 2018.

[134] F. Lu, M. Zhong, S. Tang, and M. Maggioni, "Nonparametric inference of interaction laws in systems of agents from trajectory data," *Proceedings of the National Academy of Sciences*, vol. 116, no. 29, pp. 14 424–14 433, 2019.

[135] S. H. Rudy, S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Data-driven discovery of partial differential equations," *Science Advances*, vol. 3, no. 4, e1602614, 2017.

[136] S. H. Kang, W. Liao, and Y. Liu, "IDENT: Identifying differential equations with numerical time evolution," *Journal of Scientific Computing*, 2021.

[137] P. A. Reinbold, L. M. Kageorge, M. F. Schatz, and R. O. Grigoriev, "Robust learning from noisy, incomplete, high-dimensional experimental data via physically constrained symbolic regression," *Nature Communications*, vol. 12, no. 1, pp. 1–8, 2021.

[138] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Multistep neural networks for data-driven discovery of nonlinear dynamical systems," *arXiv preprint arXiv:1801.01236*, 2018.

[139] S. H. Rudy, J. N. Kutz, and S. L. Brunton, "Deep learning of dynamics and signal-noise decomposition with time-stepping constraints," *Journal of Computational Physics*, vol. 396, pp. 483–506, 2019.

[140] T. Qin, K. Wu, and D. Xiu, "Data driven governing equations approximation using deep neural networks," *Journal of Computational Physics*, vol. 395, pp. 620–635, 2019.

[141] Z. Long, Y. Lu, X. Ma, and B. Dong, "Pde-net: Learning pdes from data," in *International Conference on Machine Learning*, PMLR, 2018, pp. 3208–3216.

[142] M. Raissi and G. E. Karniadakis, "Hidden physics models: Machine learning of nonlinear partial differential equations," *Journal of Computational Physics*, vol. 357, pp. 125–141, 2018.

[143] C. A. Bailer-Jones, D. J. MacKay, and P. J. Withers, "A recurrent neural network for modelling dynamical systems," *network: computation in neural systems*, vol. 9, no. 4, pp. 531–547, 1998.

[144] Y. Wang, "A new concept using lstm neural networks for dynamic system identification," in *2017 American Control Conference (ACC)*, IEEE, 2017, pp. 5324–5329.

[145] J. Pathak, B. Hunt, M. Girvan, Z. Lu, and E. Ott, "Model-free prediction of large spatiotemporally chaotic systems from data: A reservoir computing approach," *Physical review letters*, vol. 120, no. 2, p. 024 102, 2018.

[146] S. Mukhopadhyay and S. Banerjee, "Learning dynamical systems in noise using convolutional neural networks," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 30, no. 10, p. 103 125, 2020.

[147] A. Shalova and I. Oseledets, “Tensorized transformer for dynamical systems modeling,” *arXiv preprint arXiv:2006.03445*, 2020.

[148] M. Raissi, P. Perdikaris, and G. E. Karniadakis, “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations,” *Journal of Computational Physics*, vol. 378, pp. 686–707, 2019.

[149] S. Greydanus, M. Dzamba, and J. Yosinski, “Hamiltonian neural networks,” *NeurIPS*, 2019.

[150] T. Bertalan, F. Dietrich, I. Mezić, and I. G. Kevrekidis, “On learning hamiltonian systems from data,” *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 29, no. 12, p. 121 107, 2019.

[151] Z. Chen, J. Zhang, M. Arjovsky, and L. Bottou, “Symplectic recurrent neural networks,” in *International Conference on Learning Representations*, 2020.

[152] P. Jin, Z. Zhang, A. Zhu, Y. Tang, and G. E. Karniadakis, “SympNets: Intrinsic structure-preserving symplectic networks for identifying Hamiltonian systems,” *Neural Networks*, vol. 132, pp. 166–179, 2020.

[153] M. Lutter, C. Ritter, and J. Peters, “Deep lagrangian networks: Using physics as model prior for deep learning,” in *International Conference on Learning Representations*, 2019.

[154] P. Toth, D. J. Rezende, A. Jaegle, S. Racanière, A. Botev, and I. Higgins, “Hamiltonian generative networks,” in *International Conference on Learning Representations*, 2020.

[155] Y. D. Zhong, B. Dey, and A. Chakraborty, “Symplectic ode-net: Learning hamiltonian dynamics with control,” in *International Conference on Learning Representations*, 2020.

[156] K. Wu, T. Qin, and D. Xiu, “Structure-preserving method for reconstructing unknown hamiltonian systems from trajectory data,” *SIAM Journal on Scientific Computing*, vol. 42, no. 6, A3704–A3729, 2020.

[157] S. Xiong, Y. Tong, X. He, S. Yang, C. Yang, and B. Zhu, “Nonseparable symplectic neural networks,” *ICLR*, 2021.

[158] S.-T. Tsai, E.-J. Kuo, and P. Tiwary, “Learning molecular dynamics with simple language model built upon long short-term memory neural network,” *Nature Communications*, vol. 11, no. 1, pp. 1–11, 2020.

[159] A. C. Da Silva, *Lectures on symplectic geometry*. Springer, 2008, vol. 3575.

[160] V. I. Arnol'd, *Mathematical methods of classical mechanics*. Springer Science & Business Media, 2013, vol. GTM 60.

[161] K. Hornik, M. Stinchcombe, and H. White, "Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks," *Neural networks*, vol. 3, no. 5, pp. 551–560, 1990.

[162] D. Yarotsky, "Error bounds for approximations with deep relu networks," *Neural Networks*, vol. 94, pp. 103–114, 2017.

[163] K. T. Alligood, T. D. Sauer, and J. A. Yorke, *Chaos*. Springer, 1996.

[164] E. Sander and J. A. Yorke, "The many facets of chaos," *International Journal of Bifurcation and Chaos*, vol. 25, no. 04, p. 1 530 011, 2015.

[165] M. Hénon and C. Heiles, "The applicability of the third integral of motion: Some numerical experiments," *The Astronomical Journal*, vol. 69, p. 73, 1964.

[166] J. Pöschel, "Integrability of Hamiltonian systems on cantor sets," *Communications on Pure and Applied Mathematics*, vol. 35, no. 5, pp. 653–696, 1982.

[167] B. Quarles, G. Li, V. Kostov, and N. Haghighipour, "Orbital stability of circumstellar planets in binary systems," *The Astronomical Journal*, vol. 159, no. 3, p. 80, 2020.

[168] M. Tao, "Explicit symplectic approximation of nonseparable hamiltonians: Algorithm and long time performance," *Physical Review E*, vol. 94, no. 4, p. 043 303, 2016.

[169] B. V. Chirikov, "A universal instability of many-dimensional oscillator systems," *Physics reports*, vol. 52, no. 5, pp. 263–379, 1979.

[170] E. M. Stein and R. Shakarchi, *Complex analysis*. Princeton University Press, 2010, vol. 2.