On the Painlevé conjecture

Jinxin Xue

Abstract. In this paper, we review the circle of ideas around Painlevé conjecture–the existence of noncollision singularities i.e. noncollision finite time blowup solutions, in Newtonian N-body problem.

Contents

1. Introduction

The Newtonian N-body problem studies the motion of N particles $Q_i \in \mathbb{R}^3$ with mass $m_i,\ i = 1, \dots, n$ under the mutual gravitational forces

$$m_i \ddot{Q}_i = -\sum_{j \neq i} G m_i m_j \frac{Q_i - Q_j}{|Q_i - Q_j|^3}, \quad i = 1, \dots, N,$$

where G is the gravitational constant which will be normalized to be 1 in the paper. The system of equations has a Hamiltonian structure. It can be written as the Hamiltonian equations

$$\begin{cases} \dot{Q}_i &= \frac{\partial H}{\partial P_i} \\ \dot{P}_i &= -\frac{\partial H}{\partial Q_i} \end{cases}$$

of the Hamiltonian

$$H(P,Q)=\sum_{i=1}^{N}\frac{|P_i|^2}{2m_i}+U(Q),\quad U(Q)=-\sum_{i\neq j}\frac{m_im_j}{|Q_i-Q_j|},$$

where $P=(P_1,\dots,P_N)$, $Q=(Q_1,\dots,Q_N)$.

This is a significant nonlinear ODE system. It models the motion of the planets in our solar system, therefore its long time dynamics is of particular importance since the fate of our civilization is directly relevant. Newton solved the two-body problem and also realized that the three-body problem was very complicated. However, in more than 100 years after Newton, people still believed that the three-body problem could be "solved" in a similar manner as integrating the two-body problem. This belief of strict determinism was spoken out by Laplace known as "Laplace demon". The belief was broken by Poincaré who discovered a mechanism (homoclinic tangle) responsible for the nonintegrability and concluded that the Newtonian N-body problem is in general chaotic for $N>2$. From Poincaré's discovery of chaos, Smale introduced the horseshoe and symbolic dynamics hence gave birth to the modern theory of dynamical systems. There are many other important mathematics originated from Poincaré work on celestial mechanics, for instance, Poincaré last geometric theorem proved by Birkhoff leading to Arnold conjecture and the birth of sympletic geometry, the averaging method leading to KAM theory, the Poincaré recurrence leading to Furstenberg's multiple recurrence with applications to number theory, etc. So we have seen that the Newtonian N-body problem, as the old field of dynamics, played a significant role in the development of the field of dynamical systems and even had influences to other fields of mathematics. However, our current understanding of the dynamics of the Newtonian N-body problem is still very limited. There are a number of deep open problems that are far from being answered.

1.1. Global dynamics of Newtonian N-body problem. To fit our result into a broad global picture, we first give an overview of the global dynamics of Newtonian N-body problem. Our perspective is adopted from the general method of studying differential equations. The problems and conjectures that we list below are chosen centered around the main theme of the paper–noncollision singularity, hence is nonavoidably biased.

With the wellposedness problem understood, the main theme is to understand the long time dynamics of the system. The set of initial conditions in $\mathbb{R}^{6N}$ can be divided into two disjoint sets $\mathcal{S}$ and $\mathcal{R}$, where $\mathcal{R}$ meaning regular is the set of initial conditions such that the solution is defined on entire $\mathbb{R}$, and $\mathcal{S}$ meaning singular is the complement of $\mathcal{R}$. We can further classify $\mathcal{S}$ into two disjoint subsets $\mathcal{CS}$ and $\mathcal{NCS}$ where $\mathcal{CS}$ meaning collision singularities, is the set of initial conditions leading to a collision either forward or backward in time and $\mathcal{NCS}$ meaning noncollision singularities is the set of singularities without collisions.

Here comes the first main conjecture concerning the global dynamics of the Newtonian N-body problem.

CONJECTURE 1.1. *The set $\mathcal{S}$ has zero Lebesgue measure.*

This conjecture can be found in [**32**] as the first problem. It is known that $\mathcal{CS}$ has zero Lebesgue measure and $\mathcal{NCS}$ has zero Lebesgue measure for $N = 4$ [**31**]. However, the general $N > 4$ case the conjecture is widely open. We will elaborate on the subject of noncollision singularities starting from Section 1.2. In general, this conjecture seems very hard. To study the topological instability of the global dynamics, we may also ask the following easier companion of Conjecture 1.1.

CONJECTURE 1.2. *The set $\mathcal{S}$ is of Baire first category.*

It was known to Saari that $\mathcal{CS}$ is of first category, so again the problem is to prove that $\mathcal{NCS}$ is of first category.

PROBLEM 1.3. *Classify $\mathcal{CS}$ and $\mathcal{NCS}$.*

We will see in Appendix A that the orbit of a noncollision singularity will approach $\mathcal{CS}$ repeatedly as time approaches the singular time. So the first step towards the problem is to understand $\mathcal{CS}$. When approaching the total collapse of an N-body problem, we can blow up the dynamics to arrive at a limiting dynamical system on a collision manifold. The blowup technique will be introduced in Section 5. A special type of solution called central configurations appeared naturally as equilibria of the limiting dynamics. Denoting by $Q = (Q_1, \dots, Q_N)$ the position vector, by $U = -\sum_{i \neq j} \frac{m_i m_j}{|Q_i - Q_j|}$ the Newtonian potential and by $M = \operatorname{diag}\{m_1, \dots, m_N\}$ the mass matrix, and setting the mass center $Q_c := \sum m_i Q_i = 0$, a central configuration is an arrangement of the point masses satisfying

$$\nabla U(Q) = \lambda M Q$$

for some λ, so central configurations are critical points of U restricted to the shape sphere $\{Q^T M Q = 1\}$. For a central configuration, the Newtonian equation is simplified into $\ddot{Q} = -\lambda Q$, hence central configurations can be used to construct self-similar solutions. Thus the problem of classifying $\mathcal{CS}$ leads naturally to classifying the central configurations. Related to this, there is the following conjecture by Smale [**34**].

CONJECTURE 1.4. *For all mass ratios and all $N \geq 3$, the number of central configurations is finite.*

The central configurations in the three-body problem were known to Euler and Lagrange. There are only collinear (Euler) and equilateral (Lagrange) configurations. The problem is solved for the cases $N = 4, 5$ for generic mass ratios [**4**, **19**, **20**] and open in general. For more open problems concerning central configuration, we refer to the problem list of [**3**]. However, the finiteness still does not imply a classification and there are so

many various central configurations that a classification seems impossible. We will perform the total collapse blowup and study the dynamics on the limiting collision manifold in Section 5 where we will see that the dynamics on the collision manifold is gradient like with central configurations as fixed points. We call a central configuration *stable* if it is a sink on the collision manifold. In the $N = 3$ case the Euler configurations correspond to a sink and a source and the Lagrange configurations are saddles. The only stable one is the sink. From the perspective of studying the generic global dynamics, the following problem seems more relevant and tractable.

PROBLEM 1.5. *Classify all the stable central configurations.*

We next consider the global dynamics of points in $\mathcal{R}$. One of the main theme in dynamical system is to understand the asymptotic behavior as $t \to \pm\infty$, so the following problem is of interest.

PROBLEM 1.6. *Classify all the final motions for point $x \in \mathcal{R}$ for $t \to \pm\infty$.*

A classification was given by Chazy for three-body problem [**5**], in which all logical possible combinations ($\mathcal{H}$, $\mathcal{P}$, $\mathcal{B}$, $\mathcal{HP}$, $\mathcal{HB}$, $\mathcal{PB}$, $\mathcal{OS}$) is realized by an orbit, where $\mathcal{H}$ means hyperbolic, i.e. $|Q_k(t)| \to \infty$, $|\dot{Q}_k(t)| \to c_k > 0$ for $k = 1, 2, 3$, $\mathcal{P}$ means parabolic, i.e. $|Q_k(t)| \to \infty$, $|\dot{Q}_k(t)| \to 0$ for $k = 1, 2, 3$, $\mathcal{B}$ means bounded, i.e. $\sup_t \max_k |Q_k(t)| < \infty$, $\mathcal{OS}$ means oscillatory, i.e. $\limsup |Q(t)| = \infty$, $\liminf |Q(t)| < \infty$, and $\mathcal{HP}$ means hyperbolic-parabolic i.e. some body is hyperbolic and some other body is parabolic asymptotically, similarly for $\mathcal{HB}$ and $\mathcal{PB}$ (c.f. Chapter 2.3.4 of [**5**] for more details). There are many different special bounded motions such as periodic orbits, quasi-periodic orbits, etc. Among others, we refer to [**8**] for the remarkable figure-8 periodic orbit in three-body problem and periodic orbits realizing all free homotopy classes in planar three-body problem by Moeckel and Montgomery [**26**]. Bounded motions can even have positive measure, for instance, quasi-periodic orbits in KAM theory. The application of the KAM theory to the N-body problem was initiated in the work of Arnold [**1**], cultivated by Herman, Fejoz [**12**] and completed in the work of Chierchia-Pinzari [**9**]. However, KAM tori are only of first category. We see that a classification of final motions for general N-body problem in the spirit of Chazy may be very complicated, even modulo a zero measure set. So it is reasonable to ask which case is the predominant case of final motions for topological generic initial conditions.

Recall that the notion of nonwandering point in topological dynamical systems. Let $f : X \to X$ be a homeomorphism on a topological space X. A point $x \in X$ is called *nonwandering* if for every neighborhood U of x, there exists $n > 0$, such that $f^n(U) \cap U \neq \emptyset$.

CONJECTURE 1.7. *Fir all energy E, the non wandering set of the Hamiltonian flow on the energy level set $H^{-1}(E)$ is nowhere dense.*

This conjecture can be found in [17], where Herman attributed it to Birkhoff and Kolmogorov and called it *the oldest problem in Newtonian N-body problem*. The conjecture asserts the topological instability of the N-body problem, in contrast to the positive measure quasiperiodic motions given by KAM theory. Moreover, Herman in [17] also gives a more tractable version assuming that there is one large mass and all other particles have masses of order ε.

In a similar spirit, we make the following conjecture on the generic global dynamics in Newtonian N-body problem. This conjecture can be considered as an analogue of the soliton resolution conjecture in dispersive PDEs. Supporting evidences can be found in the numerical studies [2, 36, 35, 18] etc.

CONJECTURE 1.8. *For generic initial condition, the solution is defined on $\mathbb{R}$, and as $t \to \pm\infty$, each body in the system approaches either a linear motion with constant velocity or a Kepler elliptic motion around a center moving linearly with constant velocity.*

The conjecture anticipates that the N-body problem with $N > 3$ is such a highly nonlinear system that the long time dynamics tends to decouple the system into two-body problems. One result in this direction can be found in [27], where the authors proved that for N-body problem as $t \to \infty$, there is a dichotomy: either it is a superhyperbolic orbit; or it decouples into several subsystems moving apart linearly and each subsystem grows at most like $O(t^{2/3})$. Here a super-hyperbolic means that $\sqrt{I}(t)/t \to \infty$ as $t \to \infty$, where I is the momentum of inertia measuring the diameter of the system as defined in (A.1). Here $O(t^{2/3})$ growth is a parabolic behavior, which is expected to be of first category, so it is hopeful to improve it to bounded motion for generic initial conditions. The existence of (noncollisional) superhyperbolic is still an open problem. We can consider superhyperbolic orbits as a slow version of noncollision singularities, so showing that superhyperbolic orbits are of first category seems a similar problem to that of $\mathcal{NCS}$. Understanding noncollision singularities better is also important for understanding the global dynamics of the N-body problem.

1.2. Noncollision singularities in Newtonian N-body problem. The main theme of the present paper is to elaborate on Conjecture 1.1. We have defined noncollision singularities, but even their existence is a highly nontrivial fact.

CONJECTURE 1.9 (Painlevé, 1897). *The set $\mathcal{NCS}$ is nonempty for $N > 3$.*

The conjecture appeared first in the lecture notes of Painlevé in 1897 [28, 29], which recorded the his lectures at the University of Stockholm invited by King Oscar II of Sweden and Norway in 1895. In these lectures, Painlevé talked about the Newtonian N-body problem with a particular emphasis on the role played by singularities. He proved that in three-body

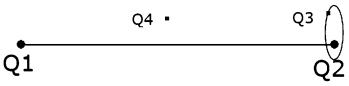

FIGURE 1. The configuration of the four-body problem

problem the only singularities are collisions and made the conjecture stated above. We will present Painlevé's theorem in Appendix A.2.

The next important work after Painlevé is the following characterization of the dynamical behavior of noncollision singularities by von Zeipel [43].

THEOREM 1.10 (von Zeipel, 1908). *If a noncollision singularity occurs at time t^*, then we have*

$$\lim_{t\to t^*} \min_{i\neq j} |Q_i(t) - Q_j(t)| \to 0, \quad \lim_{t\to t^*} \max_{i\neq j} |Q_i(t) - Q_j(t)| \to \infty.$$

We also present the proof of this theorem in Appendix A.3. From von Zeipel's theorem, we see that a noncollision singularity has nonlocal behavior and is highly counter-intuitive. Such wild behavior requires infinitely large velocities hence distinguishes the Newtonian mechanics from special relativity.

The existence of noncollision singularities had been doubted for a long time until the work of Mather-McGehee in 1975 [25]. They constructed a collinear four-body problem in which there exist initial conditions whose solutions satisfy the above characterization of von Zeipel. However, this work has a defect since it involves infinitely many double collisions before the finite singular time, which is not avoidable for a collinear problem, so it does not give a genuine noncollision singularity. Xia constructed noncollision singularities in a spatial five-body problem in 1992 [42]. While the works of Mather-McGehee and Xia use the mechanism of energy extraction from near triple collisions, Gerver pioneered another approach of extracting energy from Kepler elliptic motions. Gerver proposed a model of five-body problem without detailed proofs in 1984 [16]. Later he imposed symmetries to obtain a $3N$-body problem for N sufficiently large in which he proved the existence of noncollision singularities in 1991 [15].

It is a general belief that when a noncollision singularity is known in an N-body problem, we can try to add another remote body to get a noncollision singularity in an $(N+1)$-body problem. So the last hardest case is to prove the existence of noncollision singularities for the borderline case $N = 4$. Moreover, since all the previous works [25, 42, 15] rely crucially on certain symmetry in the model, which is easily broken by various perturbations such as adding an extra body, it is desirable to have noncollision singularities without assuming symmetry.

In [40], we prove that noncollision singularities exist in a four-body problem without symmetry assumption. Let us state the result as follows.

In the model, we have two large bodies Q_1 and Q_2 of masses 1 and two small bodies Q_3 and Q_4 of masses $\mu \ll 1$. The two large bodies Q_1 and Q_2

are set to be far away initially. One small body, say Q_3, is captured by Q_2 and doing almost Kepler elliptic motion. The other small body Q_4 travels back and forth between Q_1 and the pair Q_2-Q_3. Each time when Q_4 comes close to the pair, it has a close encounter with Q_3 to extract certain amount of angular momentum or energy. We will exhibit a rich variety of singular solutions. Fix a small ε_0. Let $\boldsymbol{\omega} = \{\omega_j\}_{j=1}^{\infty}$ be a sequence of 3s and 4s.

DEFINITION 1.11. We say that $(Q_i(t), \dot{Q}_i(t))$, $i = 1, 2, 3, 4$, is a **singular solution with symbolic sequence $\boldsymbol{\omega}$** if there exists a positive increasing sequence $\{t_j\}_{j=0}^{\infty}$ such that

- $t^* = \lim_{j\to\infty} t_j < \infty$.
- $|Q_3 - Q_2|(t_j) \le \varepsilon_0$, $|Q_4 - Q_2|(t_j) \le \varepsilon_0$.
- For $t \in [t_{j-1}, t_j]$, $|Q_{7-\omega_j} - Q_2|(t) \le \varepsilon_0$ and $\{Q_{\omega_j}(t)\}_{t\in[t_{j-1},t_j]}$ leaves the ε_0 neighborhood of Q_2, winds around Q_1 exactly once, then reenters the ε_0 neighborhood of Q_2.
- $\limsup_t |\dot{Q}_i(t)|$, $\limsup_t |Q_i(t)| \to \infty$ as $t \to t^*$, $i = 1, 2, 3, 4$.

During the time interval $[t_{j-1}, t_j]$ we refer to Q_{ω_j} as the traveling particle and to $Q_{7-\omega_j}$ as the captured particle. Thus ω_j prescribes which particle is the traveler during the jth trip.

We denote by $\Sigma_{\boldsymbol{\omega}}$ the set of initial conditions of singular orbits with symbolic sequence $\boldsymbol{\omega}$.

THEOREM 1.12. *There exists $\mu_* \ll 1$ such that for $\mu < \mu_*$ the set $\Sigma_{\boldsymbol{\omega}} \neq \emptyset$.*

Moreover there is an open set U on the zero energy level and zeroth angular momentum level, and a foliation of U by two-dimensional surfaces such that for any leaf S of our foliation $\Sigma_{\boldsymbol{\omega}} \cap S$ is a Cantor set.

In [41], we introduced a simplified model that we call two-center-two-body problem: We fix the two large masses Q_1 and Q_2 and consider the motion of the two small masses Q_3 and Q_4 in the gravitational field generated by all the four bodies. We prove that there exists orbit such that the velocities $|\dot{Q}_3|$ and $|\dot{Q}_4|$ are accelerated to infinity within finite time avoiding all early collisions. The simplified model captured most of the mathematical difficulties of the proof of Theorem 1.12. Note that the Euler two-center problem, i.e. a particle moves in the gravitational field generated by two fixed masses, is a classical integrable system. Our theorem in [41] shows that adding one more particle to the two-center system creates drastically different dynamical behaviors.

From von Zeipel's theorem, we see that noncollision singularities have remarkable nonlocal properties in contrast to most other singularities. To form a noncollision singularity, the system has to pass arbitrarily close to collisions and also gets arbitrarily far from each other repeatedly. So to construct a noncollision singularity, we need a local model near collision to extract energy, and also a mechanism for globalization so that consecutive visits to close encounters are realized by an orbit. In our construction of

noncollision singularities, we use as local model an ideal model of Gerver that achieves self-similarity after a two-step procedure (see Section 3) and we develop a framework for the globalization using hyperbolic dynamics. Our work also reveals a crucial difference between collision and noncollision singularities. Blowing up a collision singularity gives rise to a central configuration as we have seen in Section 1.1. However, for any blowup sequence along a noncollision singularity constructed in Theorem 1.12, the system is nowhere close to a central configuration. We believe that this fact deserves further attention.

The paper is organized as follows. In Section 2, we introduce the classical results on the two-body problem, including solving the two-body problem by quadrature, Kepler's three laws and the Delaunay coordinates. In Section 3, we explain Gerver's ideal model in [**13**] used in the proof of Theorem 1.12. In Section 4, we explain the proof of Theorem 1.12. In Section 5, we explain the mechanism of energy extraction from near triple collisions. We include the total collapse blowup techniques, apply it to several cases of three-body problems and explain briefly some ideas of [**25**] and [**42**]. Finally, in Appendix A, we give the proofs of Painlevé's theorem on the nonexistence of noncollision singularities for three-body problem and von Zeipel's Theorem 1.10.

2. The two-body problem

In this section we consider the two-body problem. We will solve the two-body problem by quadrature and recover Kepler's three laws from the solution. Moreover, we will introduce the action-angle coordinates called Delaunay coordinates, which will be used intensively later.

2.1. The two-body problem in Cartesian and polar coordinates. We suppose the mass center is the origin $m_1Q_1 + m_2Q_2 = 0$, which implies $P_1 + P_2 = m_1\dot{Q}_1 + m_2\dot{Q}_2 = 0$. We can thus introduce $q = Q_2 - Q_1$ as the relative position and $p = P_2$ and we see that the symplectic form

$$dP_1 \wedge dQ_1 + dP_2 \wedge dQ_2 = dp \wedge dq + d(P_1 + P_2) \wedge dQ_1$$

is preserved by the map $(P_1, Q_1, P_2, Q_2) \mapsto (p, q, P_1 + P_2, Q_1)$. In the new coordinates, the Hamiltonian system is reduced to

$$H(q,p) = \frac{p^2}{2m} - \frac{k}{|q|}, \quad (q,p) \in \mathbb{R}^3 \times \mathbb{R}^3, \tag{2.1}$$

where we have $m = \frac{m_1 m_2}{m_1+m_2}$ and $k = m_1 m_2$. The equation of motion has the form

$$m\ddot{q} = -k\frac{q}{|q|^3}.$$

It is easy to verify that the angular momentum $\Theta = q \times p$ is conserved. This implies that the orbit lies on a plane perpendicular to the angular momentum vector. So in the following we consider $q, p \in \mathbb{R}^2$.

We next introduce the polar coordinate transformation $(q,p)\mapsto(r,R,\theta,\Theta)$

$$\begin{cases} q &= (r\cos\theta, r\sin\theta), \\ p &= (R\cos\theta - \frac{\Theta}{r}\sin\theta, R\sin\theta + \frac{\Theta}{r}\cos\theta), \end{cases} \tag{2.2}$$

where r has the geometric meaning of length and θ is the polar angle, R is the radial momentum conjugate to r and $\Theta = q\times p$ is the angular momentum. It can be verified that the following symplectic form is preserved:

$$dp\wedge dq = dR\wedge dr + d\Theta\wedge d\theta.$$

The transformation is obtained as follows. First, we introduce polar coordinates for the position q. For the momentum, it is known that a linear transform $A:\ \mathbb{R}^n\to\mathbb{R}^n$, $A\in \mathrm{GL}_n\mathbb{R}$ induces a symplectic transform $(A,(A^T)^{-1})$ on $\mathbb{R}^n\times\mathbb{R}^n$. So the above symplectic transform in the momentum space is obtained by taking the transpose inverse of the derivative of the polar coordinate change in the position space.

In polar coordinates, the Hamiltonian has the form

$$H(R,r,\Theta,\theta) = \frac{1}{2m}\left(R^2 + \frac{\Theta^2}{r^2}\right) - \frac{k}{r}.$$

2.2. Solving the two-body problem. Now the angular momentum Θ is constant of motion. For fixed $\Theta\neq 0$, the Hamiltonian system has one degree of freedom governing the motion of the radial component. We introduce the effective potential $V_{\text{eff}}(r) = \frac{1}{2m}\frac{\Theta^2}{r^2} - \frac{k}{r}$. When the total energy E is negative, the radial length has to be bounded and energy conservation gives that each orbit in the phase space is closed and is given by the graph of the function

$$r\mapsto \dot r = R(r)/m = \pm\sqrt{2m^{-1}(E - V_{\text{eff}}(r))}.$$

We also have $\dot\theta = \frac{\Theta}{mr^2}$. Therefore we get

$$\frac{d\theta}{dr} = \frac{\Theta/r^2}{\sqrt{2m(E-V_{\text{eff}}(r))}},$$

$$\theta = \int \frac{\Theta/r^2\,dr}{\sqrt{2m(E-V_{\text{eff}}(r))}}$$

where the integral is taken over a period. The integral can be evaluated explicitly as

$$\theta = \arccos\frac{\frac{\Theta}{r} - \frac{mk}{\Theta}}{\sqrt{2mE + \frac{m^2k^2}{\Theta^2}}}.$$

Denoting by $e = \sqrt{1 + \frac{2E\Theta^2}{mk^2}}$ the eccentricity, we obtain from the above equation that

$$r = \frac{\Theta^2/(mk)}{1 + e\cos\theta}. \tag{2.3}$$

This is the equation for a conic section. Depending on the signs of E, the equation represents an ellipse ($E < 0$), parabola ($E = 0$) and hyperbola ($E > 0$).

2.3. Kepler's three laws. From the solution to the two-body problem, we derive Kepler's three laws. Let us now fix $E < 0$. Then we get Kepler's first law, each orbit is moving on an ellipse. Kepler's second law says that the rate of change of the area $S(t)$ swept out by the radius vector is a constant. On the infinitesimal level, we see that this area of the infinitesimal triangle is given by

$$\Delta S = \frac{1}{2} r^2 \sin(\Delta\theta) = \frac{1}{2} r^2 \dot{\theta} \Delta t.$$

This gives $\dot{S} = \frac{1}{2m}\Theta$ where Θ is the angular momentum, so Kepler's second law is really the angular momentum conservation.

We next work on the third law stating that a^3/T^2 is a constant where a is the semimajor of the ellipse and T is the period. First, in (2.3), $\theta = 0$ corresponds to the perigee and $\theta = \pi$ corresponds to the apogee. So we have

$$2a = \frac{\Theta^2/(mk)}{1+e} + \frac{\Theta^2/(mk)}{1-e} = \frac{2\Theta^2/(mk)}{1-e^2} = \frac{k}{|E|}.$$

So we get that for Kepler elliptic motion the semimajor a of the ellipse is related to the energy E through the relation

$$a = \frac{k}{2|E|}, \quad E = -\frac{k}{2a}.$$

This is an important fact used in our construction of noncollision singularities, since energy can be extracted from Kepler elliptic motion by shrinking its semimajor.

The semiminor is

$$b = a\sqrt{1-e^2} = \frac{k}{2|E|}\sqrt{\frac{2|E|\Theta^2}{mk^2}} = \frac{\Theta}{\sqrt{2m|E|}}.$$

Suppose the period is T, so the total area of the ellipse is $\int_0^T \dot{S}(t)\,dt = T\frac{\Theta}{2m} = \pi ab$. Then we obtain

$$T = \frac{2m\pi}{\Theta}\frac{k}{2|E|}\frac{\Theta}{\sqrt{2m|E|}} = \frac{\sqrt{2m}k\pi}{2|E|^{3/2}} = \frac{2\pi a^{3/2}}{(k/m)^{1/2}} = 2\pi a^{3/2}(m_1+m_2)^{1/2}.$$

This is Kepler's third law. In our solar system, let m_1 be the mass of the sun and m_2 be the mass of a planet, since we have $m_2 \ll m_1$, we see that a^3/T^2 is approximately the same constant for all the planets.

2.4. Delaunay coordinates. There is another remarkable symplectic coordinate system called Delaunay coordinates, which is the action-angle coordinates for the Kepler problem [39]. This coordinate system condenses the geometry and dynamics of the Kepler problem in a magic manner. The coordinates are denoted by $(L, \ell, G, g) \in \mathbb{R}_+ \times \mathbb{T} \times \mathbb{R} \times \mathbb{T}$, where $\mathbb{T} := \mathbb{R}/(2\pi\mathbb{Z})$, with the following physical meanings.

- L is the symplectic area enclosed by the graph of $(r \mapsto R(r))$ on the energy level $E < 0$;
- ℓ is the area swept out by the particle, called mean anomaly;
- $G = \Theta$ is the angular momentum;
- g measures the angle formed by the apapsis and the x-axis.

In Delaunay coordinates, the Hamiltonian has the form

$$H(L, \ell, G, g) = -\frac{mk^2}{2L^2}.$$

The only nontrivial Hamiltonian equation is $\dot{\ell} = \frac{mk^2}{L^3}$ which is exactly Kepler's second and third laws. The equation $\dot{L} = 0$ is the energy conservation and $\dot{G} = 0$ is the angular momentum conservation. The equation $\dot{g} = 0$ reflects a special degeneracy of the two-body problem. Double collision corresponds to $G = 0$ where the ellipse degenerates into an interval of multiplicity 2. Note that the Hamiltonian equations are nonsingular for double collision, so Delaunay coordinates have an important merit of resolving the singularity of double collision.

Comparing the above formulas for the semimajor a and semiminor b, we obtain further geometric meanings for the Delaunay variables

$$a = \frac{L^2}{mk}, \quad b = \frac{L|G|}{mk}, \quad e = \sqrt{1 - \frac{G^2}{L^2}}, \tag{2.4}$$

where e is the eccentricity. The mean anomaly ℓ can be related to the polar angle θ through the equations

$$\tan\frac{\theta}{2} = \sqrt{\frac{1+e}{1-e}} \cdot \tan\frac{u}{2}, \quad u - e\sin u = \ell.$$

Denoting the particle's position by $q = (q_1, q_2)$ and its momentum by $p = (p_1, p_2)$ we have the following formulas in the case $g = 0$

$$\begin{aligned} q_1 &= \frac{L^2}{mk}(\cos u - e), \quad q_2 = \frac{LG}{mk}\sin u, \\ p_1 &= -\frac{mk}{L}\frac{\sin u}{1 - e\cos u}, \quad p_2 = \frac{mk}{L^2}\frac{G\cos u}{1 - e\cos u}, \end{aligned} \tag{2.5}$$

Here g does not appear because the argument of apapsis is chosen to be zero. In the general case, we need to rotate the (q_1, q_2) and (p_1, p_2) using the matrix $\begin{bmatrix} \cos g & -\sin g \\ \sin g & \cos g \end{bmatrix}$.

It is remarkable that the Delaunay coordinates, defined for elliptic Kepler motions, are also defined for hyperbolic Kepler motions after slight adaptation (c.f. [11]).

$$\begin{aligned} q_1 &= \frac{L^2}{mk}\left(\cosh u - e\right), \quad q_2 = \frac{LG}{mk}\sinh u, \\ p_1 &= -\frac{mk}{L}\frac{\sinh u}{1-e\cosh u}, \quad p_2 = -\frac{mk}{L^2}\frac{G\cosh u}{1-e\cosh u}. \end{aligned} \tag{2.6}$$

where u and ℓ are related by

$$u - e\sinh u = \ell, \text{ where } e = \sqrt{1+\left(\frac{G}{L}\right)^2} \tag{2.7}$$

is the eccentricity. In the hyperbolic Delaunay coordinates (L, ℓ, G, g), the Hamiltonian for the Kepler hyperbolic motion can be written as

$$H = \frac{mk^2}{2L^2}.$$

For a hyperbola defined by $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ up to a translation or rotation, we also call a the semimajor and b the semiminor. Similar to the elliptic case, we also have $a = \frac{L^2}{mk}$, $b = \frac{L|G|}{mk}$, so one angle formed by the two asymptotes is

$$2\arctan\frac{b}{a} = 2\arctan\frac{|G|}{L}.$$

3. Gerver's ideal model

In [13], Gerver proposed an ideal local model for the construction of noncollision singularities in four-body problem, which lies in the heart of the proof of Theorem 1.12 of [40]. The model reveals a magical link of the geometry and dynamics of the Kepler problem. Since it is very interesting and elementary, in this section we explain it in details.

We need three parameters to determine an ellipse: the semimajor, semiminor and the angle of the apapsis, and similarly for a hyperbola. In Delaunay variables, these parameters are equivalent to L, G, g. The mean anomaly ℓ indicates the position of the particle on the ellipse or hyperbola.

Gerver's model considers the dynamics of the subsystem Q_2, Q_3, Q_4 in the limit case $\mu = m_3 = m_4 = 0$ with Q_1 ignored. Assuming Q_3 and Q_4 to have zero mass so they have no gravitational interaction with each other, but both are attracted by Q_2. Suppose $m_2 = 1$, then the Hamiltonians governing the motion of Q_3 and Q_4 are both (2.1) with $m = k = 1$. Indeed, we may set in (2.1) $m_2 = 1$ (corresponding to the mass of Q_2) and $m_1 = \mu$ (corresponding to the mass of Q_3 or Q_4), then we have $k = m_1 m_2 = \mu$ and $m = \frac{m_1 m_2}{m_1 + m_2} = \mu + O(\mu^2)$. Instead of the momentum y, we use the velocity $v = y/m$, then we divide the Hamiltonian by μ and let $\mu \to 0$ to obtain the Hamiltonian governing the massless particles Q_3 and Q_4 in coordinates $(x, v) \in \mathbb{R}^2 \times \mathbb{R}^2$.

We assume that

- Q_3 has elliptic motion and Q_4 has hyperbolic motion with focus Q_2;
- Q_3 and Q_4 arrive at the correct intersection point of their orbits simultaneously (see Figure 2 and 3);
- Q_3 and Q_4 do not interact unless they have an exact collision, and the collision is treated as elastic collision (energy and momentum are preserved).

Under these assumptions, Gerver designed a two-step procedure such that

- the major axis of the elliptic motion is always kept vertical;
- the incoming and outgoing asymptotes of the hyperbolic motion are always horizontal;
- after two steps of the collision procedure, the ellipse has the same eccentricity as the ellipse before the first collision, but has a smaller semi-major axis (see Figure 2 and 3).

The last bullet point is the key point of the construction since it reveals an important self-similar behavior after two steps, which is crucial for iteration and repeated energy extraction from the Kepler elliptic motion. The remaining part of this section is organized as follows. In Section 3.1, we first talk about elastic collision. In the next two subsections, we describe Gerver's two-step procedure in details. Finally, in Section 3.4, we formalize Gerver's result and provide equations to solve for Gerver's map.

3.1. Elastic collision. The interaction of Q_3 and Q_4 is desribed by the elastic collision, i.e. the energy conservation and momentum conservation laws are obeyed. Suppose two particles Q_3 and Q_4 collide at 0 with velocities v_3^- and v_4^- respectively. Suppose the velocities after the collision are v_3^+ and v_4^+ respectively. By energy conservation and momentum conservation we have

$$|v_3^+|^2 + |v_4^+|^2 = |v_3^-|^2 + |v_3^-|^2, \quad v_3^+ + v_4^+ = v_3^- + v_4^-.$$

These provide three equations and we want to solve for four variables $(v_3^+, v_4^+) \in \mathbb{R}^2 \times \mathbb{R}^2$ for given initial velocities $(v_3^-, v_4^-) \in \mathbb{R}^2 \times \mathbb{R}^2$, The solution can be expressed explicitly as

$$v_3^+ = \frac{v_3^- + v_4^-}{2} + \left|\frac{v_3^- - v_4^-}{2}\right| n(\alpha), \quad v_4^+ = \frac{v_3^- + v_4^-}{2} - \left|\frac{v_3^- - v_4^-}{2}\right| n(\alpha), \tag{3.1}$$

where $n(\alpha)$ is a unit vector making angle α with $v_3^- - v_4^-$. Here $\alpha \in [0, 2\pi)$ is a free parameter.

3.2. The first collision, angular momentum transfer. In this section, we describe the first step in Gerver's construction (see Figure 2).

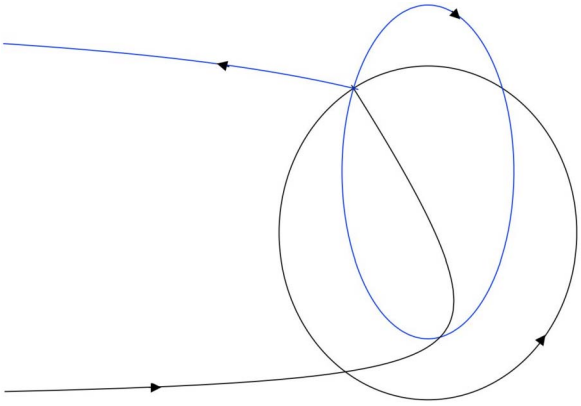

FIGURE 2. The first collision: angular momentum transfer

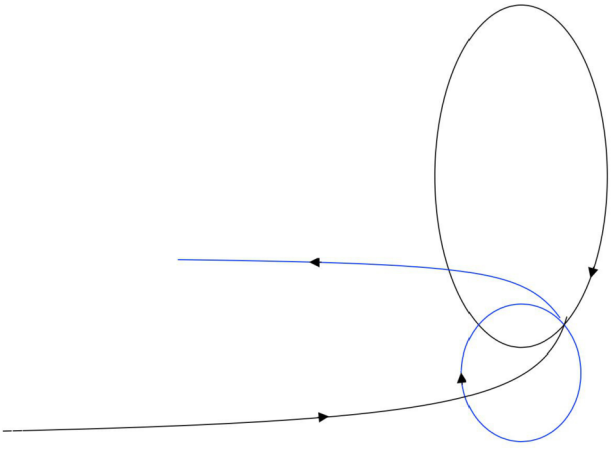

FIGURE 3. The second collision: energy transfer

Let $\varepsilon_0 \in (0, \frac{\sqrt{2}}{2})$ and $\varepsilon_1 = \sqrt{1-\varepsilon_0^2} \in (\frac{\sqrt{2}}{2}, 1)$. Suppose the initial ellipse of Q_3's orbit is

$$\frac{x^2}{\varepsilon_1^2} + (y-\varepsilon_0)^2 = 1, \tag{3.2}$$

which is an ellipse with semimajor 1 and semiminor ε_1 (hence eccentricity ε_0) and one focus at zero.

We note that the ellipse (3.2) as well as the following ellipse

$$\frac{x^2}{\varepsilon_0^2} + (y-\varepsilon_1)^2 = 1, \tag{3.3}$$

both pass through the point $Z := (X, Y) = (-\varepsilon_0\varepsilon_1, \varepsilon_0 + \varepsilon_1)$ due to the identity $\varepsilon_0^2 + \varepsilon_1^2 = 1$. The ellipse (3.3) has semimajor 1 and semiminor ε_0.

This geometric fact can be viewed dynamically. We now consider each ellipse as the orbit of a particle moving obeying the Hamiltonian system (2.1) with $m = k = 1$. Then using the geometric meanings of Delaunay coordinates (2.4), we get that for both orbits, the energy is $-\frac{1}{2}$. The angular momentum of (3.2) is $G = \varepsilon_1$ (we use the convention that G is positive if Q_3 moves counterclockwise on the ellipse) and that of (3.3) is $-\varepsilon_0$ (see Figure 2 for the orientation of the two orbits).

It is important to notice that the Y-coordinate of the intersection point Z is minus of the change in angular momentum and the X-coordinate is the product of the initial and final angular momenta.

Let v_3^- and v_3^+ be the velocities of Q_3 before and after the collision respectively. Then the above considerations gives the the equations of angular momentum and energy

$$\begin{aligned} Z \times v_3^- &= \varepsilon_1, \quad \frac{1}{2}|v_3^-|^2 - \frac{1}{|Z|} = -\frac{1}{2}, \\ Z \times v_3^+ &= -\varepsilon_0, \quad \frac{1}{2}|v_3^+|^2 - \frac{1}{|Z|} = -\frac{1}{2}. \end{aligned} \tag{3.4}$$

Note that we have four equations to solve for four variables $(v_3^-, v_3^+) \in \mathbb{R}^4$. We get the solutions explicitly taking care of signs correctly according to Figure 2

$$v_3^- = \frac{1}{1+\varepsilon_0\varepsilon_1}(-\varepsilon_1^2, -\varepsilon_0), \quad v_3^+ = \frac{1}{1+\varepsilon_0\varepsilon_1}(\varepsilon_0^2, \varepsilon_1). \tag{3.5}$$

We next find the hyperbolic orbits for Q_4 geometrically. By assumption, we require that the incoming and outgoing hyperbolas to have horizontal asymptotes. We also require that they intersect the ellipsis at Z. By energy consideration, both orbits of Q_4 have energy $\frac{1}{2}$, hence in geometric terms, both hyperbolas have semimajors 1 focused at the origin and one asymptote $y = -p$ for some $p \in \mathbb{R}$.

We consider first a standard hyperbola $\tilde{y}^2 - (\tilde{x}\tan\psi)^2 = 1$ with $\tilde{y} > 0$ and some ψ to be determined later. This is a branch of hyperbola opening up, symmetric along the $\tilde{y}$-axis and the asymptote forms an angle ψ with the $\tilde{x}$-axis. The coordinate of the focus is $(0, \csc\psi)$.

The hyperbola that we are looking for is related to this one by a rotation by angle ψ (so that one asymptote is parallel to the x-axis) and a vertical shift by $\csc\psi$ (so that the focus is shifted to the origin).

We thus have the coordinate change $\begin{cases} \tilde{x} &= x\cos\psi + y\sin\psi \\ \tilde{y} &= y\cos\psi - x\sin\psi + \csc\psi. \end{cases}$

Substituting this to the $(\tilde{x}, \tilde{y})$ equation and after some manipulations we arrive at the following equation for (x, y)

$$y^2(1-p^{-2}) - \frac{2xy}{p} + 2yp - 2x + 1 + p^2 = 1$$

where $p = \cot\psi$ and $y = -p$ is exactly the asymptote parallel to the x-axis. We arrive at

$$(p^2 + yp - x - r)(p^2 + yp - x + r) = 0, \quad r = \sqrt{x^2 + y^2}.$$

Each factor on the LHS represents one branch of the hyperbola and we choose the first one, so we have at the collision point $Z = (X, Y)$

$$p^2 + Yp - X - |Z| = 0, \quad |Z| = \sqrt{X^2 + Y^2} = 1 + \varepsilon_0\varepsilon_1.$$

We solve for p

$$p_\mp = \frac{-Y - (\mp\sqrt{Y^2 + 4(X + |Z|)})}{2},$$

where $y = -p_- < 0$ is the incoming asymptote and $y = -p_+ > 0$ is the outgoing asymptote.

We next adopt a dynamical viewpoint. By construction, we have that particles moving on both hyperbolas have energy $\frac{1}{2}$ and the semimajor $a = 1$. We next consider the angular momentum. Suppose we have a hyperbola symmetric along the x-axis, focused at the origin and opening to the left. Using hyperbolic Delaunay coordinates in Section 2.4, we have that the angle formed by the two asymptotes containing part of the x-axis is given by $2\arctan\frac{b}{a} = 2\arctan\frac{|G|}{L}$. In order to make one asymptote horizontal, we need to rotate the hyperbola by angle $g = \pm\arctan\frac{|G|}{L}$. We have $\tan g = \pm|G|$ since $L = 1$. Recall that we started with the hyperbola $\tilde{y}^2 - (\tilde{x}\tan\psi)^2 = 1$ symmetric around $\tilde{y}$-axis and opening up, which was rotated by angle ψ to have horizontal asymptotes. We conclude that $\cot\psi = \tan g$ up to a minus sign, so the value $p_\pm$ is exactly the angular momentum up to a minus sign. In other words, the vertical intercepts of the asymptotes are exactly the angular momentum, up to a sign. The sign of the angular momentum is always positive since both the incoming and outgoing orbits moves counterclockwise around the focus. So we get that the initial angular momentum is p_- and the final angular momentum is $-p_+$ and both are positive. The change of angular momentum is given by $-p_+ - p_- = Y = \varepsilon_0 + \varepsilon_1$, which is exactly the minus of the change of angular momentum for the two elliptic orbits, so the angular momentum conservation is verified.

We next let v_4^- and v_4^+ be the velocities of Q_4 before and after collision at Z respectively. The above information on angular momentum and energy gives

$$\begin{aligned} Z \times v_4^- &= p_-, \quad \frac{1}{2}|v_4^-|^2 - \frac{1}{|Z|} = \frac{1}{2}, \\ Z \times v_4^+ &= -p_+, \quad \frac{1}{2}|v_4^+|^2 - \frac{1}{|Z|} = \frac{1}{2}. \end{aligned} \tag{3.6}$$

The solution is given explicitly by

$$v_4^- = \left(1 - \frac{Y}{|Z|p_-}, \frac{1}{|Z|p_-}\right), \quad v_4^+ = \left(-1 + \frac{Y}{|Z|p_+}, -\frac{1}{|Z|p_+}\right).$$

We already have the energy conservation and angular momentum conservation by construction. However, to show that the interaction is an elastic collision, we have to verify the momentum conservation $v_3^- + v_4^- = v_3^+ + v_4^+$. Note that we have $\frac{1}{p_-} + \frac{1}{p_+} = \frac{p_- + p_+}{p_- p_+} = \frac{Y}{X+|Z|} = Y$. This gives

$$v_4^+ - v_4^- = \left(-2 + \frac{Y}{|Z|}\left(\frac{1}{p_-} + \frac{1}{p_+}\right), -\frac{1}{|Z|}\left(\frac{1}{p_-} + \frac{1}{p_+}\right)\right) = \left(\frac{-1}{|Z|}, -\frac{Y}{|Z|}\right).$$

This is exactly $v_3^- - v_3^+$ from (3.5), so we get the momentum conservation.

3.3. The second collision, energy transfer. In this section, we describe the second step in Gerver's construction (see Figure 3). The second collision transfers energy from Q_3 to Q_4. We start with the Q_3 ellipse after the first collision. The second collision point is chosen at $Z = (X, Y) = (\varepsilon_0^2, 0)$ and the equations for the initial and final ellipses are respectively

$$\begin{aligned} &\frac{x^2}{\varepsilon_0^2} + (y - \varepsilon_1)^2 = 1, \\ &\frac{\varepsilon_1^2}{\varepsilon_0^4} x^2 + \left(\frac{\varepsilon_1^2}{\varepsilon_0^2} y + \varepsilon_0\right)^2 = 1. \end{aligned} \tag{3.7}$$

The final ellipse has semimajor $a = \frac{\varepsilon_0^2}{\varepsilon_1^2}$ and semiminor $b = \frac{\varepsilon_0^2}{\varepsilon_1}$ (hence eccentricity $\varepsilon_0 = \sqrt{1 - \frac{b^2}{a^2}}$, the same as the initial ellipse in step 1). Equation (2.4) implies that the energy of the final Q_3 orbit is $-\frac{1}{2a} = -\frac{1}{2}\frac{\varepsilon_1^2}{\varepsilon_0^2}$ and the angular momentum is $-\varepsilon_0$.

Again the angular momentum and energy conservations

$$\begin{aligned} &Z \times v_3^- = -\varepsilon_0, \quad \frac{1}{2}|v_3^-|^2 - \frac{1}{|Z|} = -\frac{1}{2}, \\ &Z \times v_3^+ = -\varepsilon_0, \quad \frac{1}{2}|v_3^+|^2 - \frac{1}{|Z|} = -\frac{\varepsilon_1^2}{2\varepsilon_0^2}. \end{aligned} \tag{3.8}$$

yield the velocities v_3^- and v_3^+ of Q_3 before and after the collision

$$v_3^- = (-\frac{\varepsilon_1}{\varepsilon_0}, -\frac{1}{\varepsilon_0}), \quad v_3^+ = (1, -\frac{1}{\varepsilon_0}).$$

We next consider the Q_4-hyperbola. Similarly, the incoming asymptote $y = -p_-$ of Q_4 is solved from the equation

$$p^2 + Yp - X - |Z| = 0, \quad |Z| = \varepsilon_0^2.$$

So $p_- = \sqrt{2}\varepsilon_0$ so that the incoming asymptote is $y = -\sqrt{2}\varepsilon_0$ and the angular momentum is $\sqrt{2}\varepsilon_0$.

The velocity $v_4^- = (1, \frac{\sqrt{2}}{\varepsilon_0})$ of Q_4 before the collision can be solved from the angular momentum and energy equations

$$Z \times v_4^- = p_-, \quad \frac{1}{2}|v_4^-|^2 - \frac{1}{|Z|} = \frac{1}{2}.$$

After the collision, the energy of Q_4 is $\frac{\varepsilon_1^2}{2\varepsilon_0^2}$ so that the semimajor of the hyperbola is $\varepsilon_0^2/\varepsilon_1^2$. We perform a rescaling $(x,y)\mapsto(\bar{x},\bar{y}):=(x,y)\frac{\varepsilon_1^2}{\varepsilon_0^2}$ so that the semimajor is rescaled to 1, and apply our p-equation to $(\bar{x},\bar{y})$. The outgoing asymptote is then solved as $\bar{y}=-p_+$ where p_+ solves the equation

$$p^2+Yp\frac{\varepsilon_1^2}{\varepsilon_0^2}-X\frac{\varepsilon_1^2}{\varepsilon_0^2}-|Z|\frac{\varepsilon_1^2}{\varepsilon_0^2}=0.$$

Then we get $p_+=-\sqrt{2}\varepsilon_1$ so the outgoing asymptote is $y=\sqrt{2}\frac{\varepsilon_0^2}{\varepsilon_1}$ and the angular momentum is $\sqrt{2}\varepsilon_0$.

The velocity $v_4^+=(-\frac{\varepsilon_1}{\varepsilon_0},\frac{\sqrt{2}}{\varepsilon_0})$ after the collision is solved from the angular momentum and energy equations

$$Z\times v_4^+=\sqrt{2}\varepsilon_0,\quad \frac{1}{2}|v_4^+|^2-\frac{1}{|Z|}=\frac{\varepsilon_1^2}{2\varepsilon_0^2}.$$

Again, it can be verified that the velocities $v_{3,4}^{\pm}$ satisfies the momentum conservation. Indeed, we have

$$v_3^+-v_3^-=(1+\frac{\varepsilon_1}{\varepsilon_0},0)=v_4^--v_4^+.$$

We see that after the two steps, we get a final Q_3-ellipse that has the same eccentricity with the initial one, but smaller semimajor.

3.4. Formalizing Gerver's result. We summarize the above construction of Gerver into the following table and statement.

	1st collision	@$(-\varepsilon_0\varepsilon_1,\varepsilon_0+\varepsilon_1)$	2nd collision	@$(\varepsilon_0^2,0)$
	Q_3	Q_4	Q_3	Q_4
energy	$-\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}\to-\frac{\varepsilon_1^2}{2\varepsilon_0^2}$	$\frac{1}{2}\to\frac{\varepsilon_1^2}{2\varepsilon_0^2}$
angular momentum	$\varepsilon_1\to-\varepsilon_0$	$-p_-\to-p_+$	$-\varepsilon_0$	$\sqrt{2}\varepsilon_0$
eccentricity	$\varepsilon_0\to\varepsilon_1$		$\varepsilon_1\to\varepsilon_0$	
semimajor	1	1	$1\to\left(\frac{\varepsilon_0}{\varepsilon_1}\right)^2$	$1\to\frac{\varepsilon_1^2}{\varepsilon_0^2}$
semiminor	$\varepsilon_1\to\varepsilon_0$	$\lvert p_-\rvert\to\lvert p_+\rvert$	$\varepsilon_0\to\frac{\varepsilon_0^2}{\varepsilon_1}$	$\sqrt{2}\varepsilon_0\to\sqrt{2}\varepsilon_1$

We next formalize the above construction of Gerver into a statement.

We use the variables (e_3,g_3,E_3) to describe the shape of the Q_3-ellipse focused at the origin, where (e_3,g_3,E_3) are eccentricity, argument of apapsis and energy respectively. For the Q_4-hyperbola, we can use similarly (e_4,g_4,E_4) to describe it. However, we can eliminate E_4 since we have energy conservation $E_4=-E_3$ and we can eliminate g_4 since we assume one asymptote is horizontal. So e_4 is enough to determine the shape of the Q_4-hyperbola and also the collision point that is the intersection point of Q_3-ellipse and Q_4-hyperbola. Note also that we have a freedom to choose either Q_3 or Q_4 to move on the elliptic orbit and the other hyperbolic after each collision. We thus obtain a map $\mathbf{G}_{e_4,j,\omega}$ that we call Gerver's map that maps from the

space of ellipses to itself, depending on parameter e_4 the eccentricity of Q_4, $j = 1, 2$ the first or second collision and $\omega = 3, 4$ the escaping body. For the sake of definiteness we always choose $\omega = 4$ in the following statement.

PROPOSITION 3.1. *Assume that the total energy of the Q_2, Q_3, Q_4 system is zero, i.e. $E_3 + E_4 = 0$, and fix the incoming and outgoing asymptotes of the hyperbola to be horizontal.*

(a) *For $E_3^* = -\frac{1}{2}$, $g_3^* = \frac{\pi}{2}$ and for any $e_3^* \in (0, \frac{\sqrt{2}}{2})$, there exist $e_4^*, e_4^{**}, \lambda_0 > 1$ such that*

$$(e_3, g_3, E_3)^{**} = \mathbf{G}_{e_4^*,1,4}\,(e_3, g_3, E_3)^*, \quad (e_3, -g_3, \lambda_0 E_3)^* = \mathbf{G}_{e_4^{**},2,4}\,(e_3, g_3, E_3)^{**},$$

where $E_3^{**} = E_3^* = -\frac{1}{2}$, $g_3^{**} = g_3^* = \frac{\pi}{2}$ and $e_3^{**} = \sqrt{1 - e_3^{*2}}$.

(b) *There is a constant $\bar{\delta}$ such that if (e_3, g_3, E_3) lie in a $\bar{\delta}$ neighborhood of (e_3^*, g_3^*, E_3^*), then there exist smooth functions $e_4'(e_3, g_3)$, $e_4''(e_3, g_3)$, and $\lambda(e_3, g_3, E_3)$ such that*

$$e_4'(e_3^*, g_3^*) = e_4^*, \quad e_4''(e_3^*, g_3^*) = e_4^{**}, \quad \lambda(e_3^*, g_3^*, E_3^*) = \lambda_0,$$

$$(\bar{e}_3, \bar{g}_3, \bar{E}_3) = \mathbf{G}_{e_4'(e_3,g_3),1,4}\,(e_3, g_3, E_3)\,,$$
$$(e_3^*, -g_3^*, \lambda(e_3, g_3, E_3) E_3^*) = \mathbf{G}_{e_4''(e_3,g_3),2,4}\,(\bar{e}_3, \bar{g}_3, \bar{E}_3)\,.$$

Part (a) is the main content of [**13**] described above in Section 3.2 and 3.3, which gives a two-step procedure to decrease the energy of the elliptic Kepler motion and maintain the self-similar structure (See Figure 2 and 3). We call the collision points in part (a) the *Gerver's collision points*. Part (b) says that once the ellipse gets deformed slightly away from the standard case in Figure 2 after the first collision, we can correct it by changing the phase of Q_3 slightly at the next collision to guarantee that the ellipse that we get after the second collision is standard.

Gerver's construction in Section 3.2 and 3.3 finds a special solution of the map $\mathbf{G}$. We next introduce a set of implicit equations written in polar coordinates in order to solve the Gerver map $\mathbf{G}$ for general initial data.

We use the variable θ for polar angle with the positive y axis as the axis $\theta = 0$. Recall the formula (2.3) $r = \frac{G^2}{1 - e\cos\theta}$ for conic sections in which the apapsis lies on the axis $\theta = \pi$. So we have the following two equations describing the Q_3-ellipse and Q_4-hyperbola before $(-)$ and after $(+)$ the collision.

$$\begin{cases} r_3^{\pm} = \dfrac{(G_3^{\pm})^2}{1 - e_3^{\pm}\sin(\theta_3^{\pm} + g_3^{\pm})}, \\ r_4^{\pm} = \dfrac{(G_4^{\pm})^2}{1 - e_4^{\pm}\sin(\theta_4^{\pm} - g_4^{\pm})}. \end{cases} \tag{3.9}$$

We have energy conservation and angular momentum conservation.

$$E_3^+ + E_4^+ = E_3^- + E_4^-\,, \tag{3.10}$$

$$G_3^+ + G_4^+ = G_3^- + G_4^-\,, \tag{3.11}$$

Momentum conservation gives the following equation in addition to the angular momentum conservation (see the derivation below)

(3.12)
$$\frac{e_3^+}{G_3^+}\cos(\theta_3^+ + g_3^+) + \frac{e_4^+}{G_4^+}\cos(\theta_4^- - g_4^-) = \frac{e_3^-}{G_3^-}\cos(\theta_3^- + g_3^-) + \frac{e_4^-}{G_4^-}\cos(\theta_4^- - g_4^-),$$

The fact that the four curves (Q_3 ellipses and Q_4 hyperbolas before and after the collision) intersect at the same point gives the following equations

$$\frac{(G_3^+)^2}{1 - e_3^+\sin(\theta_3^+ + g_3^+)} = \frac{(G_3^-)^2}{1 - e_3^-\sin(\theta_3^- + g_3^-)}, \tag{3.13}$$

$$\frac{(G_3^+)^2}{1 - e_3^+\sin(\theta_3^+ + g_3^+)} = \frac{(G_4^+)^2}{1 - e_4^+\sin(\theta_4^+ - g_4^+)}, \tag{3.14}$$

$$\frac{(G_3^-)^2}{1 - e_3^-\sin(\theta_3^- + g_3^-)} = \frac{(G_4^-)^2}{1 - e_4^-\sin(\theta_4^- - g_4^-)}, \tag{3.15}$$

$$\theta_3^+ = \theta_3^-, \quad \theta_4^- = \theta_3^-, \quad \theta_4^+ = \theta_3^+. \tag{3.16}$$

Gerver's map can be solved from the equations (3.10)–(3.16) as follows. We have the incoming variables $Z_- := (E_3, \theta_3, G_3, g_3, E_4, \theta_4, G_4, g_4)^-$ and outgoing variables $Z_+ := (E_3, \theta_3, G_3, g_3, E_4, \theta_4, G_4, g_4)^+$. For given Z_-, we first eliminate $\theta_3^+ = \theta_4^+$ from Z_+ using (3.16). We determine $g_4^\pm = \pm\arctan\frac{G_4^\pm}{L_4^\pm}$ by the condition of horizontal asymptotes. There are five remaining variables $(E_3, G_3, g_3, E_4, G_4)^+$ in Z_+ which can be solved from the remaining five equations (3.10)–(3.14) for given Z_-.

We next derive equation (3.12) from the conservation of momentum as follows. Represent the position vector as $\mathbf{r} = r\mathbf{e}_r$. Then the velocity is $\dot{\mathbf{r}} = \dot{r}\mathbf{e}_r + r\dot{\theta}\mathbf{e}_\theta$. Conservation of momentum gives $(\dot{\mathbf{r}}_3)^- + (\dot{\mathbf{r}}_4)^- = (\dot{\mathbf{r}}_3)^+ + (\dot{\mathbf{r}}_4)^+$. Taking the radial component and using the polar representation of the ellipse $r = \frac{G^2}{1-e\sin(\theta+g)}$, we get

$$\dot{r} = \frac{G^2}{(1 - e\sin(\theta+g))^2}e\cos(\theta+g)\dot{\theta} = \frac{r^2}{G^2}e\cos(\theta+g)\frac{G}{r^2} = \frac{e}{G}\cos(\theta+g).$$

4. Noncollision singularities in four-body problem

In this section, we explain the proof of Theorem 1.12. We first show heuristically how Gerver's construction may give a noncollision singularity. We have learned in the last section that Gerver's construction gives a smaller Q_3-ellipse with the same eccentricity after a two-step procedure. We can then zoom in the picture to unit size and repeat the procedure. For Kepler elliptic motion, energy E_3 of Q_3 is related to the semimajor a through the relation $E_3 = -\frac{1}{2a}$. Suppose we can iterate the procedure for infinitely many steps, we get that the energy of Q_3 grows to $-\infty$ like $E_3 \sim -\lambda^n$, where $\lambda > 1$ is the ratio of the semimajors of the initial and final ellipses. Energy conservation implies that the energy of Q_4 grows to ∞ like $E_4 \sim \lambda^n$. For most of the

time Q_4 is far from the two large bodies so that most of its energy is kinetic energy and the speed of Q_4 grows exponentially like $|\dot{Q}_4| \sim \lambda^{n/2}$. Suppose the masses of Q_4 and Q_3 are $\mu \ll 1$, the masses of Q_1 and Q_2 are 1 and the initial distance of Q_1 and Q_2 is χ. Then when Q_4 turns around Q_1, its velocity changes direction by almost π, then momentum conservation gives that $|\dot{Q}_1| \simeq \mu|\dot{Q}_4|$, hence a fixed proportion of energy is transferred to Q_1. Under iteration, the speed of Q_1 grows like $|\dot{Q}_1| \simeq \mu\lambda^{n/2}$, which is still much slower than $|\dot{Q}_4| \simeq (1-\mu)\lambda^{n/2}$. So during the n-th iteration, the particle Q_4 can still complete a return within time $\chi(\lambda - c\mu)^{-n/2}$ for some constant c. for some constant c. Note that the return time decays exponentially fast so infinitely many returns can be completed in finite time and a non collision singularity can be constructed.

As we have explained in the introduction, to construct a noncollision singularity, we need first a local model for energy exchange, and also a mechanism for globalization to glue different local pieces. Gerver's construction provides us a local model, but we still need to perform perturbative analysis around Gerver's solution. On the other hand, essential difficulty comes from the globalization part. The globalization issue in the models of [**25, 42**] is relatively easy since their models has symmetry, which is absent in our four-body problem.

The mechanism of globalization in our work goes essentially back to the concept of structural stability in hyperbolic dynamics. Recall that the Arnold's cat map $A := \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} : (\mathbb{R}/\mathbb{Z})^2 \to (\mathbb{R}/\mathbb{Z})^2$ is unstable in the sense that it has positive Lyapunov exponent $\lim \frac{1}{n} \log \|A^n\|$, which means that two nearby points may diverge exponentially so that the long time dynamics is not predictable. However, this map enjoys remarkable structural stability property. One manifestation of such structural stability is the shadowing property: let $\varepsilon > 0$ be a small number, for any pseudo orbit $\{x_0, x_1, \ldots, x_n\}$ satisfying $|Ax_i - x_{i+1}| < \varepsilon$, $i = 0, \ldots, n-1$, there is a real orbit $\{A^i y,\ i = 0, 1, \ldots, n\}$ such that $|A^i y - x_i| < C\varepsilon$, $i = 0, 1, \ldots, n$, for some constant C independent of n and ε. In the globalization part of our work, we shall exploit the strong hyperbolicity to find real orbit shadowing different local pieces.

The remaining part of this section is organized as follows. In Section 4.1, we give the heuristics of the two strongly expanding directions. In Section 4.2, we give our hyperbolicity framework, assuming which we prove the main Theorem 1.12. In Section 4.3, we explain the hyperbolicity of the tangent dynamics by some simple but essential arguments. In Section 4.4 we discuss some further issues such as excluding collisions etc.

4.1. Two sources of hyperbolicity. To carry out Gerver's strategy for infinitely many steps in the $\mu > 0$, $\chi < \infty$ case, we need to utilize the hyperbolicity of the tangent dynamics. We notice that there are two sources of strong expansion.

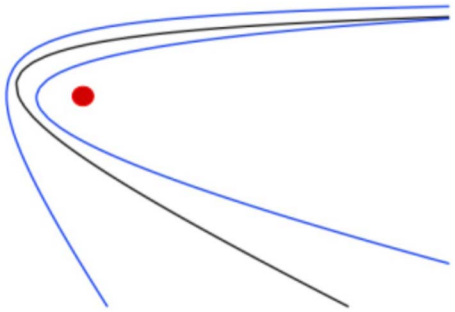

FIGURE 4. The first expanding direction

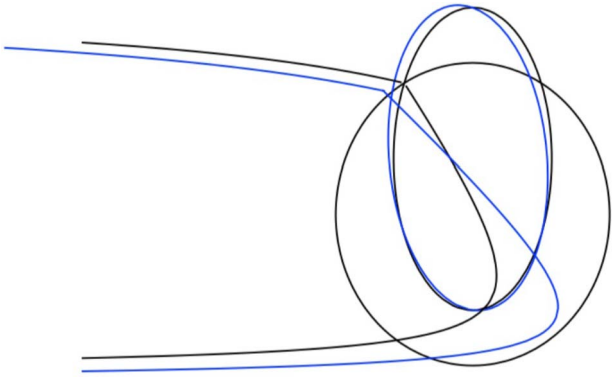

FIGURE 5. The second expanding direction

4.1.1. *The first expanding direction, prescribing angular momentum.* The first expanding direction comes from the hyperbolicity of the hyperbolic Kepler motion. Namely, if we shoot a bunch of orbits of Q_4 towards Q_1 with the same initial velocities and slightly different y components of the initial positions, then the bunch of orbits will diverge after turning around Q_1 (see Figure 4). This actually enables us to change the vertical component of Q_4 arbitrarily when Q_4 comes to a neighborhood of the pair Q_2-Q_3. Recall in Section 3.2 and 3.3, the vertical component of Q_4 is also its angular momentum, so the angular momentum of Q_4 can be prescribed arbitrarily.

4.1.2. *The second expanding direction, phase stretching and synchronization.* The second expanding direction is a bit subtler. Let us use $\theta \in [0, 2\pi)$ to denote the phase Q_3 on its ellipse when colliding with Q_4. Let us look at Figure 5. Imagine that Q_4 and Q_3 collide at a point with phase θ slightly different from Gerver's collision point if the initial condition of Q_4 is slightly perturbed. We require again the outgoing asymptote of Q_4 to be horizontal. We expect that the semi major a_3^+ of the Q_3 ellipse after collision is slightly

different from that of Gerver's case, i.e.

$$\frac{\partial a_3^+}{\partial \theta} \neq 0. \tag{4.1}$$

Kepler's law $\frac{a^3}{T^2} = \frac{1}{(2\pi)^2}$ implies that the two ellipses (blue and black in Figure 5) have different periods, hence within one period of Gerver's standard elliptic orbit, the particle Q_3 on the other ellipse has a small phase difference. It takes long time (of order χ) for Q_4 to complete a return and come to the next collision with Q_3. During this long time, two different elliptic motions have completed many periods hence accumulated a huge phase difference. This is the second expanding direction along which a small initial phase difference is stretched to a huge final phase difference. The second expanding direction allows us to synchronize Q_3 and Q_4. Namely, by adjusting the phase of collision in the present step slightly, we can arrange that Q_3 and Q_4 come to the same point at the same time in the next step. We also note that if we adjust the phase in the present step slightly further, Q_3 and Q_4 may miss each other in the next step, but if we continue to adjust the phase in the present step furthermore, it is possible to control Q_3 and Q_4 to come to close encounter again in the next step, but differ from the previous encounter by 2π in phase. All these adjustments are done in a very small interval (of order $1/\chi$) of the phase variable since the phase stretching rate is huge, as we will see in the next section. This is the reason why we get a Cantor set as initial conditions.

4.2. Outline of the proof of Theorem 1.12. In this section, we sketch the proof of the main theorem based on two technical lemmas. The idea is to introduce the Poincaré return map, compute its derivative and show that there is a two dimensional invariant subspace that is strongly expanding under the differential of the Poincaré map. These two strongly expanding directions are described in the last subsection.

The phase space is obtained by reducing the translation invariance, fixing the zeroth energy level and picking the Poincaré section $\{x = -2\}$. We will introduce coordinates later in Section 4.3 to parametrize the phase space of the four-body problem as $T^*(\mathbb{T}^3 \times \mathbb{R}^2)$. We will not consider all the points in the section $\{x = -2\}$ as initial condition. Instead, we pick a small number δ and define $U_1(\delta)$, $U_2(\delta)$ as δ neighborhoods of Gerver's first and second collision points respectively, traced back to the section $\{x = -2\}$ along the flow. The Poincaré section $\{x = -2\}$ cuts the orbit into two different pieces. The right piece defines the local map $\mathbb{L}$ and the left piece is defines the global map $\mathbb{G}$ by following the orbits starting from and ending at the section $\{x = -2\}$. The Poincaré return map is defined as the composition $\mathcal{P} = \mathbb{G} \circ \mathbb{L}$ whose domain is contained in $U_j(\delta),\ j = 1, 2$. We also define a renormalization map $\mathcal{R}$ to zoom in the position space by λ and slow down the velocity by dividing by $\sqrt{\lambda}$, where λ is the energy of the Q_3 after two steps of interactions in Gerver's construction.

4.2.1. *C^1 control: the derivatives of local and global maps.*

LEMMA 4.1 (Lemma 3.1 of [40]). *If the incoming asymptote θ^- and the outgoing asymptotes θ^+ satisfy $|\theta^-| \le C\mu$ and $|\theta^+ - \pi| \le C\mu$ for some constant C, then for $x \in U_j$, $j = 1, 2$ being a initial point on the section $\{x = -2\}$, there exist linear functionals $\mathbf{l}(x)$, a vector fields $\mathbf{u}(x)$ and matrices $B(x)$ which are uniformly bounded such that*

$$d\mathbb{L}(x) = \frac{1}{\mu}\mathbf{u}(x) \otimes \mathbf{l}(x) + B(x) + o(1),\ 1/\chi \ll \mu \to 0. \tag{4.2}$$

LEMMA 4.2 (Lemma 3.2 of [40]). *If the y coordinates of the initial and final positions of Q_4 are bounded when applying $\mathbb{G}$, Then there exist linear functionals $\bar{\mathbf{l}}(x)$ and $\bar{\bar{\mathbf{l}}}(x)$ and vectorfields $\bar{\mathbf{u}}(y)$ and $\bar{\bar{\mathbf{u}}}(y)$ such that*

$$d\mathbb{G}(x) = \chi\bar{\mathbf{u}}(y) \otimes \bar{\mathbf{l}}(x) + \chi^2\bar{\bar{\mathbf{u}}}(y) \otimes \bar{\bar{\mathbf{l}}}(x) + O(\mu\chi),\ 1/\chi \ll \mu \to 0, \tag{4.3}$$

where we denote x the initial point and $y = \mathbb{G}(x)$ the final point.

We will explain the heuristics of the two lemmas in Section 4.3, where we can see that all the vectors $\mathbf{u}$, $\bar{\mathbf{u}}$, $\bar{\bar{\mathbf{u}}}$, $\mathbf{l}$, $\bar{\mathbf{l}}$, $\bar{\bar{\mathbf{l}}}$ and matrix B can be worked out explicitly.

Let us ignore the $o(1)$, $O(\mu\chi)$ perturbations in $d\mathbb{L}$, $d\mathbb{G}$ respectively for a moment. After application of $d\mathbb{G}$ we get a plane $\mathrm{span}\{\bar{\mathbf{u}}, \bar{\bar{\mathbf{u}}}\}$. We next apply $d\mathbb{L}$ to get a plane $\mathrm{span}\{\mathbf{u}, BY\}$ where $Y \in (\mathrm{Ker}\mathbf{l}) \cap \mathrm{span}\{\bar{\mathbf{u}}, \bar{\bar{\mathbf{u}}}\}$. To apply $d\mathbb{G}$ again, we want to guarantee that the plane $\mathrm{span}\{\mathbf{u}, BY\}$ is not collapsed into a line or a point so that we need the following transversality condition

$$(\mathrm{Ker}\bar{\mathbf{l}} \cap \mathrm{Ker}\bar{\bar{\mathbf{l}}}) \pitchfork \mathrm{span}\{\mathbf{u}, BY\}. \tag{4.4}$$

This condition is equivalent to $\det \begin{pmatrix} \bar{\mathbf{l}}\mathbf{u} & \bar{\mathbf{l}}BY \\ \bar{\bar{\mathbf{l}}}\mathbf{u} & \bar{\bar{\mathbf{l}}}BY \end{pmatrix} \neq 0$, which can be verified by working out the vectors and matrices explicitly.

4.2.2. *The Cantor set construction.* With the above lemmas on $d\mathbb{G}$, $d\mathbb{L}$ and the transversality condition, we establish the strong expansion of the Poincaré map.

LEMMA 4.3 (Lemma 2.17 of [40]). *There are cone fields $\mathcal{K}_1$ on $T_x(T^*(\mathbb{T}^3 \times \mathbb{R}^2))$, $x \in U_1(\delta)$ and $\mathcal{K}_2$ on $T_x(T^*(\mathbb{T}^3 \times \mathbb{R}^2))$, $x \in U_2(\delta)$, each of which contains a two dimensional plane, such that*

- *Invariance: $d\mathcal{P}(\mathcal{K}_1) \subset \mathcal{K}_2$, $d(\mathcal{R} \circ \mathcal{P})(\mathcal{K}_2) \subset \mathcal{K}_1$.*
- *Expansion: If $v \in \mathcal{K}_1$, then $\|d\mathcal{P}(v)\| \ge c\chi\|v\|$. If $v \in \mathcal{K}_2$, then $\|d(\mathcal{R} \circ \mathcal{P})(v)\| \ge c\chi\|v\|$.*

The cone can be defined as the set of vectors forming some small angle η with the plane $\mathrm{span}\{\bar{\mathbf{u}}, \bar{\bar{\mathbf{u}}}\}$. It is used to handle the error terms in Lemma 4.1 and 4.2. The leading terms of the vectors $\bar{\mathbf{u}}$ and $\bar{\bar{\mathbf{u}}}$ can be worked out explicitly as ∂_{ℓ_3} and $\partial_{G_4} - \frac{L_4}{L_4^2+G_4^2}\partial_{g_4}$ respectively in Delaunay coordinates, which will become clear in Section 4.3.3 and 4.3.4. The strong expansion in item (2) of Lemma 4.3 implies that we can prescribe the ℓ_3 (the phase

of Q_3 on the ellipse) and G_4 (the angular momentum of Q_4) components arbitrarily after each application of the Poincaré map.

DEFINITION 4.4. We call a C^1 surface $S_1 \subset U_1(\delta)$ (respectively $S_2 \subset U_2(\delta)$) admissible if $TS_1 \subset \mathcal{K}_1$ (respectively $TS_2 \subset \mathcal{K}_2$).

Lemma 4.3 shows that admissible surfaces are always mapped to admissible surfaces and are strongly expanded. To see the idea clearer, we consider the following toy model. Let $E_\lambda : \mathbb{T} \to \mathbb{T}$, $x \mapsto \lambda x \bmod 1$, $\lambda \in \mathbb{N}$ is a large number. For any fixed interval $[a, b] \subsetneq \mathbb{T}$, the set of points p such that $E_\lambda^n p \in [a, b]$ for all $n \in \mathbb{N}$ is a Cantor set, since it can be constructed by a process of open interval deletion for infinitely many steps.

With these preparations, we can now construct the Cantor set in Theorem 1.12. As Q_1 gets farther and farther, the orbit of Q_4 gets more and more horizontal when away from Q_1 and Q_2-Q_3. Moreover, we can control the orbit of Q_3 such that it always stays close to the standard ellipses in Gerver's model (see Section 4.4.2 later). Thus the configuration is close to Gerver's model and orbits can always be found to intersect the sets $U_i(\delta)$, $i = 1, 2$. Since the Poincaré map maps admissible surfaces to admissible ones, we can restrict to the admissible surfaces to reduce the map to a two dimensional one. From the explicit form of $\bar{\mathbf{u}}$ and $\bar{\bar{\mathbf{u}}}$, we can parametrize the admissible surfaces by ℓ_3 and G_4 and write the restricted Poincaré map in these coordinates. The strong expansion of the restricted Poincaré map enables us to iterate for infinitely many steps. Take a piece of admissible surface S_1 and look at the pre-image $(\mathcal{RP}^2)^{-1}S_1$, which consists of many copies of tiny (due to expansion in Lemma 4.3) pieces of admissible (due to invariance in Lemma 4.3) surfaces. The reason why we have many copies is because one of the expanding direction, the phase variable θ, is defined up to 2π. Finally, our Cantor set is constructed as $\lim_j (\mathcal{RP}^2)^{-j} S_j$, where S_j is a piece of admissible surface at the $2j$-th step.

4.3. The tangent dynamics. In this section, we explain the Lemma 4.1 and 4.2 on the derivatives of local and global maps under various simplifying assumptions. These simple derivations lie in the heart of our lengthy calculations in [**40**].

4.3.1. *The coordinates.* The proof of the technical Lemma 4.1 and 4.2 involves estimating the fundamental solution to the variational equations of the Hamiltonian equations in certain suitable coordinates. We explain in this section our choice of coordinates.

The first step in the choice of coordinates is to introduce Jacobi coordinates. We explain in details Jacobi coordinates for three-body problem in Appendix A.1. We introduce two sets of Jacobi coordinates to eliminate the translation invariance of the Hamiltonian system of our four-body problem. When Q_4 is closer to Q_2 than Q_1 (see the upper figure in Figure 6), we denote $x_3 = Q_3 - Q_2$, define x_4 to be the distance from Q_4 to the mass center of Q_3 and Q_2 and define x_1 to be the distance from Q_1 to the mass center

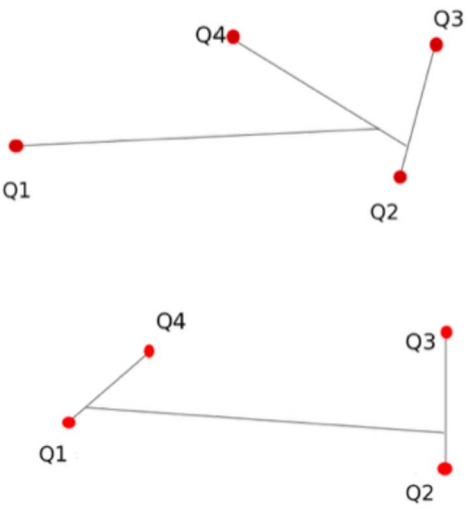

FIGURE 6. The Jacobi coordinates

of Q_4, Q_3, Q_2. There is a corresponding linear change of coordinates on the momentum side v_1, v_3, v_4 to make the new set of coordinates symplectic in the reduced phase space. Similarly, we construct coordinates when Q_4 is closer to Q_1 (see the lower figure of Figure 6) by defining $x_3 = Q_3 - Q_2$, $x_4 = Q_4 - Q_1$ and x_1 to be the distance from the mass center of Q_1-Q_4 to the mass center of Q_2-Q_3. We do not give explicit formula for the coordinate change in this paper. Interested readers can refer to [**40**] for details. The advantage of the coordinates is that they reduce the Hamiltonian system into three Kepler motions with "controllable" perturbations. To see the meaning of "controllable", we show one example. There is a term $\mu/|Q_3 - Q_4|$ in the potential. When we integrate over time t assuming Q_4 moves away from Q_3 linearly in t, the integral blows up like $\ln t$ as $t \to \infty$. However, in Jacobi coordinates, some cancelations occur and the leading contribution to interaction between x_3, x_4 is given by a term of the form $\mu\langle x_3, x_4\rangle/|x_4|^3$ whose t integral is now convergent as $t \to \infty$ knowing that x_3 is bounded and x_4 is linear in t.

Now we have perturbed Kepler motions for $(x, v)_{3,1,4}$. We next introduce the classical Delaunay coordinates (L_4, ℓ_4, G_4, g_4) for the hyperbolic motion (x_4, v_4) and (L_3, ℓ_3, G_3, g_3) for the elliptic motion (x_3, v_3). By fixing an energy level, for instance 0, and picking Poincaré sections, we can eliminate L_4, ℓ_4 from our list of variables by solving L_4 as a function of other variables and treating ℓ_4 as the new time. Moreover, we stick to x_1, v_1 without reducing the rotational invariance, so we get totally ten variables $(x_1, v_1, L_3, \ell_3, G_3, g_3, G_4, g_4)$ to describe the dynamics and the phase space in this set of coordinates can be written as $T^*(\mathbb{T}^3 \times \mathbb{R}^2)$.

4.3.2. *The Rutherford scattering.* In this section, we explain the proof of Lemma 4.1. One of Gerver's assumption is to model the interaction of Q_3 and

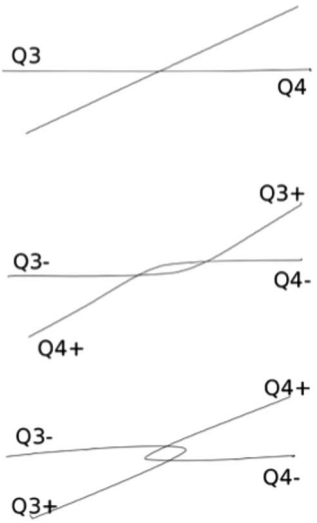

FIGURE 7. Shadowing elastic collision by Kepler hyperbolic motions

Q_4 as elastic collision. When we turn on the Newtonian interaction between the two particles, we get hyperbolic Kepler motions to shadow the elastic collision picture (see the lower two figures in Figure 7). The momentum and energy are still conserved, so we still have (3.1). The rotation angle α, now measured as the angle formed by two asymptotes of the hyperbolas, is no longer a free variable but determined by the initial conditions. To determine the rotation angle α, we introduce the Rutherford's scattering formula that is well known in physical literature. We consider two particles of masses m_3, m_4 interacting via the Newtonian potential $-k/|Q_-|$, where $Q_- = Q_3 - Q_4$ is the relative position. The relative motion is governed by the Hamiltonian (2.1) with $m = \frac{m_3 m_4}{m_3+m_4} = \frac{\mu}{2}$ and $k = \mu^2$. Denoting by $P_- = \mu v_-$ the relative momentum where $v_- = v_3 - v_4$ the relative velocity, we can write the Hamiltonian explicitly as $H = \frac{P_-^2}{\mu} - \frac{\mu^2}{|Q_-|} = \mu v_-^2 - \frac{\mu^2}{|Q_-|}$. We then divide the Hamiltonian by μ, to obtain $H/\mu = v_-^2 - \frac{\mu}{|Q_-|}$. Noting that the Poincaré-Cartan form $v_- dQ_- - H/\mu dt = \frac{1}{\mu}(P_- dQ_- - H dt)$, so we can treat H/μ as a Hamiltonian and (v_-, Q_-) as the canonical variables. We next convert (v_-, Q_-) to Delaunay variables (L, ℓ, G, g). We get $H/\mu = -\frac{\mu^2}{4L^2}$ and the angular momentum $G = v_- \times Q_-$. For convenience, we introduce $\mathcal{L} = L/\mu$. The rotation angle formed by the two asymptotes of the hyperbola is given by

$$\alpha = 2 \arctan \frac{b}{a} = 2 \arctan \frac{G}{\mu \mathcal{L}}. \tag{4.5}$$

From the Hamiltonian we have $|v_-|^2 = O(1)$ hence $\mathcal{L} = O(1)$ as $\mu \to 0$. Thus we get also $G = O(\mu)$ since the rotation angle α is bounded away from $0, \pi$. The variable G is the most important quantity in the scattering theory

called the *impact parameter*, which is a measurement of the closest distance between the two particles during the scattering process. When we compute how the outgoing velocities v_3^+, v_4^+ depend on incoming velocities v_3^-, v_4^-, we substitute the expression α (4.5) into (3.1). We notice that v_3^+, v_4^+ has implicit dependence on v_3^-, v_4^- through G and also dependence through $\mathcal{L}$ and explicit dependence. So we compute the derivative as follows

$$\begin{aligned}\frac{\partial\alpha}{\partial G} &= \frac{2}{\mu}\frac{1/\mathcal{L}}{1+(G/(\mu\mathcal{L})^2)} = O(\mu^{-1})\\ \frac{\partial\alpha}{\partial\mathcal{L}} &= \frac{2G}{\mu}\frac{-1/\mathcal{L}^2}{1+(G/(\mu\mathcal{L})^2)} = O(1)\\ \frac{\partial(v_3,v_4)^+}{\partial(v_3,v_4)^-} &= O\left(\frac{1}{\mu}\right)\frac{\partial(v_3,v_4)^+}{\partial\alpha}\otimes\frac{\partial G}{\partial(v_3,v_4)^-}\\ &\quad + (\text{derivative not through } G).\end{aligned}$$

This calculation shows that the most effective way to change (v_3^+, v_4^+) significantly is to change the impact parameter G. This simple derivation explains the structure of the derivative of the local map in Lemma 4.1. Moreover, both the tensor part and the B part in $d\mathbb{L}$ can be computed explicitly.

4.3.3. *The $O(\chi)$ term in $d\mathbb{G}$, the shears.* The $O(\chi)$ term in $d\mathbb{G}$ involves mainly the motion of Q_3. We forget about the perturbation coming from Q_4 for simplicity to see the ideas. The Hamiltonian for elliptic Kepler motion in Delaunay coordinates $H_3 = \frac{-1}{2L_3^2}$. We integrate the Hamiltonian equations from time 0 to time T to get

$$\begin{cases}\dot{L}_3 &= 0,\\ \dot{\ell}_3 &= \frac{1}{L_3^3},\\ \dot{G}_3 &= 0,\\ \dot{g}_3 &= 0,\end{cases}\Longrightarrow\begin{cases}L_3(T) &= L_3(0),\\ \ell_3(T) &= \ell_3(0)+\frac{T}{L_3^3(0)},\\ G_3(T) &= G_3(0),\\ g_3(T) &= g_3(0).\end{cases}$$

The derivative matrix has the following decomposition

(4.6)

$$\frac{\partial(L,\ell,G,g)_3(T)}{\partial(L,\ell,G,g)_3(0)} = \begin{pmatrix}1 & 0 & 0 & 0\\ -\frac{3T}{L_3^4(0)} & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix} = -\frac{3T}{L_3^4(0)}\begin{pmatrix}0\\1\\0\\0\end{pmatrix}\otimes(1,0,0,0)+O(1),$$

if we choose $T = O_{\chi\to\infty}(\chi)$ as the time for Q_4 to complete a return. This gives the $\chi\bar{\mathbf{u}}\otimes\bar{\mathbf{l}}$ part of the $d\mathbb{G}$. We see that the vectors $\bar{\mathbf{u}}$ and $\bar{\mathbf{l}}$ can be obtained explicitly. It turns out that in the limit $1/\chi \ll \mu \to 0$, we have $\bar{\mathbf{u}} \to \frac{\partial}{\partial\ell_3}$ and $\bar{\mathbf{l}} \to dL_3$.

Notice the above matrix (4.6) is only a shear, which has no hyperbolicity. Under iteration, it grows only linearly but not exponentially. There is a mechanism called shear-induced chaos, namely, hyperbolicity can be created by a shear combined with a kick. To see why, we multiply another matrix

(kick) to a shear matrix to get

$$\begin{pmatrix} 1 & 0 \\ \chi & 1 \end{pmatrix} \begin{pmatrix} 1 & \mathbf{1} \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ \chi & \chi+1 \end{pmatrix}. \tag{4.7}$$

The resulting matrix has two eigenvalues, $O(\chi)$ and $O(1/\chi)$. Actually, to get an eigenvalue of order χ, we only need to require that in the above second matrix the bold **1** entry is nonzero and independent of χ, while the other entries can be arbitrary. In our case, $d\mathbb{G}$ gives us a shear matrix and $d\mathbb{L}$ gives us a kick so that the composition of local and global map gives rise to an expansion with rate χ. We have a transversality condition (4.4), which is essentially $\frac{\partial a_3^+}{\partial \theta} \neq 0$ in (4.1) (see the proof of the transversality condition in Section 3.4 of [**41**]). This requirement amounts to requiring the bold **1** entry is nonzero in (4.7).

4.3.4. *The $O(\chi^2)$ term in $d\mathbb{G}$, the scattering.* In this section, we explain the $O(\chi^2)$ term in $d\mathbb{G}$. The χ^2 term in $d\mathbb{G}$ comes mainly from the motion of Q_4 as described in Section 4.1.1.

We pick two more Poincaré sections $\{x = -\chi/2\}$ to cut the orbit of the global map into three pieces denoted by (I), (III), (V) (see Figure 8). The (I), (V) pieces of orbits are considered as a perturbed hyperbolic Kepler motion focused at the mass center of Q_2 and Q_3, while the (III) piece of orbit is treated as a perturbed hyperbolic Kepler motion focused at Q_1. We need two coordinate changes when the orbit crosses the sections denoted by (II), (IV). Since L, ℓ are reduced by fixing an energy level and picking a Poincaré section, we have only Delaunay variables G, g to characterize the motion of Q_4, whose meanings are respectively the angular momentum and the argument of apapsis (direction) of the hyperbolic motion. We define the angle of asymptotes of the hyperbola as

$$\theta = g \pm \arctan G/L, \tag{4.8}$$

since g is the direction of the symmetric axis of the hyperbola and $2\arctan b/a = 2\arctan G/L$ is the angle formed by the two asymptotes. We make the following simplifying assumptions.

- Assume L is a constant, say, 1.
- Assume $v_4 = (\pm 1, \theta)$.
- Assume the two hyperbolas to the left and right of the section $\{x = -\chi/2\}$ share the same asymptote angle θ.

Let us now look at Figure 8. When we change coordinates from (I) to (III), we are changing the focus of the hyperbolic motion of Q_4. During this coordinate change, the two different hyperbolic motions share nearly the same asymptotes, which we identify. So we first convert (G, g) on the right to variables (G, θ) on the right, then we convert (G, θ) on the right to (G, θ) on the left, and the last step is to convert (G, θ) back to (G, g) on the left. We summarize the steps as follows. The coordinates changes (II) from

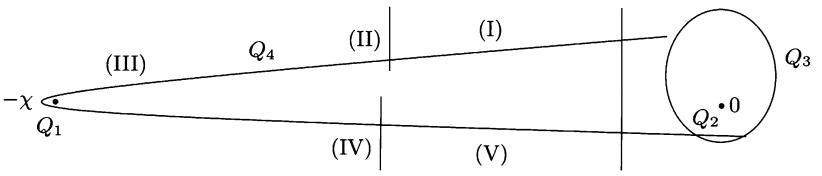

FIGURE 8. The Poincaré sections

the right (subscript R) to the left (subscript L) is the composition of

$$(G,g)_R \xrightarrow{(i)} (G,\theta)_R \xrightarrow{(ii)} (G,\theta)_L \xrightarrow{(iii)} (G,g)_L,$$

where the maps (i), (ii), (iii) are given explicitly as follows:

$$(i): \begin{cases} G_R &= G_R \\ \theta_R &= g_R - \arctan G_R \end{cases},$$

$$(ii): \begin{cases} G_L &= G_R + \chi\theta \\ \theta_R &= \theta_L \end{cases},$$

$$(iii): \begin{cases} G_L &= G_L \\ g_L &= \theta_L - \arctan G_L \end{cases}.$$

The g and θ relations are obtained from (9), while $G_L = G_R + \chi\theta$ comes from the definition of angular momentum since we move the origin from $(0,0)$ to $Q_1 = (-\chi, 0)$,

$$G_L = v_L \times x_L = v_R \times (x_R - (\chi,0)) = G_R - v_R \times (\chi,0) = G_R + \chi\theta.$$

The derivatives for $(II) = (iii)(ii)(i)$ is

$$D[(iii)(ii)(i)] = \begin{pmatrix} 1 & 0 \\ \sharp & 1 \end{pmatrix} \begin{pmatrix} 1 & \chi \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \sharp & 1 \end{pmatrix},$$

where we use $\sharp$ to denote some constants that can be computed explicitly. Similarly, for matrix (IV) going from the left to the right, we get

$$D[(iii')(ii')(i')] = \begin{pmatrix} 1 & 0 \\ \sharp & 1 \end{pmatrix} \begin{pmatrix} 1 & -\chi \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \sharp & 1 \end{pmatrix},$$

Next we look at the derivative of the composition $(IV)(III)(II)$ assuming $D(III) = id$, since G, g are constants of motion for the Kepler problem (Unfortunately, this assumption is wrong. It turns out that the perturbation from the pair Q_2, Q_3 has a nontrivial contribution to $D(III)$. We can get an explicit expression for G, g part of $D(III)$, which is not id but does not cause any trouble).

$$
\begin{aligned}
&D[(IV)(III)(II)]\\
&=\begin{pmatrix}1&0\\ \sharp&1\end{pmatrix}\begin{pmatrix}1&-\chi\\ 0&1\end{pmatrix}\begin{pmatrix}1&0\\ \sharp&1\end{pmatrix}\cdot\begin{pmatrix}1&0\\ \sharp&1\end{pmatrix}\begin{pmatrix}1&\chi\\ 0&1\end{pmatrix}\begin{pmatrix}1&0\\ \sharp&1\end{pmatrix}\\
&=\begin{pmatrix}1&0\\ \sharp&1\end{pmatrix}\begin{pmatrix}1+\sharp\chi&-\sharp\chi^2\\ \sharp&-\sharp\chi+1\end{pmatrix}\begin{pmatrix}1&0\\ \sharp&1\end{pmatrix}\\
&=\sharp\begin{pmatrix}1&0\\ \sharp&1\end{pmatrix}\left(\chi^2\begin{pmatrix}1\\ 1/\chi\end{pmatrix}\otimes(1/\chi,1)+O(1)\right)\begin{pmatrix}1&0\\ \sharp&1\end{pmatrix}.
\end{aligned}
$$

This calculation shows the χ^2 part of $d\mathbb{G}$. We see that the $\bar{\bar{\mathbf{u}}}$, $\bar{\bar{\mathbf{l}}}$ can again be calculated explicitly. It turns out that in the limit $1/\chi \ll \mu \to 0$, we have $\bar{\bar{\mathbf{l}}} \to d\theta_4^+$, where θ_4^+ is the outgoing asymptote of the Q_4 hyperbola after close encounter with Q_3, and $\bar{\bar{\mathbf{u}}} \to \left(0_{1\times 8}; 1, -\frac{L_4}{L_4^2+G_4^2}\right)$, which shows that after the application of the global map there is a linear relation between G_4 and g_4 forced by the fact that the asymptote of Q_4 must be close to horizontal when Q_4 comes close to Q_2.

4.4. Further issues.

4.4.1. *Collision exclusion and the existence of returning orbit.* We need to exclude the possibility of the collisions between the pair Q_3-Q_4 and the pair Q_1-Q_4. The pair Q_3-Q_4 is easily done using the formula (4.5) since the rotation angle α is not close to π in Gerver's construction so that $G/(\mu\mathcal{L})$ is bounded away from zero. Next, we explain how to exclude the Q_1-Q_4 collision. Recall that in the two-body problem, if two bodies collide they will bounce back. Suppose we have a collision of Q_1 and Q_4, then we reverse the time for the piece of orbit coming to collision and compare it with the bouncing back orbit. We work in Delaunay coordinates so that the collisional singularity is resolved. We can measure the difference of the two orbits by integrating the variational equations (derivative of the Hamiltonian equations). It turns out that the deviation of the two orbits is at most $O(\mu)$ when we trace the orbit to the section $\{x = -2\}$. However, the returning orbit that we want should stay close to Gerver's model and the y coordinates of Q_4 for two consecutive visits to the section $\{x = -2\}$ differ by a $O(1)$ number. So we conclude that there is no collision between Q_1 and Q_4.

4.4.2. *How to control the shape of Q_3 ellipse.* In order to control the phase space dynamics such that the image of the admissible surface always visits the fixed neighborhood $U(\delta)$, we look at our list of variables $(L_3, \ell_3, G_3, g_3; x_1, v_1; G_4, g_4)$. First L_3 is always rescaled to the unit size by the renormalization and x_1, v_1 can be control by the angular momentum. Indeed, the total angular momentum is assumed to be zero. The renormalization always rescales G_3 and G_4 to order 1 so $G_1 = x_1 \times v_1$ has also order 1 by $G_1 + G_3 + G_4 = 0$. The magnitude $|x_1|$ can be estimated as $O(\chi)$ which grows exponentially since each renormalization introduces a λ factor, and $|v_1| = O(1)$ so we get the angle $\angle(x_1, v_1) = O(1/\chi) \to 0$ in order to have

$G_1 = O(1)$. Next the variables ℓ_3, G_4, which can be used to parametrized the admissible surface, can be chosen arbitrarily due to the strong expansion, and g_4 is also determiend since the asymptote of Q_4 is almost horizontal. The only remaining variables are G_3, g_3, which we want to control such that they stay close to Gerver's values. We notice that the Q_3 ellipse may deviate from Gerver's standard case and lose self-similarity. For this, we invoke item (b) of Proposition 3.1. Consider again Gerver's ideal model. We have four variables L, ℓ, G, g to characterize the elliptic motion. The variable ℓ is almost the same as the phase θ which can be controlled by the strong hyperbolicity and L is related to the semi major which can always be rescaled to one applying the renormalization $\mathcal{R}$. It remains to control G and g such that they do not deviate too far. The point is that during each collision, there is a phase of the collision point of Q_3 and Q_4 that can be adjusted. So after the two steps in Gerver's model, we get two phases θ_1, θ_2 on which the final orbit parameters G^+, g^+ depend smoothly. Part (b) of Proposition 3.1 is proved by verifying $\frac{\partial(G^+,g^+)}{\partial(\theta_1,\theta_2)} \neq 0$, so we can control G^+, g^+ through adjusting the phases θ_1, θ_2.

4.4.3. *How to switch the roles of Q_3 and Q_4.* We see from Figure 7 that there are two hyperbolic motions shadowing the same elastic collision picture. This implies that by choosing the rotation angle α correctly, one can switch the roles of the messenger and the captured particle after each Q_3 and Q_4 close interaction. For this reason, for any given symbolic sequence $\boldsymbol{\omega}$, a Cantor set $\Sigma_{\boldsymbol{\omega}}$ of non collision singularities can be constructed.

4.4.4. *The measure and Hausdorff dimension of the Cantor set.* We notice that each time when we apply the renormalization $\mathcal{R}$ by zooming in the configuration space by λ, the distance between Q_1 and Q_2 get multiplied by λ, hence χ in Lemma 4.2, as a measurement of the distance between Q_1 and Q_2, grows exponentially to infinity. Since χ is the expansion rate in Lemma 4.3. In each step of the Cantor set construction, we preserves only $1/\chi^2$ of the total measure on the initial admissible surface. Since $1/\chi^2$ decays exponentially to zero, we conclude that the Hausdorff dimension of the Cantor set restricted to each two dimensional admissible surface is 0. Hence our Cantor set of non collision singularities is a zero measure set of codimension 2 in the zero angular momentum level set.

5. Triple collision blowup

We have shown in the introduction that central configurations appear naturally when blowing up a total collapse. So it is natural to use central configuration as the local model for the near collision dynamics in the construction of noncollision singularities. In this section, we show that the near multiple collision dynamics can provide another mechanism of acceleration.

To analyze the near multiple collision dynamics, we introduce a blowup technique developed by McGehee [**21**]. The point of this technique is to introduce scaling invariant variables based on polar coordinates and derive

their equations of motion. It turns out that there is a limiting dynamical system when we let $I = \sum m_i|Q_i|^2 \to 0$, which is shadowed by the near collision dynamics. The limiting dynamics can be quite well understood, so can the near collision dynamics.

This section is organized as follows. We first give a general formalism for N-body total collapse in Section 5.1. We then apply the general formalism in three different settings of the three-body problem. In Section 5.2, we talk about the isosceles three-body problem and explain Devaney's classification of final motions for near collision dynamics. We also explain the mechanism of acceleration for the near triple collision dynamics, so this subsection is the heart of the current section. In Section 5.3, we talk about the collinear three-body problem and explain the work [**25**]. In Section 5.4, we talk about the Sitnikov-Alekseev model and explain briefly the work [**42**].

5.1. The general formalism of the total collision blowup. The general idea is as follows. Recall the Lagrange-Jacobi identity from Section A.1. We denote $r^2 = I = \sum m_i|Q_i|^2$. Then we get $2r\dot{r} = \dot{I} = J := 2\sum m_i Q_i \cdot \dot{Q}_i$ and $\dot{J} = K + H$ where K is the kinetic energy and H is the Hamiltonian.

We introduce the inner product $\langle \cdot, \cdot \rangle$ on $\mathbb{R}^{3N}$ defined by

$$\langle u, v \rangle = \sum m_i u_i v_i, \quad \|u\| := \langle u, u \rangle^{1/2}, \quad \forall\ u, v \in \mathbb{R}^{3N}.$$

We then introduce the normalized configuration $\mathsf{s} = r^{-1}Q = r^{-1}(Q_1, \dots, Q_N)$, hence we have that $\|\mathsf{s}\| = 1$. Then we introduce the normalized velocity $\mathsf{v} = \sqrt{r}\dot{Q}$ and

$$v = \langle \mathsf{s}, \mathsf{v} \rangle = r^{-1/2}\langle Q, \dot{Q} \rangle = r^{-1/2} J.$$

We next split the normalized velocity v into the radial component $v\mathsf{s}$ with $v = \langle \mathsf{v}, \mathsf{s} \rangle$ and the component w tangent to the shape sphere $\mathcal{E} = \{\|s\| = 1\}$

$$\mathsf{v} = v\mathsf{s} + \mathsf{w}, \quad \langle \mathsf{s}, \mathsf{w} \rangle = 0.$$

Then we obtain the following set of equations of motion where $'$ means the derivative $\frac{d}{d\tau}$ with the time rescaling $dt = r^{3/2}d\tau$

$$\begin{cases} r' &= vr, \\ v' &= \frac{1}{2}v^2 + \|\mathsf{w}\|^2 + U(\mathsf{s}), \\ \mathsf{s}' &= \mathsf{w}, \\ \mathsf{v}' &= -M^{-1}\nabla U(\mathsf{s}) - U(\mathsf{s})\mathsf{s} - \frac{1}{2}v\mathsf{w} - \|\mathsf{w}\|^2\mathsf{s}, \end{cases} \tag{5.1}$$

where $U(\mathsf{s}) = rU(x)$ and $\nabla U(\mathsf{s}) = r^2 \nabla U(x)$. Note that all variables except r is rescaling invariant. We can also introduce the equation $\mathsf{w}' = \mathsf{v}' - (v'\mathsf{s} + v\mathsf{s}')$ instead of the v' equation above to form a closed set of equations for the variables $(r, v, \mathsf{s}, \mathsf{w})$.

This is the general equations of motion for the N-body problem in the blowup coordinates. The total energy has the form

$$rH = \frac{1}{2}v^2 + \frac{1}{2}\|\mathsf{w}\|^2 + U(\mathsf{s}).$$

In the limit $r = 0$ corresponding to the total collapse, we get that the r equation becomes trivial. We then call the resulting phase space for the variables $(v, \mathsf{s}, \mathsf{w})$ the *collision manifold* when $r = 0$. Since we have $rH \to 0$ when $r \to 0$, we get

$$v' = \frac{1}{2}\|\mathsf{w}\|^2 + rH \to \frac{1}{2}\|\mathsf{w}\|^2 \geq 0,$$

which follows essentially from the Lagrange-Jacobi identity $\ddot{I} = K + H$. This implies that v provides us a Lyapunov function that is nondecreasing along an orbit on the collision manifold. At fixed points of the ODE system, we have

$$r = 0,\ \mathsf{w} = 0,\ v^2 + 2U(\mathsf{s}) = 0, \quad -M^{-1}\nabla U(\mathsf{s}) - U(\mathsf{s})\mathsf{s} = 0,$$

where the last equation implies that the fixed points are normalized central configurations, which are critical points of U restricted to the sphere $\mathcal{E}$.

In practice we may need to understand better the dynamics of s, w. The general strategy is to introduce angular coordinates or spherical coordinates on $\mathcal{E}$ for s, dual to which, the variable w is represented by normalized (by $r^{-1/2}$) angular momentum dual to the spherical variables respectively.

5.2. The isosceles three-body problem. To illustrate the above general formalism, we implement it in a special three-body problem in the plane called isosceles three-body problem studied first by Devaney [**10**].

Consider the three-body problem as follows. Let Q_1 and Q_2 be two equal mass particles parallel to the x-axis with masses normalized to 1 and let Q_3 be a particle on the y-axis with mass m. The three particles form an isosceles triangle and remains to be isosceles for all time in the plane.

The main result of [**10**] is a classification of the final motions of the system near triple collision. There are two types of final motions except a total collapse,

(1) the binary Q_1-Q_2 moves to infinity parallel to the x-axis in opposite directions and Q_3 oscillates on the y-axis;
(2) Q_3 and the mass center of the pair move to infinity along the y-axis in opposite directions with velocities that can be as large as we wish, and the binary undergoes infinitely many double collisions.

This result was proved by blowing up the triple collision and analyzing the dynamics on the collision manifold.

5.2.1. *Triple collision blowup in the isosceles three-body problem.* We next perform the blow up procedure. We first introduce the Jacobi coordinates as in (A.3), under which the Hamiltonian has the following form

$$H = \frac{y_1^2}{2M_1} + \frac{y_0^2}{2M_0} + V(x_0, x_1), \quad V(x_0, x_1) = -\frac{m}{|x_1 - \frac{x_0}{2}|} - \frac{m}{|x_1 + \frac{x_0}{2}|} - \frac{1}{|x_0|},$$

where the reduced masses are $M_0 = \frac{1}{2}$, $M_1 = \frac{2m}{m+2}$. The system has two degrees of freedom. To study the triple collision, we carry out the above general blowup scheme as follows.

Introducing $M = \mathrm{diag}\{M_0, M_1\}$, we denote

$$x = (x_0, x_1) \in \mathbb{R}^2,\ \mathsf{v} = r^{1/2}\dot{x},\ \mathsf{s} = x/r \in \mathbb{S}^1, \quad v = \langle \mathsf{s}, \mathsf{v}\rangle, \quad \mathsf{w} = \mathsf{v} - v\mathsf{s}.$$

The equations of motion are

$$\begin{cases} r' &= rv, \\ \mathsf{s}' &= \mathsf{w}, \\ v' &= \|\mathsf{w}\|^2 + \frac{1}{2}v^2 + \bar{V}(\mathsf{s}), \\ \mathsf{w}' &= -\frac{1}{2}v\mathsf{w} - \|\mathsf{w}\|^2\mathsf{s} - M^{-1}\nabla\bar{V}(\mathsf{s}) - \bar{V}(\mathsf{s})\mathsf{s}. \end{cases} \tag{5.2}$$

Here $\bar{V}(\mathsf{s}) = rV(x)$, $\nabla\bar{V}(\mathsf{s}) = r^2\nabla V(\mathsf{x})$. We also have the energy relation

$$rH = \frac{1}{2}\|\mathsf{w}\|^2 + \frac{1}{2}v^2 + \bar{V}(\mathsf{s}).$$

Instead of the vector valued variables s and w, we further introduce $\psi = \arctan\frac{\sqrt{M_1}x_1}{\sqrt{M_0}x_0}$ in place of $\mathsf{s} = M^{-1/2}(\cos\psi, \sin\psi)$ and $w = r^{1/2}\langle \mathsf{v}, \mathsf{e}_w\rangle$ in place of w where $\mathsf{e}_w = (-\sin\psi(0,1), \cos\psi(1,0))$.

Then in the variables (r, ψ, v, w), we obtain the following equations of motion

$$\begin{cases} r' &= rv \\ v' &= w^2 + \frac{1}{2}v^2 + \bar{V}(\psi), \\ \psi' &= w, \\ w' &= -\frac{1}{2}vw - \bar{V}'(\psi). \end{cases} \tag{5.3}$$

where $\bar{V}(\psi) = -\frac{\sqrt{2}}{\cos\psi} - \frac{4m_1}{\sqrt{\frac{1}{2}\cos^2\psi + \frac{8m_1}{m_1+2}\sin^2\psi}}$.

The potential $\bar{V}$ blows up when $\psi = \pm\pi/2$ corresponding to the double collision. It is standard to regularize the double collision after Sundman. Let us introduce $\hat{w} = \cos\psi w$ as well as a new change of time $\frac{d\tau}{d\hat{\tau}} = \cos\psi$. The equations of motion becomes

$$\begin{cases} \frac{dr}{d\hat{\tau}} &= rv\cos\psi, \\ \frac{dv}{d\hat{\tau}} &= \big(2rH - \frac{1}{2}v^2 - \bar{V}(\psi))\big)\cos\psi \\ \frac{d\psi}{d\hat{\tau}} &= \hat{w} \\ \frac{d\hat{w}}{d\hat{\tau}} &= -\frac{1}{2}v\hat{w}\cos\psi - \bar{V}'(\psi)\cos^2\psi - \hat{w}\sin\psi, \end{cases} \tag{5.4}$$

with the energy relation

$$rH = \frac{\hat{w}^2}{2\cos^2\psi} + \frac{1}{2}v^2 + \bar{V}(\psi).$$

In this new coordinates, the system is an analytic vector field defined on $[0,\infty)\times\mathbb{R}\times[-\pi/2,\pi/2]\times\mathbb{R}$ and the double collision becomes an elastic collision so that the singularity is regularized. Denote by $\mathcal{M}_0$ the collision manifold with $r = 0$.

We state the main result of Devaney as follows.

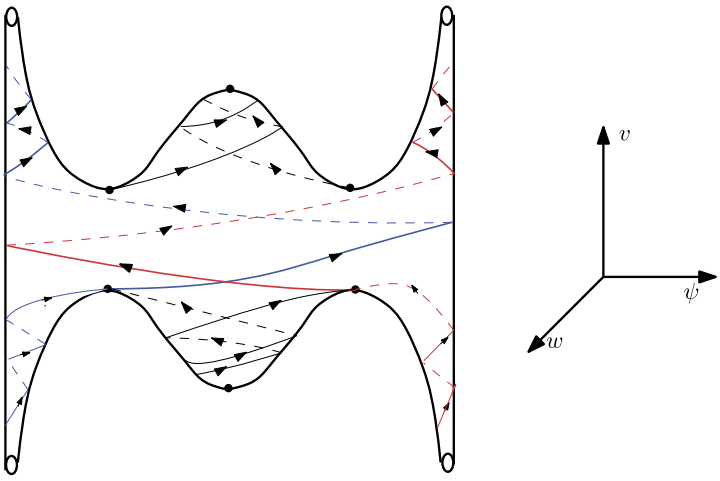

FIGURE 9. The collision manifold

THEOREM 5.1 ([10]). (1) *The collision $\mathcal{M}_0$ is a topological 2-sphere with four punctures (called arms) with each puncture corresponding to $\psi = \pm\pi/2$, i.e. the binary is infinitely far away from Q_1. The manifold $\mathcal{M}_0$ is symmetric with respect to $\psi \mapsto -\psi$ or $w \mapsto -w$.*

(2) *There are six fixed points on $\mathcal{M}_0$: a sink, a source and four saddles (see Figure 9). The sink and the source correspond to Euler collinear central configurations and the four saddles correspond to Lagrangian equilateral central configurations. The eigenvalues of the sink and sources are real for some masses, but can be complex for the others.*

(3) *v is a Lyapunov function in the sense that $v' \geq 0$ along the flow on $\mathcal{M}_0$ and $v' = 0$ iff at the fixed points.*

5.2.2. *Devaney's classification of final motions.* Devaney's classification of the final motions for the isosceles three-body problem mentioned at the beginning of this subsection can be seen as follows. For each orbit getting close to the triple collision but not a total collapse, its projection to the collision manifold corresponds to an orbit which either escapes along one of the two upper arms or dies in the sink. The sink lies in the upper half space and corresponds to an Euler central configuration. For an orbit dying in the sink, we have that $v > 0$ hence $r = r(0)e^{\int v}$ grows exponentially. Therefore we get the first type of final motion. For an orbit escaping along an upper arm, we will have $x_0/x_1 \to 0$ therefore the pair moves away from the third body experiencing infinitely many double collisions corresponding to $\psi = \pm\pi/2$.

5.2.3. *The acceleration mechanism.* We next examine closer the second type of final motion in Devaney's classification. In this section, we show that

it provides a mechanism of acceleration different from the one we explained in Section 3.

Consider the situation that the binary Q_1-Q_2 comes from the above towards the near triple collision with Q_3. We hope that after the near triple collision the binary moves back in the positive y-direction and Q_3 moves to the negative y-direction. In this case the relative position x_1 from the mass center of the binary to Q_3 has a negative sign before and after the near triple collision, therefore the variable ψ determined from $\mathsf{s} = M^{-1/2}(\cos\psi, \sin\psi)$ and $x = (x_0, x_1) = r\mathsf{s}$ should also carry a negative sign before and after the near triple collision. The following result of [**33**, **37**] gives a description of the stable and unstable manifolds of the Lagrangian fixed points on the collision manifold.

THEOREM 5.2. *Assume $m < 55/4$ in which case the Euler fixed points are sink and source with complex eigenvalues. Then there exists $\varepsilon_1 \simeq 0.379$ and $\varepsilon_2 \simeq 2.662$ such that the following holds for the left lower Lagrangian fixed point:*

(1) *when $\varepsilon_1 < m < \varepsilon_2$, the two unstable manifolds escape from the two upper arms respectively;*
(2) *when $m > \varepsilon_2$, one of the unstable manifolds escapes from the left upper arm and the other dies at the upper Euler fixed point.*

Note that the above two cases include the equal mass case $m = 1$.

Now consider the three-body problem with mass ratio as in the theorem, say $m = 1$ and suppose the binary approaches Q_3 as above and gets very close to the left lower Lagrange fixed point O with v-coordinate $v_\star < 0$. Denote by $W^s(O)$ the two dimensional stable manifold of O of the three-body problem (including the r-component, not restricted to the collision manifold) on an energy level. It is standard result from the hyperbolic dynamics that by choosing the initial condition sufficiently close to the stable manifold $W^s(O)$, one can arrange the orbit to stay within a small neighborhood of the fixed point for a time that can be as long as we wish before leaving the neighborhood along the unstable manifold. When this happens, from the r-equation, we get $r(\tau) = r(0)e^{\int_0^\tau v(s)\,ds} \simeq r(0)e^{v_\star \tau}$ so r can be as small as we wish provided we make the sojourn time τ large. For small r, the dynamics of the three-body problem shadows that on the collision manifold, so the orbit leaves a neighborhood of the triple collision along the unstable manifold of the left lower Lagrange fixed point. When the orbit arrives at a section $\{v = v_+\}$ for some $v_+ > 0$ large, we see from the definition $v = r^{-1/2}\langle x, \dot{x}\rangle$ of v that $|\dot{x}| \sim O(r^{-1/2})$ that can be made as large as we wish by making r small.

5.3. The collinear three-body problem and the work of Mather-McGehee. In this section, we consider the triple collision blowup for a collinear three-body problem studied by [**21**]. Suppose on $\mathbb{R}$ there are three

bodies ordered from left to right as Q_1, Q_2, Q_3 with masses m_1, m_2, m_3 respectively. We fix the mass center at the origin so that we have $\sum m_i Q_i = 0$. Therefore the set $\mathcal{E} = \{\|\mathsf{s}\| = 1\}$ is a circle. The general formalism in Section 5.1 applies to yield the blowup coordinates and the equations of motion formally similar to (5.3). We skip the details and refer readers to [**21**]. Here we only describe the dynamics on the collision manifold. As in the case of isosceles three-body problem, the Hamiltonian system of the collinear three-body problem has two degrees of freedom. Hence the collision manifold has two dimensions after imposing the energy relation and letting $r \to 0$. It is a sphere with four punctures corresponding to double collisions of the binaries Q_1-Q_2 or Q_2-Q_3, symmetric under the reflections as in Theorem 5.1. There are two critical points, each corresponds to an Euler central configuration and both are saddles, which is the main difference from Devaney's isosceles three-body problem. There are masses such that the unstable manifold of the lower critical point coincides with the stable manifold of the upper critical point, which is called the totally degenerate case. If this case does not happen, then the stable manifolds of the lower critical point come in from the lower arms and the unstable manifolds escape from the upper arms. Taking into consideration of the r-component, the stable manifold of the lower fixed point O is two dimensional, denoted by $W^s(O)$, when restricted to a fixed energy level. We can further take a section defined by $\mathcal{S}_- = \{v = v_-\}$ where v_- is sufficiently negative number so that before entering the section, the system is away from triple collision and can be considered as perturbed Kepler problems.

With this, we explain the idea of the proof of [**25**]. Consider a collinear four-body problem in which the four particles are labeled as Q_1, Q_2, Q_3, Q_4 from left to right. In this model, the binary Q_3-Q_4 has negative energy $E_{34} := \frac{|P_3|^2}{2m_3} + \frac{|P_4|^2}{2m_4} - \frac{m_3 m_4}{|Q_3 - Q_4|}$, so it can be considered as an almost Kepler elliptic motion, unless the triple Q_2-Q_3-Q_4 is close to a triple collision. We only need two variables L, ℓ to characterize the relative motion of Q_3 and Q_4, where L is the semi-major and ℓ characterizes the relative position of the two bodies. The two other variables (G, g) degenerate to zero in the collinear case. With the two variables L, ℓ, we still have the Hamiltonian equations $\begin{cases} \dot{L} = 0 \\ \dot{\ell} = \frac{mk^2}{L^2} \end{cases}$ for the Q_3-Q_4 dynamics up to perturbations from other bodies, so that the same analysis as in Section 4.3.3 applies here, that is, by adjusting the energy of the binary Q_3-Q_4 slightly after the near triple collision, one can adjust the relative phase of the binary significantly when Q_2 comes to the next $\mathcal{S}_-$.

Suppose next that Q_1 is far to the left and we focus on the triple Q_2-Q_3-Q_4. We introduce a variable ψ as coordinate on $\mathcal{E}$ measuring the relative positions of the triple on $\mathcal{E}$, similar to the one used in Section 5.2.1. The value of ψ on the section $\mathcal{S}_-$ can be adjusted by varying the relative phase of Q_3-Q_4. Moreover, we can use ψ as a measurement of the distance

to the stable manifold $W^s(O)$, thus by varying ψ, we can adjust how much time the near collision orbit stays within a neighborhood of the lower fixed point O before escaping along an upper arm shadowing an unstable manifold. Therefore, for the same reason as Section 5.2.3, the velocities of the triple can be as large as possible after the near triple collision. This gives the acceleration mechanism for the collinear four-body problem.

5.4. The Sitnikov-Alekseev model. The Sitnikov-Alekseev model is a name for the spatial isosceles three-body problem.

5.4.1. *Triple collision blowup for the spatial isosceles three-body problem.* In this section, we consider the triple collision blowup for a spatial isosceles three-body problem studied by [**42**].

Consider three particles Q_1, Q_2, Q_3 with coordinates $Q_1 = (x, y, z)$ then $Q_2 = (-x, -y, z)$ and $Q_3 = (0, 0, z_3)$ with $2mz + m_3 z_3 = 0$. The three bodies form an isosceles configuration. Let $M = \mathrm{diag}\{2m, 2m, 2m(1+2\alpha)\}$, where $\alpha = m/m_3$ is the mass ratio. Using the formalism in Section 5.1, we introduce blowup coordinates

$$\mathsf{x} = (x, y, z),\ r = \|\mathsf{x}\|,\ \mathsf{s} = r^{-1}\mathsf{x},\ \mathsf{v} = r^{1/2}\dot{\mathsf{x}},\ v = \langle \mathsf{v}, \mathsf{s}\rangle,\ \mathsf{w} = \mathsf{v} - v\mathsf{s},$$

and rescale time $dt = r^{3/2}d\tau$. We use $'$ to denote the time derivative with respect to τ. We first obtain the equations of motion for $(r, \mathsf{s}, v, \mathsf{w})$ by adapting (5.1). We next introduce spherical coordinates ($\psi \in [-\pi/2, \pi/2]$ and $\theta \in [0, 2\pi)$) for s and denote

$$\begin{aligned} \mathbf{u}_1 &= \mathsf{s} = M^{-1/2}(\cos\theta\cos\psi, \sin\theta\cos\psi, \sin\psi), \\ \mathbf{u}_2 &= \partial_\theta \mathsf{s} = M^{-1/2}(-\sin\theta\cos\psi, \cos\theta\cos\psi, 0), \\ \mathbf{u}_3 &= \partial_\psi \mathsf{s} = M^{-1/2}(-\cos\theta\sin\psi, -\sin\theta\sin\psi, \cos\psi). \end{aligned} \tag{5.5}$$

The three vectors for orthogonal (but not orthonormal) basis for $\mathbb{R}^3$ except for $\psi \neq \pm\pi/2$. We introduce variables v, w_2, w_3 as the coefficients of the decomposition of v in this basis with respect to the inner product $\langle\cdot,\cdot\rangle$, i.e. $\mathsf{v} = v\mathsf{s} + w_2\mathbf{u}_2 + w_3\mathbf{u}_3$. We thus obtain the equations of motion for the variables $(r, v, \psi, \theta, w_2, w_3)$ as follows.

$$\begin{cases} r' &= vr, \\ v' &= \frac{1}{2}v^2 + w_2^2\cos^2\psi + w_3^2 + U(\psi), \\ \psi' &= w_3, \\ \theta' &= w_2, \\ w_2' &= -\frac{1}{2}vw_2 + 2w_2w_3\tan\psi \\ w_3' &= -U'(\psi) - \frac{1}{2}vw_3 - w_2^2\cos^2\psi\tan\psi. \end{cases}$$

with

$$U(\psi) = -\frac{1}{\sqrt{2}}m^{3/2}m_3[\alpha\sec\psi + 4(1 + 2\alpha\sin^2\psi)^{-1/2}].$$

The system has singularities at $\psi = \pm\pi/2$ corresponding to double collisions of Q_1 and Q_2. The RHS of the system is independent of θ, reflecting the

angular momentum conservation of the pair Q_1-Q_2. We resolve this singularity by introducing $w = w_3 \cos\psi$ and $u = w_2 \cos^2\psi$ and a time change $d\hat{\tau} = \cos\psi d\tau$. In the new coordinates, we obtain the equations of motion as follows

(5.6)
$$\begin{cases} \frac{d}{d\hat{\tau}} r &= vr\cos\psi \\ \frac{d}{d\hat{\tau}} v &= (2rH - \frac{1}{2}v^2 - U(\psi))\cos\psi \\ \frac{d}{d\hat{\tau}} \psi &= w \\ \frac{d}{d\hat{\tau}} w &= -U'(\psi)\cos^2\psi - \frac{1}{2}vw\cos\psi - (2U(\psi) + 2rH - v^2)\sin\psi\cos\psi \\ \frac{d}{d\hat{\tau}} u &= -\frac{1}{2}vu\cos\psi \end{cases}$$

with the energy relation

$$\frac{1}{2}(v^2\cos^2\psi + w^2 + u^2) + U(\psi)\cos^2\psi = rH\cos^2\psi,$$

where we skip the θ-equation since the RHS of the system has no dependence on θ. This resolves the singularity of double collision and result in an analytic vector field on the phase space.

The dynamics on the collision manifold is obtained by setting $r = 0$ in the above equations. The resulting collision manifold is a three dimensional sphere with four punctures, each of which corresponds to the double collision between the pair Q_1-Q_2. When $u = 0$, the $\frac{du}{d\hat{\tau}}$ equation also becomes trivial. The variable u can be considered as a variant of the rescaled angular momentum of the pair Q_1-Q_2, where the rescaling is to multiply by $r^{-1/2}$, so the vanishing of u means that the two bodies Q_1 and Q_2 move on a line. Therefore the case $u = 0$ corresponds to the planar isosceles three-body problem of Devaney. So we may understand the collision manifold of the spatial isosceles three-body problem as a fattening of the collision manifold of the planar isosceles three-body problem with the extra dimension parametrized by u. The equations of motion (5.6) is reduced to (5.4) by setting $u = 0$. In the planar case, we have been focusing on the Lagrangian fixed points in the lower half space (Section 5.2.3), whose v-value is negative and $\psi \neq \pm\pi/2$. From the u-equation, we see that if we stay close to the lower Lagrangian fixed points for a long time, the value of the u-variable will grow exponentially, meaning that the system will be pushed away from a neighborhood of the planar isosceles three-body problem. The physical meaning of this exponential growth of u is as follows. When Q_3 gets expelled with a large velocity, the semimajor of the binary Q_1-Q_2 will become smaller due to the loss of energy. Since the angular momentum of Q_1-Q_2 is conserved, the eccentricity $\sqrt{1 - \frac{G^2}{L^2}}$ becomes farther away from 1. If we rescale the system such that the semimajor is rescaled to unit size, then the semiminor becomes larger than that before the close encounter.

The model of [42] is to consider a five-body problem $Q_i,\ i = 1, \dots, 5$ such that two pairs Q_1-Q_2 and Q_4-Q_5 each has orbital plane parallel to the x-y plane and Q_3 travels back and forth between the two pairs. When Q_3 gets close to each pair, the triple gets close to triple collision to get accelerated using a similar mechanism to Section 5.2.3 and 5.3. Moreover, the timing of the close encounter of Q_3 with each pair can be controlled using similar phase stretching mechanism that we explained in Section 4.3.3 and 5.3. Then the new difficulty is the exponential growth of the u-variable, since large u will bound the orbit away from the collision manifold so that further acceleration is not possible. There is an observation in [42] that gives a mechanism to decrease u. Indeed, each pair is highly eccentric so can be thought of being collinear for simplicity. When the two lines form an angle that is not 0, $\pi/2$, say, $\pi/4$, each pair gives a nontrivial torque to the other. We suppose initially the angular momentum of both pairs cancel so they do so for all later times. The torque implies that the time derivative of the positive angular momentum carries a negative sign so that the positive angular momentum decreases. However, as time approaches the singular time, the two pairs get farther and farther so the torque gets weaker and weaker. So one should start with very tiny initial angular momenta of the pairs, whose values are a delicate balance of the growth of u and the decay of the torque along the singular orbit, which in turn depends on how close the orbit approaches the lower Lagrange fixed point during each visit to near triple collision.

5.4.2. *The oscillatory motion in Sitnikov-Alekseev model.* As a digression, we mention that this spatial isosceles three-body problem exhibits another interesting dynamics called oscillatory motion in the classification of Chazy (see Section 1.1). In other words, there exists an orbits $Q(t)$ in the configuration space such that

$$\limsup_{t\to\infty} |Q(t)| = \infty, \quad \liminf_{t\to\infty} |Q(t)| < C.$$

Such motion was first discovered by Sitnikov in the following model: a pair of equal masses Q_1 and Q_2 moving on the x-y plane along Kepler elliptic orbits and a massless particle moving on the z-axis attracted by the pair. Alekseev then showed that the same phenomenon occurs in the spatial isosceles three-body problem removing the massless assumption. The mechanism of having the oscillatory motion is to treat the Kepler parabolic orbit as a homoclinic orbit to the (degenerate) hyperbolic fixed point at infinity. Then a small perturbation will cause separatrix splitting known to Poincaré. Suppose we consider the motion of a massless particle Q_3 moving on the z-axis under the gravitation of two equal masses Q_1 and Q_2 on the x-y plane doing Kepler elliptic motion. We shall treat the motion of Q_3 as a Kepler problem attracted by the mass center of Q_1 and Q_2, then perturbed periodically by Q_1-Q_2. The Kepler parabolic orbit can be considered as a homoclinic orbit to infinity and a small periodic perturbation will create chaos from it. Let

us first consider the Kepler two-body problem written in polar coordinates

$$H(r,R,\theta,\Theta)=\frac{1}{2}R^2-\frac{1}{r}+\frac{\Theta^2}{2r^2}$$

with equations of motion $\begin{cases}\dot r=R\\ \dot R=-\frac{1}{r^2}+\frac{\Theta^2}{r^3}\end{cases}$. We next introduce the McGehee transform $r=u^{-2}$ so that $r=\infty$ corresponds to $u=0$. In terms of u and R we have the following equations of motion $\begin{cases}\dot u=-\frac{1}{2}u^3R\\ \dot R=-u^4+\Theta^2u^6\end{cases}$. If we make a time rescaling $d\tau=u^3dt$, then we get $\begin{cases}\frac{du}{d\tau}=-\frac{1}{2}R\\ \frac{dR}{d\tau}=-u+\Theta^2u^3\end{cases}$. Therefore the point $(u,R)=(0,0)$ is a hyperbolic fixed point with coinciding stable and unstable manifold corresponding to the Kepler parabolic orbits. Under the periodic perturbation by Q_1-Q_2, the stable and unstable manifold generally split. The existence of oscillatory motion is then proved by evaluating the Melnikov function along the parabolic Kepler orbit $\text{Graph}\left\{R(r)=\pm\sqrt{\frac{2}{r}-\frac{\Theta^2}{r^2}}\right\}$. We refer readers to [5, 24] for more details.

Appendix A. Painlevé's and von Zeipel's theorems

In this section, we present the proof of two early results of Painlevé and von Zeipel on noncollision singularities. The proofs of the two theorems contain some deep insights about the global aspects of the set of noncollision singularities.

A.1. Preliminary: Jacobi coordinates and moment of inertia. We introduce the following important quantity called momentum of inertia

$$I=\sum m_i|Q_i|^2=\frac{1}{\sum m_i}\sum_{i,j}m_im_j|Q_i-Q_j|^2. \tag{A.1}$$

The time derivative of $\dot I$ is as follows

$$J:=\dot I/2=\sum m_iQ_i\cdot\dot Q_i.$$

$$\dot J=\sum m_i|\dot Q_i|^2+m_iQ_i\cdot\ddot Q_i=\sum m_i|\dot Q_i|^2+Q_i\cdot(-\nabla_iU)=\sum m_i|\dot Q_i|^2+U$$

by Euler identity since U is (-1)-homogeneous. Denoting by $K=\frac{1}{2}\sum m_i|\dot Q_i|^2$ the kinetic energy, we thus have the important Lagrange-Jacobi identity

$$\dot J=\ddot I/2=2K+U=K+H=2H-U. \tag{A.2}$$

When $H\geq 0$, we get that $\dot J\geq 0$ so J is a nondecreasing Lyapunov function.

For N-body problem, due to the translation invariance, we can fix the mass center of the N bodies at zero, i.e. $\sum m_iQ_i=0$. This implies the total momentum conservation $\sum P_i=0$ by taking time derivative. Therefore we can remove (Q_i,P_i) for some i to reduce the dimension of the phase space.

There is a classical way to introduce a new set of coordinates maintaining the diagonal quadratic form of the kinetic energy part and effectively removing one redundant body.

We introduce the Jacobi coordinates for three-body problem for simplicity.

$$\text{(A.3)}\qquad \begin{cases} x_0 = Q_1 - Q_2 \\ x_1 = Q_3 - \frac{m_1 Q_1 + m_2 Q_2}{m_1+m_2} \end{cases}, \quad \begin{cases} y_0 = P_2 - \frac{m_1}{m_1+m_2} P_3 \\ y_1 = P_3 \end{cases}.$$

The physical meaning is as follows. We first take the relative position of Q_1 and Q_2 as one position vector x_0 and then take the second position vector x_1 to be the distance fo Q_3 to the mass center of the Q_1 and Q_2. The momentum vectors y_0 and y_1 are chosen to make the symplectic form invariant (using the fact $\sum P_i = 0$)

$$dx_0 \wedge dy_0 + dx_1 \wedge dy_1 = \sum dQ_i \wedge dP_i.$$

Moreover, when introducing the reduced masses $M_0 = \frac{m_1 m_2}{m_1+m_2}$, $M_1 = \frac{m_3(m_1+m_2)}{m_1+m_2+m_3}$, we have the moment of inertia in Jacobi coordinates

$$I = \sum m_i |Q_i|^2 = M_0 |x_0|^2 + M_1 |x_1|^2.$$

In this coordinate system the Hamiltonian has the form

$$\text{(A.4)}\qquad \begin{aligned} H &= \sum_{i=1}^{3} \frac{P_i^2}{2m_i} - \sum_{i\neq j} \frac{m_i m_j}{|Q_i - Q_j|} \\ &= \frac{y_0^2}{2M_0} + \frac{y_1^2}{2M_1} - \frac{m_1 m_2}{|x_0|} - \frac{m_1 m_3}{|x_1 - \alpha_1 x_0|} - \frac{m_2 m_3}{|x_1 + \alpha_0 x_0|}, \end{aligned}$$

where $\alpha_0 = \frac{m_1}{m_1+m_2}$, $\alpha_1 = \frac{m_2}{m_1+m_2}$.

A.2. Painlevé's theorem. In this section, we prove the following theorem of Painlevé.

THEOREM A.1 (Painlevé). *For $N = 3$, all the singularities are collisions.*

PROOF. We first state a lemma which is valid for general N-body problem.

LEMMA A.2. *In the N-body problem with $N > 2$, suppose a singularity occurs as $t \to t^*$, then we have*

$$\lim_{t\to t^*} \min_{i\neq j} |Q_i(t) - Q_j(t)| = 0.$$

With this lemma, we prove Painlevé's theorem.

If we also had $\limsup_{t\to t^*} \max_{i\neq j} |Q_i(t) - Q_j(t)| = 0$, then this implies that the singularity is a triple collision. So we have proved the statement. We now assume

$$\text{(A.5)}\qquad \limsup_{t\to t^*} \max_{i\neq j} |Q_i(t) - Q_j(t)| > a > 0.$$

We next show that for $|t-t^*|<\delta$ sufficiently small, we have the fixed pair, say Q_1 and Q_2, has $|Q_1(t)-Q_2(t)|\to 0$. Otherwise, by Lemma A.2, there is a moment sufficiently close to t^* such that $\min_{i\neq j}|Q_i(t)-Q_j(t)|$ is attained by different pairs, say $|Q_1(t)-Q_2(t)|=|Q_1(t)-Q_3(t)|=\varepsilon<a/3$. Then by triangle inequality, we get a contradiction to (A.5).

Thus we get that for δ small enough,

$$|Q_1(t)-Q_2(t)|<\varepsilon,\ \forall\ t>t^*-\delta$$

and also there is a subsequence $t_n>t^*-\delta$ such that

$$|Q_3(t_n)-Q_i(t_n)|>a,\quad i=1,2.$$

By Lagrange-Jacobi identity (A.2), we get $\ddot{I}>0$ for $t>t^*-\delta$ since the potential $U\to-\infty$ by Lemma A.2. This implies $I(t)$ has a limit (may be $+\infty$) as $t\to t^*$. Taking into account (A.5) and (A.1), we get $I(t)>a^2\min\{m_i\}$ for all $t>t^*-\delta$. By triangle inequality and (A.1), this implies $|Q_3(t)-Q_i(t)|>Ca$, $i=1,2$. for all $t>t^*-\delta$, where the constant depends only on the masses. Then from the Hamiltonian equation, we get $|\ddot{Q}_3|\leq C/a$. Integrating it twice, we get that the position $Q_3(t)$ has a limit $Q_3(t^*)$ that is not infinity. Therefore, the singularity at t^* is a collision. This is a contradiction. □

We next work on the proof of Lemma A.2, which needs the following lemma.

Lemma A.3. *In the N-body problem, suppose a singularity occurs as $t\to t^*$, then we have*

$$\liminf_{t\to t^*}\min_{i\neq j}|Q_i(t)-Q_j(t)|=0.$$

Proof of Lemma A.2. From Lemma A.3, it is enough to prove that

$$\limsup_{t\to t^*}\min_{i\neq j}|Q_i(t)-Q_j(t)|=0.$$

Suppose the limsup is greater than zero and we choose a number $a>0$ less than the limsup. Then for each δ, there exists $t_1\in[t^*-\delta,t^*)$ such that $|Q_i(t_1)-Q_j(t_1)|>3/4a$ for all $i\neq j$ and there exists $t_2\in(t_1,t^*)$ such that $|Q_i(t_2)-Q_j(t_2)|<1/4a$ for all $i\neq j$. We choose t_2 to be the least of such time so that in the interval $[t_1,t_2]$ we have $|Q_i(t)-Q_j(t)|\geq 1/4a$ for all $i\neq j$. This implies by the energy conservation that in this time interval

$$\sum P_i^2/(2m_i)\leq H+4N^2M^2/a.$$

hence each $\dot{Q}_i=P_i/m_i$ is bounded from above by a constant C depending only on the masses, H, a, N. Within time δ, the oscillations of the positions Q_i are bounded by $C\delta$. For δ sufficiently small, this gives $|Q_i(t_2)-Q_j(t_2)|\geq 1/2a$ for all $i\neq j$ if we have $|Q_i(t_1)-Q_j(t_1)|>3/4a$. This contradicts to the assumption that $|Q_i(t_2)-Q_j(t_2)|<1/4a$ for all $i\neq j$. □

PROOF OF LEMMA A.3. Suppose $\liminf = a > 0$ and choose $0 < b < a$. Then there exists $\delta > 0$ sufficiently small such that $|Q_i(t) - Q_j(t)| > b$ for all $i \neq j$ and $t \in [t^* - \delta, t^*)$. From the Hamiltonian we obtain

$$\frac{1}{2}\sum_i P_i^2/m_i \leq H + N^2M^2/b.$$

So each velocity has an upper bound. Moreover, from the equation $\dot{P}_i = \sum_{i\neq j} \frac{m_i m_j (Q_j - Q_i)}{|Q_j - Q_i|^3}$, we obtain that $|\dot{P}_i| \leq NM^2/b^2$. Within time δ, the oscillation of Q_i and P_i are then both of order δ. This means $\lim_t Q_i(t)$, $\lim_t P_i(t)$ both exists as $t \to t^*$ and there is no collision at all. This is a contradiction. □

A.3. von Zeipel's theorem. In this section, we present the proof of von Zeipel's theorem.

THEOREM A.4 (von Zeipel). *If a noncollision singularity occurs at t^* then we have*

$$\lim_{t\to t^*} \min_{i\neq j} |Q_i(t) - Q_j(t)| = 0, \quad \lim_{t\to t^*} \max_{i\neq j} |Q_i(t) - Q_j(t)| = \infty.$$

PROOF. The first part is given by Lemma A.2, so it is enough to prove the second one. We prove it by contradiction. Suppose we have an noncollision singularity at time t^* such that $|Q_i| < C$ for all $i = 1, \ldots, N$. Let $\Delta^* \subset \mathbb{R}^{3N}$ be the set of accumulation points of $Q(t) = (Q_1, \ldots, Q_N)(t)$ as $t \to t^*$. Then we have for each $Q \in \Delta^*$, there is some $i \neq j$ such that $Q_i = Q_j$.

It is clear that Δ^* is a nonempty compact set. Moreover, it cannot be a single point which can only happen for a total collision.

The hierarchy decomposition with respect to a partition. For each point $p \in \Delta^*$, we introduce a partition $\xi(p)$ of $\{1, \ldots, N\}$ as follows. The label i and j are in the same element of ξ if $Q_i = Q_j$ at p. It may happen that an element of $\xi(p)$ has only one label. Moreover, the partition may not be unique for each point p. In that case, we choose a partition with minimal cardinality.

For a given partition $\xi(p)$, we perform a decomposition of the moment of inertia and the potential as follows. Let $S \in \xi(p)$ be an element of the partition, then we have

$$I = I^\xi + I_\xi, \quad I_\xi = \sum_{S\in\xi} I_S, \tag{A.6}$$

where $I_S = \sum_{i\in S} m_i |Q_i - c_S|^2$ is the moment of inertia for S, where we have introduced the mass center $c_S = \frac{1}{m_S}\sum_{i\in S} m_i Q_i$, $m_S = \sum_{i\in S} m_i$, and $I^\xi = \sum_{S\in\xi} m_S |c_S|^2$ is the moment of inertia on the level of partition. We

next introduce a similar decomposition of the potential $U = -\sum_{i<j} \frac{m_i m_j}{|Q_i - Q_j|}$

$$U = U^\xi + \sum_{S \in \xi} U_S$$

where $U_S = -\sum_{i,j} \frac{m_i m_j}{|Q_i - Q_j|}$ where the sum is taken over all $i, j \in S$ with $i < j$ and $U^\xi = -\sum_{i,j} \frac{m_i m_j}{|Q_i - Q_j|}$ where the sum is taken over all i and j in different elements of the partition ξ with $i < j$.

We next introduce an orthogonal decomposition of $\mathbb{R}^{3N}$ into $E_0(\xi) \oplus E_-(\xi)$ with respect to the metric $\langle u, v\rangle = \sum m_i u_i v_i,\ \forall\ u,\ v \in \mathbb{R}^{3N}$. For each $Q = (Q_1, \dots, Q_N) \in \mathbb{R}^{3N}$, we decompose it into $Q = Q_0 + Q_-$ where $Q_0 = (c_S) \in E_0(\xi)$ collects the mass centers for each element $S \in \xi$ and $Q_- = ((Q_i - c_S)_{i \in S}) \in E_-(\xi)$ is the collections of relative positions to the mass center c_S for each element S. Note that for each S, the vectors $(Q_i - c_S)_{i \in S}$ are not linearly independent unless we remove one vector from them. We do so and still denote by Q_- the resulting vector. The above equation (A.1) is the Pythagorean law about the orthogonality of $E_0(\xi)$ and $E_-(\xi)$. Now the transformation from $Q \mapsto (Q_0, Q_-)$ is linear and we have a dual linear transformation for the momentum $P \mapsto (P_0, P_-)$ so that $dP \wedge dQ = dP_0 \wedge dQ_0 + dP_- \wedge dQ_-$. This is actually the Jacobi coordinates for each element S of the partition ξ. Note that in this decomposition U_S depends only on vectors from Q_- and U^ξ depends on vectors from both Q_0 and Q_-.

Claim 1. *At a noncollision singularity, the orbit shadows more than one partition for points in* Δ^*.

Proof. Suppose there is only one such partition. Then using the above hierarchy decomposition, we see that $|U^\xi|$ is bounded as $t \to t^*$ since $|Q_i - Q_j|$ is bounded from below on the compact set Δ^* for i, j not in the same element of the partition. This also holds in a neighborhood of Δ^*. Then we get $M_S \ddot{c}_S = -\nabla_{c_S} U^\xi$. The boundedness of $|\nabla_{c_S} U^\xi|$ implies that within time δ, the oscillation of c_S is of order δ so that $Q_0(t) \to Q_0(t^*)$ as $t \to t^*$. On the other hand, by definition we have $Q_-(t) \to 0$ as $t \to t^*$. This gives a total collapse for each element $S \in \xi$ at time t^*. □

Claim 2. *If at a noncollision singularity we have*

$$\limsup_{t \to t_*} \max_{i \neq j} |Q_i(t) - Q_j(t)| < \infty,$$

then $I(t)$ *converges to* $I_* < \infty$.

Proof. By Lagrange-Jacobi identity, we have $\ddot{I} = K + H$, since we have $\min_{i \neq j} |Q_i(t) - Q_j(t)| \to 0$ by Lemma A.2, this implies $K \to \infty$ so $\ddot{I} > 0$ for t close to t^*. So I is convex hence continuous, so it has a limit I_* as $t \to t^*$, which by assumption is bounded. □

Let $\Delta^*_\xi := \Delta^* \cap E_0(\xi)$, then $Q(t)$ has accumulation points both in and not in the set Δ^*_ξ as $t \to t^*$.

CLAIM 3. *If ξ is a minimal partition then the set Δ^*_ξ is closed.*

PROOF. Let p be a an accumulation point of Δ^*_ξ. Then there is a sequence $p_n \in \Delta^*_\xi$ such that $p_n \to p$. By definition, each p_n itself is an accumulation point of $Q(t), t \to t^*$. Taking a Cantor diagonal subsequence we see that p is also accumulated by $Q(t), t \to t^*$. This implies that $p \in \Delta^*$. It remains to prove that p has the partition ξ. For each $i, j \in S \in \xi$, we get for $Q_i = Q_j$ the point p since we have $Q_i = Q_j$ for all the points p_n. This implies that the partition corresponding to the point p is the same as ξ or coarser. Since we have chosen ξ to be minimal, the partition for p can only be ξ. This implies that $p \in E_0(\xi)$ so we have $p \in \Delta^*_\xi$. □

By assumption Δ^* is bounded so Δ^*_ξ is compact. We choose a neighborhood V_0 of Δ^* in $E_0(\xi)$ such that both $|U^\xi|$ and $|\nabla_{c_S} U^\xi|$ are bounded. We next extend V_0 to a neighborhood of Δ^*_ξ in $\mathbb{R}^{3n}$ of the form $V = V_0 \times B_\sigma$ where B_σ is a ball of radius σ centered at zero in the space $E_-(\xi)$. We choose σ small to achieve the following two goals.

(1) By Claim 1, $Q(t)$ has to enter and exit the neighborhood V infinitely often as $t \to t^*$.
(2) $\partial V_0 \times \bar{B}_\sigma$ does not intersect Δ^* so that the orbit $Q(t)$ does not intersect $\partial V_0 \times \bar{B}_\sigma$ for $t \in [t^* - \delta, t^*)$ for δ small enough. So the orbit can only visit V through the boundary $V_0 \times \partial B_\sigma$

The construction gives the following claim.

CLAIM 4. *There exists infinitely time intervals $[t_1, t_3] \subset [t^* - \delta, t^*)$ such that $I_\xi(t_1) = I_\xi(t_3) = \sigma^2$ and $t_2 \in (t_1, t_3)$ such that $I_\xi(t_2) = \sigma^2/2$.*

Choose δ small such that $|I(t) - I^*| < \sigma^2/6$ for all $t \in [t^* - \delta, t^*)$ and let t_1, t_2, t_3 be as in Claim 4. Then we have the following claim.

CLAIM 5. *Within each interval (t_1, t_3), there exists $\hat{t}$ such that $I^\xi(\hat{t}) \geq I^\xi(t_2)$ and $I^\xi(t_3) \geq I^\xi(\hat{t}) - \frac{C}{2}\delta^2$.*

PROOF. We have

$$I^\xi(t_2) = I(t_2) - I_\xi(t_2) \geq I^* - \frac{\sigma^2}{6} - \frac{\sigma^2}{2} = I^* - \frac{2\sigma^2}{3},$$

$$\max\{I^\xi(t_1), I^\xi(t_3)\} \leq I^* + \frac{\sigma^2}{6} - \sigma^2 = I^* - \frac{5\sigma^2}{6}.$$

This implies that there is a $\hat{t} \in [t_1, t_3]$ such that $I^\xi(t)$ attains a local max at $\hat{t}$ so we have $\dot{I}^\xi(\hat{t}) = 0$. We bound $I^\xi(t)$ by $I^* + \frac{\sigma^2}{6}$ for all $t \in [\hat{t}, t^*)$ hence get a bound for $|c_S|$. Within a small neighborhood U of Δ^*, we have $|U^\xi|$ and $|\nabla_{c_S} U^\xi|$ are both bounded, so we have

$$\ddot{I}^\xi = \sum_{S \in \xi} M_S |\dot{c}_S|^2 + c_S \cdot \nabla_{c_S} U^\xi \geq -C$$

where $C > 0$ bounds $|c_S \cdot \nabla_{c_S} U^\xi|$. Integrating over the time interval $[\hat{t}, t_3)$, we find that $I^\xi(t_3) \geq I^\xi(\hat{t}) - \frac{C}{2}\delta^2$. □

Now it is clear that Claim 2, 4, 5 give a contradiction by choosing δ sufficiently small and σ fixed. The contradiction implies that $\limsup_{t\to t_*} \max_{i\neq j} |Q_i(t) - Q_j(t)| = \infty$, i.e. $\limsup_{t\to t_*} I(t) \to \infty$. In the proof of Claim 2, we have proved that $\ddot{I} > 0$. The convexity of I implies that we have indeed

$$\lim_{t\to t_*} \max_{i\neq j} |Q_i(t) - Q_j(t)| = \infty. \qquad \square$$

Acknowledgment

I would like to thank the organizers of Current Developments in Mathematics 2020 for inviting me to take part in the conference. This work is supported by NSFC (Significant project No.11790273) in China and Beijing Natural Science Foundation (Z180003).

References

[1] V. I. Arnold, *Proof of a theorem by A.N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations of the Hamiltonian.* Russian Math. Surveys 18 (1963) 9–36. MR 0163025

[2] Agekyan, T. A., and Zh P. Anosova. *A Study of the Dynamics of Triple Systems by Means of Statistical Sampling.* AZh 44 (1967): 1261.

[3] Albouy, Alain, Hildeberto E. Cabral, and Alan A. Santos. *Some problems on the classical n-body problem.* Celestial Mechanics and Dynamical Astronomy 113.4 (2012): 369–375. MR 2970201

[4] Albouy, Alain, and Vadim Kaloshin. *Finiteness of central configurations of five bodies in the plane.* Annals of Mathematics (2012): 535–588. MR 2925390

[5] Arnold, Vladimir I., Valery V. Kozlov, and Anatoly I. Neishtadt. *Mathematical aspects of classical and celestial mechanics.* Vol. 3. Springer Science & Business Media, 2007. MR 2269239

[6] M. Brin, Stuck, *Introduction to Dynamical Systems*, CUP, 2002. MR 1963683

[7] Chenciner, Alain. *A l'infini en temps fini.* Séminaire Bourbaki 832 (1997): 323–353. MR 1627117

[8] Chenciner, Alain, and Richard Montgomery. *A remarkable periodic solution of the three-body problem in the case of equal masses.* Annals of Mathematics – Second Series 152.3 (2000): 881–902. MR 1815704

[9] Chierchia, Luigi, and Gabriella Pinzari. *The planetary N-body problem: symplectic foliation, reductions and invariant tori.* Inventiones Mathematicae 186.1 (2011): 1–77. MR 2836051

[10] Devaney, Robert L. *Triple collision in the planar isosceles three body problem.* Inventiones Mathematicae 60.3 (1980): 249–267. MR 0586428

[11] L. Floria, *a simple derivation of the hyperbolic Delaunay variables*, The Astronomical Journal, 110, No 2, (1995), 940–942.

[12] Féjoz, Jacques. *Démonstration du théorème d'Arnold sur la stabilié du système planétaire (d'après Herman).* Ergodic Theory and Dynamical Systems 24.5 (2004): 1521–1582. MR 2104595

[13] Gerver, Joseph L. *Noncollision Singularities: Do Four Bodies Suffice!* Experimental Mathematics 12.2 (2003): 187–198. MR 2016705

[14] Gerver, Joseph L. *Noncollision singularities in the n-body problem*, in Dynamical systems. Part I, 57–86, Pubbl. Cent. Ric. Mat. Ennio Giorgi, Scuola Norm. Sup., Pisa, 2003. MR 2071232

[15] Gerver, Joseph L. *The existence of pseudo-collisions in the plane*, J. Differential Equations 89 (1991), 1–68. MR 1088334
[16] Gerver, Joseph L. *A possible model for a singularity without collisions in the five-body problem*, J. Differential Equations 52 (1984), 76–90. MR 0737965
[17] Herman, Michael. *Some open problems in dynamical systems.* Proceedings of the International Congress of Mathematicians. Vol. 2. 1998. MR 1648127
[18] Hut, P., Bahcall, J. N. *Binary-single star scattering. I – Numerical experiments for equal masses.* Astrophys. J. 268, 319–341 (1983).
[19] Hampton, Marshall, and Anders Jensen. *Finiteness of spatial central configurations in the five-body problem.* Celestial Mechanics and Dynamical Astronomy 109.4 (2011): 321–332. MR 2783101
[20] Hampton, Marshall, and Richard Moeckel. *Finiteness of relative equilibria of the four-body problem.* Inventiones Mathematicae 163.2 (2006): 289–312. MR 2207019
[21] McGehee, Richard. *Triple collision in the collinear three-body problem.* Inventiones Mathematicae 27.3 (1974): 191–227. MR 0359459
[22] Moeckel, Richard. *Lectures on central configurations.* Notes of the Centre de Recerca Matematica, Barcelona (2014). MR 3469182
[23] Moser, Jürgen. *Dynamical Systems, past and present.* In Proc. Int. Congress of Math., vol. 1, pp. 381–402 (1998). MR 1660656
[24] Moser, J. *Stable and random motions in dynamical systems*, volume 77 of Annals of Mathematics Studies (1973). MR 0442980
[25] J. Mather, R. McGehee, *Solutions of the collinear four body problem which become unbounded in finite time*, Lecture Notes in Physics (J. Moser, ed.), vol. 38, pp. 575–597, Springer-Verlag, Berlin/New York, 1975. MR 0495348
[26] Moeckel, Richard, and Richard Montgomery. *Realizing all reduced syzygy sequences in the planar three-body problem.* Nonlinearity 28.6 (2015): 1919. MR 3350615
[27] Marchal, Christian, and Donald G. Saari. *On the final evolution of the n-body problem.* Journal of Differential Equations 20.1 (1976): 150–186. MR 0416150
[28] Paul Painlevé, *Leçons sur l'intégration des équations différentielles de la méchanique et applications.* Hermann, Paris, 1897.
[29] Paul Painlevé, *Leçons sur la théorie analytique des équations différentielles.* Hermann, Paris, 1897.
[30] D. G. Saari, *Singularities and collisions of Newtonian gravitational systems*, Arch. Rational Mech. Anal. 49 (1972) 311–320. MR 0339596
[31] D. G. Saari, *A global existence theorem for the four body problem of Newtonian mechanics*, J. Differential Equations 26 (1977) 80–111. MR 0478863
[32] Simon, Barry. *Fifteen problems in mathematical physics.* Perspectives in mathematics, Birkhäuser, Basel 423 (1984). MR 0779685
[33] Simó, C, *Analysis of triple collision in the isosceles problem*, Classical Mechanics and Dynamical Systems, 203–224, Marcel Dekker (1981). MR 0640127
[34] Smale, Steve. *Mathematical problems for the next century.* The Mathematical Intelligencer 20.2 (1998): 7–15. MR 1631413
[35] Standish, E. M. *The dynamical evolution of triple star systems.* Astron. Astrophys. 21, 185–191 (1972).
[36] Šuvakov, M., Dmitrašinović, V. *Three classes of Newtonian three-body planar periodic orbits.* Phys. Rev. Lett. 110, 114301 (2013).
[37] Simó, Carles, and Regina Martinez. *Qualitative study of the planar isosceles three-body problem.* Celestial Mechanics 41.1-4 (1987): 179–251. MR 0954883
[38] Siegel, Carl L., and Jürgen K. Moser. *Lectures on celestial mechanics.* Springer Science & Business Media, 2012. MR 1345153
[39] Wintner, Aurel. *The analytical foundations of celestial mechanics.* Courier Corporation, 2014. MR 0005824

[40] Xue, Jinxin, *Noncollision singularities in a planar 4-body prooblem*, Acta Mathematica. Volume 224 (2020), Number 2, 253–388. MR 4117052
[41] Xue, Jinxin, and Dmitry Dolgopyat. *Non-Collision Singularities in the Planar Two-Center-Two-Body Problem.* Communications in Mathematical Physics 345.3 (2016): 797–879. MR 3519584
[42] Z. Xia, *The existence of noncollision singularities in Newtonian systems*, Ann. of Math. (2) 135 (1992) 411–468. MR 1166640
[43] H. Von Zeipel, *Sur les singularites du probleme des n corps*, Ark. Mat. Astro. Fys. (4) 32 (1908) 1–4.

Yau Mathematical Sciences Center & Department of Mathematics, Tsinghua University, Beijing, China, 100084
Email address: jxue@tsinghua.edu.cn