Published in final edited form as:

Astron J. 2018 October ; 156(4): . doi:10.3847/1538-3881/aad95c.

On the Dynamics of the Inclination Instability

Ann-Marie Madigan¹, Alexander Zderic¹, Michael McCourt², Jacob Fleisig¹

¹JILA and Department of Astrophysical and Planetary Sciences, CU Boulder, Boulder, CO 80309, USA

²Astronomy Department, Broida Hall, University of California, Santa Barbara, CA 93106, USA

Abstract

Axisymmetric disks of eccentric Kepler orbits are vulnerable to an instability that causes orbits to exponentially grow in inclination, decrease in eccentricity, and cluster in their angle of pericenter. Geometrically, the disk expands to a cone shape that is asymmetric about the mid-plane. In this paper, we describe how secular gravitational torques between individual orbits drive this “inclination instability”. We derive growth timescales for a simple two-orbit model using a Gauss N -ring code, and generalize our result to larger N systems with N -body simulations. We find that two-body relaxation slows the growth of the instability at low N and that angular phase coverage of orbits in the disk is important at higher N . As $N \rightarrow \infty$, the e-folding timescale converges to that expected from secular theory.

Keywords

celestial mechanics-minor planets, asteroids; general-planets and satellites; dynamical evolution and stability

1. INTRODUCTION

In Madigan & McCourt (2016), we introduced a new dynamical instability in axisymmetric Kepler disks of eccentric orbits. We showed that secular (orbit-averaged) gravitational torques between orbits in the disk drive exponential growth of their inclinations. As the orbits’ inclinations grow, they tilt in the same way with respect to the disk plane. This leads to clustering in their angles of pericenter and the initially thin disk expands into a cone shape. Concurrently, the orbital eccentricities decrease and perihelion distances increase. We proposed that this could be at work in the outer solar system between minor planets scattered to large orbital eccentricities by the giant planets. It can explain the observed clustering in the angle of pericenter and detached, high perihelia objects (Trujillo & Sheppard 2014).

This paper serves to explain the physics behind the inclination instability. In section 2 we describe the mechanism of the instability through a simple, yet predictive, two-orbit toy model. We show analytically that inclinations grow exponentially in disks of eccentric

orbits, and numerically derive the associated growth timescale. We discuss our results and their application to the outer solar system in section 3.

2. MECHANISM OF INSTABILITY

We consider a thin, axisymmetric disk of particles on eccentric orbits ($e \gtrsim 0.6$) around a much more massive central object. All orbits take part in the inclination instability which changes their orientation. However, to illustrate the instability's physical mechanism, we consider just two orbits within the disk and see how this system responds to a small perturbing force. The response of the system to the small perturbing force is best described with a set of angles different than the traditionally used Kepler elements (inclination i , longitude of ascending node Ω , and argument of pericenter ω). These new angles are i_a , i_b , and i_e , and we originally introduced them in Madigan & McCourt (2016). The angles represent rotations of the orbit about its semi-major ($\hat{a}$) axis, semi-minor ($\hat{b} \equiv \hat{j} \times \hat{a}$) axis, and angular momentum vector ($\hat{j}$), respectively,¹, i.e.,

$$i_a = \arctan \left[\frac{\hat{b}_z}{\sqrt{1 - \hat{b}_z^2}} \right], \quad (1a)$$

$$i_b = \arctan \left[-\frac{\hat{a}_z}{\sqrt{1 - \hat{a}_z^2}} \right], \quad (1b)$$

$$i_e = \arctan[\hat{a}_y, \hat{a}_x]. \quad (1c)$$

The subscripts x , y , and z denote an inertial Cartesian reference frame with unit vectors, $\hat{x}$, $\hat{y}$, and $\hat{z}$. The angles i_a , i_b , and i_e are equivalent to that of an aircraft's roll, pitch and yaw. We give a graphical representation of i_a and i_b in figure 1 and convert between these variables and Kepler elements in the Appendix.

2.1. Two-orbit toy model

We consider how two initially coplanar, anti-aligned eccentric orbits respond to a perturbing force, as depicted in figure 1.

1. We perturb orbit 1 with a uniform force in the $+\hat{z}$ -direction. Such a force could arise from, e. g., a temporary fluctuation in the center of mass of the disk. Because of the orbit's symmetry around the $\hat{a}$ axis, all torques in the $\hat{a}$ -direction cancel. However, because of the orbit's asymmetry around the $\hat{b}$ axis, the vertical

¹Note that these are not true rotation angles. See the Appendix for more information.

force perturbation creates a net torque in the $\hat{b}$ -direction. This asymmetry between the torques is further exacerbated by the time-averaged torque, $\bar{\tau} = \overline{\mathbf{r} \times \mathbf{f}}$, at apocenter being greater than at pericenter because the particle spends a greater fraction of its time at apocenter, ($\sim [(1+e)/(1-e)]^{3/2}$). Thus, the net time-averaged torque points in the $+\hat{b}$ -direction. Consequently, the angular momentum vector $\mathbf{j}$ of orbit 1 rotates counter clockwise about its semi-major axis, rolling the orbital plane such that $\dot{i}_a < 0$.

2. The roll of orbit 1 induces out-of-plane forces along the sides of orbit 2. These forces are asymmetric around the $\hat{a}$ axis and imply a time-averaged torque on orbit 2 in its $+\hat{a}$ -direction. This torque rotates orbit 2 about its semi-minor axis, pitching the orbital plane such that $\dot{i}_b > 0$.
3. This rotation raises orbit 2's apocenter, and thus its center of mass, above the original orbital plane. The raising of orbit 2's apocenter raises the center of mass of the disk, which reinforces the initial perturbation. Thus, the disk is unstable to the initial perturbation, and the inclination will run away to large values.

As the orbits incline off the disk plane, conservation of angular momentum requires that they decrease in eccentricity. They do so via the same secular gravitational torques that drive growth in inclination. Let us take orbit 1 as an example. After orbit 1 begins to roll over its major axis, its angular momentum vector is no longer parallel to the initial perturbing force. A component of the force acts therefore in its orbital plane. This component causes a positive time-averaged torque resulting in a decrease in eccentricity. A similar analysis shows that orbit 2 must also decrease in eccentricity.

Though we have only considered two orbits in the disk, all orbits take part in the instability. We show this in figure 2 in which we plot the median values of $\dot{i}_a$ and $\dot{i}_b$ (top), and eccentricity (bottom) for a representative N -body simulation. We refer the reader to figure 4 to clearly see exponential growth of $\dot{i}_b$. Here we use the REBOUND code (Rein & Liu 2012), with the IAS15 integrator (Rein & Spiegel 2015). We distribute $N=100$ massive particles in a thin axisymmetric disk in the potential of a central body of mass M . The orbit eccentricities are initialized to $e=0.9$, semi-major axes are uniformly distributed in the narrow range $[0.9, 1.1]$. Mean anomaly, argument of pericenter, and angle of ascending node are distributed uniformly in $(0, 2\pi]$.

The two-orbit toy model makes several testable predictions confirmed by N -body experiments (see also figure 2 of Madigan & McCourt (2016)):

1. If the magnitudes of the torques are linear in $\dot{i}_a$ and $\dot{i}_b$, as one might expect for small deviations from an equilibrium, this instability should grow exponentially in time.
2. The orbits should grow their $|\dot{i}_a|$ and $|\dot{i}_b|$ values with opposite signs and $\dot{i}_b/\dot{i}_a$ in constant ratio.
3. A constant ratio $\dot{i}_b/\dot{i}_a$ implies a constant angle of pericenter, as for small inclinations $\omega(\dot{i}_a, \dot{i}_b) \sim \arctan |\dot{i}_b/\dot{i}_a|$ ($+\pi$ if $\dot{i}_a < 0$). As $\dot{i}_a$ and $\dot{i}_b$ depend on the

eccentricities of orbits, $\omega(i_a, i_b)$ should cluster to one of two eccentricity-dependent values, depending on the direction of the initial perturbing force.

2.2. Growth timescale of instability

We estimate the e-folding timescale of the instability between two orbits as follows. For a given orbit with its $\hat{a}$ and $\hat{b}$ axes, two components of the torque that rotate its orbital plane are:

$$\tau_a \equiv \boldsymbol{\tau} \cdot \hat{\mathbf{a}}, \quad (2a)$$

$$\tau_b \equiv \boldsymbol{\tau} \cdot \hat{\mathbf{b}}. \quad (2b)$$

We denote the torque on orbit 2 due to orbit 1 as $\boldsymbol{\tau}^{(2,1)}$ and the torque on orbit 1 due to orbit 2 as $\boldsymbol{\tau}^{(1,2)}$. Torque acting on orbit 2 can be expressed as

$$\tau_a^{(2,1)}(i_a^{(1)}) = \frac{dj^{(2)}(i_b^{(2)})}{dt} = j^{(2)} \frac{di_b^{(2)}}{dt}, \quad (3)$$

(where the last step is simply vector rotation), and expanded via Taylor series and linearized

$$\tau_a^{(2,1)}(i_a^{(1)}) = \left(\frac{d\tau_a^{(2,1)}}{di_a^{(1)}} \right) i_a^{(1)}. \quad (4)$$

Since i_a represents a rotation of angular momentum in the $(-\hat{\mathbf{b}})$ direction and i_b represents a rotation of angular momentum in the $\hat{\mathbf{a}}$ direction, we have the linearized equations of motion:

$$j^{(2)} \frac{di_b^{(2)}}{dt} = \tau_a^{(2,1)}(i_a^{(1)}) = \left(\frac{d\tau_a^{(2,1)}}{di_a^{(1)}} \right) i_a^{(1)} = \alpha j^{(2)} i_a^{(1)}, \text{ and} \quad (5a)$$

$$j^{(1)} \frac{di_a^{(1)}}{dt} = -\tau_b^{(1,2)}(i_b^{(2)}) = - \left(\frac{d\tau_b^{(1,2)}}{di_b^{(2)}} \right) i_b^{(2)} = -\beta j^{(1)} i_b^{(2)} \quad (5b)$$

where in the last step we have defined the coefficients

$$\alpha \equiv \frac{1}{j^{(2)}} \left(\frac{d\tau_a^{(2,1)}}{di_a^{(1)}} \right) = \frac{1}{j^{(1)}} \left(\frac{d\tau_a^{(1,2)}}{di_a^{(2)}} \right), \quad (6a)$$

$$\beta \equiv \frac{1}{j^{(1)}} \left(\frac{d\tau_b^{(1,2)}}{di_b^{(2)}} \right) = \frac{1}{j^{(2)}} \left(\frac{d\tau_b^{(2,1)}}{di_b^{(1)}} \right). \quad (6b)$$

There are similar expressions for $i_a^{(1)'}(t)$ and $i_b^{(2)'}(t)$. In the following, we assume the parameters α and β are constants, which is consistent with N -body simulations of the instability. We then differentiate equation 5a and combine it with equation 5b to yield an equation of motion for i_b :

$$i_b^{(2)''}(t) = \alpha i_a^{(1)'}(t) = -\alpha\beta i_b^{(2)}. \quad (7)$$

For small inclinations, the inclination angles evolve in time as

$$\{i_a, i_b\} \propto \exp(\gamma t), \quad (8)$$

$$\gamma \equiv \sqrt{-\alpha\beta}, \quad (9)$$

where γ is the growth rate of the instability. Equation 7 shows that the instability exponentially grows orbit inclinations through mutual secular gravitational torques between eccentric orbits.

We numerically evaluate α and β from equation 6 using Gauss' approximation, as described in Gürkan & Hopman (2007).² We estimate the time-averaged torques between two orbits by smearing the bodies out into “wires” with a linear mass density along the orbit inversely proportional to the instantaneous Kepler velocity; we then split each wire into n segments equally spaced in eccentric anomaly, and we directly sum the n^2 gravitational torques of all pairs of segments. We repeatedly double the number of segments until the total torque converges. Though our numeric estimate for the time-averaged torques between orbits does not require gravitational softening, we include it in our calculation as arbitrarily large torques can arise for certain orientations of the pair of orbits. This is unsatisfactory from a theoretical perspective because the large torques arising over small spatial scales destroy the linearity presumed in equation 5. These large torques are also unphysical, because they would drive scattering or diffusion of the orbits on a timescale shorter than the instability

²Our source code is publicly available at <https://github.com/mkmcc/orbit-torque>.

growth timescale. They therefore should not be included in our calculation, and gravitational softening is one way to approximate this behavior. We adopt a fiducial force-softening length $s = (3 \times 10^{-2}) a$, chosen to exclude close encounters that would drive diffusion on a timescale shorter than $1/\gamma$. With this choice, the linear theory we present applies over the majority of the parameter space. The boundaries between stable and unstable regions do not depend significantly on our chosen value of s . However, the maximum growth rate scales with the softening length as $\gamma_{\max} \sim s^{-3/4}$.

Using this numerical approach, we calculate the squared growth rate, γ^2 , as a function of orbital eccentricity and angular separation, $\theta = \arccos(\hat{e}_1 \cdot \hat{e}_2)$. We initialize the equal-mass orbits with similar semi-major axes. The growth rate is normalized to the orbital period, so that γ is the number of e-foldings per orbit. We show γ^2 as a function of θ and e in figure 3. At every eccentricity, the orbits are unstable in some range of angular separations. The largest growth rates occur at the highest orbital eccentricities and small-to-intermediate angular separations. In general, as eccentricity increases, the range of stable angular configurations shrinks. At very small angular separations, the two orbits are unstable only for very high eccentricities. Interestingly, our results show that the angular separation ($\theta = \pi$) used by the toy model in section 2.1 is the least-unstable unstable region of figure 3. The inclination instability only occurs in high eccentricity disks ($e \gtrsim 0.6$) because below this boundary most orbits are stable with respect to one another; tests with N -body simulations confirm this.

There is a marked boundary in $e - \theta$ space between stable and unstable configurations. The location of this boundary can be understood from the two-orbit model: For low θ , orbits exert stabilizing vertical forces on each other because their sides overlap. As θ increases, the distance between their sides grows until the force acting on each side reverses and becomes unstable. This reversal happens at smaller separations for high eccentricity orbits as the width of an orbit scales as $\sqrt{1 - e^2}$. Orbits of nonzero eccentricity are unstable at $\theta = \pi/2$. As θ increases towards π , the growth rate of the instability decreases due to the increasing distance and decreasing gravitational force between the orbits. Low-eccentricity orbits are the most stable as forces at pericenter become stronger than those at apocenter. The squared growth rate γ^2 is negative in the stable regions which implies that the two orbits may be undergoing oscillatory behavior. We will explore the dynamical behavior of the orbits in the regions of stability in a future study.

2.3. Generalization to N bodies

To apply the instability to real astrophysical systems with more than two bodies, we need to know the growth rate for large N systems. We start with the hypothesis that the inclination instability acts on a secular timescale,

$$t_{\text{sec}} \sim \frac{1}{2\pi} \frac{M}{M_{\text{disk}}} P, \quad (10)$$

where P is the orbital period. Here t_{sec} is derived as the time it takes to change an orbit's angular momentum by order of its circular angular momentum using a specific torque over one orbital period of $\tau \sim Gm/a$. To test this hypothesis, we use numerical N -body simulations with the REBOUND code. We simulate the instability for a range of disk masses ($M_{\text{disk}} = [10^{-2}, 10^{-3}, 10^{-4}]$, where the central mass $M = 1$), and number of particles ($N = [100, 200, 500, 1000]$), averaging many tens of simulations for each parameter set to reduce the noise. Each particle has $m = M_{\text{disk}}/N$, $e = 0.7$, and $i = 10^{-4}$ rad. Semi-major axes are distributed uniformly in $[0.9, 1.1]$. Mean anomaly, argument of pericenter, and angle of ascending node are randomly distributed in $(0, 2\pi]$. We do not use gravitational softening.

We show the results of our simulations in figure 4. The top plot shows the mean inclination of particles as a function of time for simulations with a fixed total mass $M_{\text{disk}} = 10^{-2}M$. The various lines correspond to simulations with differing numbers of particles. A surprising result is how dependent the inclination growth rate is on the number of particles per simulation. The instability proceeds faster with increasing N . The bottom plot extends this result for different disk masses, showing the growth rate γ as a function of N for $M_{\text{disk}} = [10^{-2}, 10^{-3}, 10^{-4}]$. We derive the growth rate in each simulation from the linear parameter of a quadratic model fit to the median $\log i_b$ value³. We optimize the fit parameters with a linear least-squares method and find the error using a bootstrapping method. Our data is Gaussian distributed so a least-squares optimization is equivalent to a maximum likelihood estimator (MLE) approach.

If the e-folding time of the instability matched the secular timescale given by equation 10, the data in figure 4 would align with the dashed gray lines. This is clearly not the case however, particularly at low N and large M_{disk} . We identify three reasons for the deviation of the instability growth rate from a secular one.

1. Angular phase coverage of orbits in disk—In the two-orbit model, the instability growth rate strongly depends on the angular separation of orbits (see figure 3). In fact, large sections of phase space are stable ($\gamma^2 < 0$). Generalizing to N bodies, the angular phase coverage of orbits in a disk greatly impacts the overall dynamics, including whether or not the disk is stable or unstable. Non-axisymmetric disks are stable; a good example being the eccentric nuclear disk of stars in the Andromeda nucleus (Madigan et al. 2018). Furthermore, in the limit of small N , few angular separations are realized and axisymmetric disks can also be stable; tests show that at $N \lesssim 20$ the instability doesn't occur within the predicted timescale. As we randomly select the angular orientation of the orbits in our simulations, the unstable region of phase space gets filled at a rate $\propto N^{1/2}$. Assuming the most unstable angular separation dominates the growth rate, we would then expect the growth rate to follow an $N^{1/2}$ dependence.

2. Apsidal precession—In limit of $M_{\text{disk}} \rightarrow M$, the Kepler orbit approximation breaks down. Orbits precess rapidly due to the gravitational potential of the disk, and we expect

³We use i_b instead of i_a to calculate the exponential growth rate as it is the less noisy parameter (it requires more force for an orbit to change its i_b value as it must lift its apocenter).

orbit-averaged torques to weaken significantly. Hence the timescale for the instability should strongly deviate from the secular timescale at large M_{disk}/M values, as seen in figure 4.

3. Two-body relaxation—In a near-Keplerian system, the two-body relaxation time is given by

$$t_{\text{rel}} \equiv \frac{M^2}{m^2 N \ln(\Lambda)} \quad (11)$$

(Rauch & Tremaine 1996), where $\ln(\Lambda)$ is the Coulomb logarithm. Therefore, the ratio of the two-body timescale to the secular timescale is

$$\frac{t_{\text{rel}}}{t_{\text{sec}}} \sim \frac{M}{m \ln(\Lambda)} \propto m^{-1}. \quad (12)$$

As m decreases, the two-body relaxation timescale increases relative to the secular timescale. Conversely, for larger values of m , strong two-body scattering perturbs particles off their orbits and slows the rate of the secular instability. Thus, we expect the growth rate to have a m^{-1} dependence from two-body relaxation. For a fixed M_{disk} , this will emerge as an N^{-1} dependence.

All three of these physical effects coexist simultaneously in our simulations. To explore their relative importance, we parameterize deviation from the secular growth rate with an arbitrary function α ,

$$\gamma^{-1} = \alpha(N, M_{\text{disk}}) t_{\text{sec}}. \quad (13)$$

We fit our simulation data using two different models, $\alpha = aN^{-1/2} + b$ and $\alpha = aN^{-1} + b$, for disk phase-coverage and two-body relaxation, respectively. The phase coverage model most accurately fits the large N evolution of the instability timescale ($N = 200$ – 1000).

Additionally, it predicts that α converges to the same value for all M_{disk} ($b \approx 0.4$). However, the two-body scattering model more accurately fits the low N data points. Therefore, we see a mixture of phase coverage and two-body scattering for the N we simulated. In both models, the value of the a parameter decreases with M_{disk} . That is, the growth rate's dependence on N weakens as M_{disk} decreases. Apsidal precession is the likely source of this additional M_{disk} dependence. In summary, our exploration into the growth rate of the instability for large N , low M_{disk} systems finds that the e-folding timescale is

$$t_{\text{e-fold}} \sim \frac{0.2}{\pi} \frac{M}{M_{\text{disk}}} P. \quad (14)$$

3. DISCUSSION

In this paper, we describe the mechanism behind the inclination instability in axisymmetric, eccentric Kepler disks as introduced by Madigan & McCourt (2016). The instability results in an exponential growth of orbital inclinations, a decrease in orbital eccentricities, and clustering in the argument of pericenter during its linear phase. We explain how the instability works with a simple toy model, examining the secular torques between just two orbits (section 2.1). Using linearized equations of motion, we analytically demonstrate the exponential growth of the instability and numerically derive the e-folding rate (section 2.2). We expand to large N systems (section 2.5) with a suite of N -body simulations. Our results show that two-body scattering, angular phase coverage of orbits within the disk, and apsidal precession reduce the growth rate of the instability. At low N ($\approx 10^2$), two-body scattering is the dominant physical process decreasing the growth rate. At large N , the angular phase coverage of orbits is important. Apsidal precession further decreases the growth rate as $M_{\text{disk}} \rightarrow M$. This result is general to all N -body simulations involving secular dynamical behavior. Thus, we conclude that N -body simulations (which for computational efficiency typically use low N and high M_{disk}/M values) generally underestimate the strength of secular dynamics.

3.1. Outer Solar System

The results of section 2.3 can be used to apply the inclination instability to real astrophysical systems that are currently impossible to directly simulate. Studies of inner Oort cloud formation in a stellar cluster suggest that 1–10 Earth masses of cometary material may exist at hundreds of au (Brasser et al. 2012). The comets are scattered outward by Jupiter and Saturn, and their pericenters are later decoupled from the planets by perturbations from cluster gas and nearby stars. We compare important dynamical timescales for bodies in this region, taking the orbital parameters of the minor planet Sedna as an example (Brown et al. 2004, $e \approx 0.85$, $a \approx 507$ au, $i \approx 12^\circ$).

The induced motion of Sedna's orbit due to the four giant planets can be approximately calculated by assigning to the Sun an artificial quadrupole moment,

$$J_2 = \frac{1}{2} \sum_i^4 \frac{m_i}{M_\odot} \left(\frac{a_i}{R_\odot} \right)^2. \quad (15)$$

The apsidal and nodal precession frequencies are then given by

$$\dot{\omega} = \frac{3J_2}{4} n \left(\frac{R_\odot}{a} \right)^2 \left(\frac{5 \cos^2 i - 1}{(1 - e^2)^2} \right), \quad (16)$$

and

$$\dot{\Omega} = -\frac{3J_2}{2}n\left(\frac{R_\odot}{a}\right)^2\left(\frac{\cos i}{(1-e^2)^2}\right), \quad (17)$$

where $n = 2\pi/P$ is the mean motion of Sedna's orbit (Brasser et al. 2006). With $J_2 = 2.6149 \times 10^3$, the corresponding precession timescales are

$$t_{\text{apsidal}} = 2\pi/\dot{\omega} \approx 1.4 \text{ Gyr}, t_{\text{nodal}} = 2\pi/\dot{\Omega} \approx 2.6 \text{ Gyr}.$$

The Kozai-Lidov timescale for Sedna due to the galactic potential is,

$$t_{\text{KL}} \sim 4.0 \left(\frac{0.1 M_\odot \text{pc}^{-3}}{\rho_G} \right) \left(\frac{10^4 \text{au}}{a} \right)^{3/2} \text{Gyr} = 360 \text{ Gyr}, \quad (18)$$

where ρ_G is the mass density of the Galaxy in the solar neighborhood, and the angular momentum diffusion timescale for Sedna from stellar passages is,

$$t_{\text{diff}} \sim 25 \left(\frac{q}{30 \text{ au}} \right) \left(\frac{10^4 \text{au}}{a} \right)^2 \text{Myr} = 25 \text{ Gyr}, \quad (19)$$

with the timescales taken from Brasser et al. (2008)⁴. The timescale for apsidal precession induced by the giant planets is the shortest of these three. The inclination e-folding timescale for Sedna is,

$$t_{\text{inst}} \sim 0.4 \frac{1}{2\pi} \left(\frac{M_\odot}{M_{\text{disk}}} \right) P = 725 \left(\frac{M_\odot}{M_{\text{disk}}} \right) \text{yr}. \quad (20)$$

If $M_{\text{disk}} \approx 0.18 M_\oplus$, the instability timescale is of the same order of magnitude as the apsidal precession timescale. Thus, we find that in this scenario the self-gravity of small bodies in the Sedna region is the dominant dynamical driver⁵.

If we instead hypothesize that the minor planets in the Sedna region came from the scattered disk (a natural starting point for the instability, which requires only eccentric orbits in an axisymmetric disk), and detached themselves via self-gravitational torques, the more relevant comparison is that of the instability e-folding timescale to the energy diffusion time. The energy diffusion time due to planetary perturbations is given by Duncan et al. (1987)

⁴The angular momentum diffusion timescale and Kozai-Lidov timescale were derived for $e \sim 1$ orbits in Brasser et al. (2008).

⁵Of course this statement does not hold if a massive planet 9 (Batygin & Brown 2016) exists in this region.

$$t_{D(x)} \approx 1 \times 10^6 \left(\frac{10^4 \text{ au}}{a} \right)^{1/2} \left(\frac{10^{-4} \text{ au}^{-1}}{D(x)} \right)^2 \text{ yr}, \quad (21)$$

where $D(x)$ is the typical energy perturbation per perihelion passage. For a body originating in the scattered disk with perihelion distance of $p \approx 35 \text{ au}$ and a low orbital inclination, $D(x) \approx 10^{-5} \text{ au}^{-1}$ (figure 1 of Duncan et al. 1987). This yields an energy diffusion time at Sedna's semi-major axis of $t_{D(x)} \approx 450 \text{ Myr}$, consistent with results from Bannister et al. (2017). If $M_{\text{disk}} \approx 0.5 M_{\oplus}$, the instability timescale equals the energy diffusion timescale. Thus in this scattered disk scenario, the self-gravity of small bodies in the Sedna region is the dominant dynamical driver if there exists about half an Earth mass of material at hundreds of au. In general, self-gravity in the inner Oort Cloud acts to increase the clearing timescale due to planetary perturbations. Gravitational torques increase the orbital inclination and the perihelion distance of minor planets, thus weakening the strength of scattering events. Minor planets can be more effectively retained.

Fan & Batygin (2017) recently presented simulations of self-gravitating planetesimal disks within the context of the Nice Model. They found no evidence of the inclination instability, attributing its absence to mass depletion via ejections, and rapid precession of orbits due to the presence of the giant planets. We note however that the simulations ran for only 20 Myr during the most violent phase of solar system evolution. We do not expect the inclination instability to dominate while the giant planets are migrating tens of au. Furthermore, there is the important issue of particle number. If the original disk mass of $30 M_{\oplus}$ depletes to $\approx 0.5 M_{\oplus}$, the number of particles reduces from $N = 10^3$ to $N \approx 16$. The instability cannot be captured at this low particle number, as discussed in section 2.3. Thus, we cannot conclude from the Fan & Batygin (2017) results that the Nice model is incompatible with the inclination instability⁶.

If the observed clustering of arguments of pericenter (ω) of minor planets is a real effect (see Shankman et al. 2017), we need to either (i) be catching the instability during its linear phase, or (ii) be in the nonlinear, saturated phase and have something maintaining the clustering. It is interesting to note that the cone-shaped gravitational potential produced by the instability induces precession of eccentric orbits such that $\dot{\omega} < 0$, while the presence of the giant planets induces the opposite, $\dot{\omega} > 0$ (for $\cos i < 1/\sqrt{5}$ see equation 16). There may well be an interesting region of phase space where the two cancel and clustering is maintained (see Sefilian & Tوما (2018) for a similar idea involving longitude of perihelion (ϖ) clustering with a massive, apse-aligned disk). We have, so far, focused on the linear phase of the inclination instability in an idealized set-up. In future studies, we will directly apply the inclination instability to the solar system and examine its nonlinear evolution (Zderic et al. 2019, in preparation.). We will include the gravitational influence of the giant planets, and vary our initial conditions to reflect the scattered disk population and the inner

⁶It's interesting that the success ratio of the Nice model in reproducing the architecture of the current day solar system is a factor of ~ 10 lower than in simulations without self-gravity. Self-gravity of minor planets would appear to be a crucial component in the dynamical evolution of the solar system.

Oort cloud formation in a stellar cluster. Finally, instead of presenting results averaged for many particles, we will focus on individual minor planets, particularly those that flip to retrograde orbits.

Acknowledgments

We thank the anonymous referee for their helpful comments. We gratefully acknowledge support from NASA Solar System Workings (SSW) under grant number 80NSSC17K0720. Resources were provided by the NASA High-End Computing (HEC) Program through the NASA Advanced Supercomputing (NAS) Division at Ames Research Center. We also used the RMACC Summit supercomputer, which is supported by the National Science Foundation (awards ACI-1532235 and ACI-1532236), the University of Colorado Boulder, and Colorado State University. Simulations in this paper made use of the REBOUND code which can be downloaded freely at <https://github.com/hannorein/rebound>. We used the open-source software TIOGA to make our plots.

APPENDIX

A. ELEMENT CONVERSIONS

The inclination instability is due primarily to forces acting *out* of the orbital plane. These forces exert torques on the orbit, $\boldsymbol{\tau} = \mathbf{r} \times \mathbf{f} = d\mathbf{j}/dt$, which result in rotations of the $\mathbf{j}$ vector. Whereas none of the standard Kepler elements remain constant during the instability, certain combinations of them do. Hence in Madigan & McCourt (2016) we introduced two new coordinates that represent rotation angles⁷ about the semi-major axis of the orbit, i_a , and about the semi-minor axis⁸ of the orbit, i_b . To complete a transformation from Kepler elements, we also introduced i_e , the angle between the positive $\hat{x}$ -axis and the projection of the semi-major axis (or eccentricity vector) in the xy -plane. In the following conversions, we use $\hat{a}$ and $\hat{e}$ interchangeably as they are equivalent.

Inclination angles i_a , i_b , and $i_e \leftrightarrow$ orbit vectors $\hat{a}$, $\hat{j}$, and $\hat{b} (\equiv \hat{j} \times \hat{e})$:

$$\cos i = \cos i_a \cos i_b \left[1 - (\tan i_a \tan i_b)^2 \right]^{1/2} \quad (\text{A1a})$$

$$\hat{a} = \begin{bmatrix} \cos i_b \cos i_e \\ \cos i_b \sin i_e \\ -\sin i_b \end{bmatrix} \quad (\text{A1b})$$

⁷The inclination angles i_a and i_b are defined in equation 1; with this definition, i_a and i_b are independent quantities, but they are not formally rotation angles. They are rotation angles to first order, however, so this difference does not affect our analysis in this paper.

⁸Strictly speaking, an orbit does not rotate about its minor axis, rather about a chord parallel to its minor axis which runs through the focus of the ellipse, the latus rectum.

$$\hat{\mathbf{b}} = \begin{bmatrix} \tan i_b \sin i_a \cos i_e - (\cos i / \cos i_b) \sin i_e \\ \tan i_b \sin i_a \sin i_e + (\cos i / \cos i_b) \cos i_e \\ \sin i_a \end{bmatrix} \quad (\text{A1c})$$

$$\hat{\mathbf{j}} = \begin{bmatrix} \tan i_b \cos i \cos i_e + (\sin i_a / \cos i_b) \sin i_e \\ \tan i_b \cos i \sin i_e - (\sin i_a / \cos i_b) \cos i_e \\ \cos i \end{bmatrix} \quad (\text{A1d})$$

Kepler elements $\leftrightarrow$ orbit vectors $\hat{\mathbf{e}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{b}}$:

$$\hat{\mathbf{e}} = \begin{pmatrix} \cos \omega \cos \Omega - \sin \omega \cos i \sin \Omega \\ \cos \omega \sin \Omega + \sin \omega \cos i \cos \Omega \\ \sin \omega \sin i \end{pmatrix} \quad (\text{A2a})$$

$$\hat{\mathbf{j}} = \begin{pmatrix} \sin i \sin \Omega \\ -\sin i \cos \Omega \\ \cos i \end{pmatrix} \quad (\text{A2b})$$

$$\hat{\mathbf{b}} = \begin{pmatrix} -\sin \omega \cos \Omega - \cos \omega \cos i \sin \Omega \\ -\sin \omega \sin \Omega + \cos \omega \cos i \cos \Omega \\ \sin i \cos \omega \end{pmatrix} \quad (\text{A2c})$$

$$\mathbf{n} = \hat{\mathbf{z}} \times \mathbf{j} \quad (\text{A3a})$$

$$i = \arccos[\hat{j}_z / j] \quad (\text{A3b})$$

$$\Omega = \arctan[\hat{n}_y, \hat{n}_x] \quad (\text{A3c})$$

$$\omega = \arctan\left[\operatorname{sgn}(\hat{e}_z)\left[1 - (\hat{n} \cdot \hat{e})^2\right]^{1/2}, \hat{n} \cdot \hat{e}\right] \quad (\text{A3d})$$

Kepler elements $\leftrightarrow$ inclination angles i_a , i_b , and i_c (note that these definitions are valid for prograde orbits only):

$$i_a = \arcsin(\sin i \cos \omega) \quad (\text{A4a})$$

$$i_b = \arcsin(-\sin i \sin \omega) \quad (\text{A4b})$$

$$i_c = \Omega - \arctan[\cos \omega, \cos i \sin \omega] + \pi/2 \quad (\text{A4c})$$

$$i = \arccos\left[\cos i_a \cos i_b \left[1 - (\tan i_a \tan i_b)^2\right]^{1/2}\right] \quad (\text{A5a})$$

$$\omega = \arctan\left[-\operatorname{sgn}(i_b) \sqrt{1 - \left(\frac{\sin i_a}{\sin i}\right)^2}, \frac{\sin i_a}{\sin i}\right] \quad (\text{A5b})$$

$$\Omega = i_c + \arctan\left[\tan i_b \cos i, \frac{\sin i_a}{\cos i_b}\right] \quad (\text{A5c})$$

REFERENCES

- Bannister MT, Shankman C, Volk K, et al. 2017, AJ, 153, 262
 Batygin K, & Brown ME 2016, AJ, 151, 22
 Brasser R, Duncan MJ, & Levison HF 2006, Icarus, 184, 59—.2008, Icarus, 196, 274
 Brasser R, Duncan MJ, Levison HF, Schwamb ME, & Brown ME 2012, Icarus, 217, 1
 Brown ME, Trujillo C, & Rabinowitz D 2004, ApJ, 617, 645
 Duncan M, Quinn T, & Tremaine S 1987, AJ, 94, 1330
 Fan S, & Batygin K 2017, ApJ, 851, L37
 Gürkan MA, & Hopman C 2007, MNRAS, 379, 1083
 Madigan A-M, Halle A, Moody M, et al. 2018, ApJ, 853, 141
 Madigan A-M, & McCourt M 2016, MNRAS, 457, L89
 Rauch KP, & Tremaine S 1996, New Astronomy, 1, 149
 Rein H, & Liu S-F 2012, A&A, 537, A128

- Rein H, & Spiegel DS 2015, MNRAS, 446, 1424
Sefilian AA, & Tوما JR 2018, ArXiv e-prints, arXiv:1804.06859
Shankman C, Kavelaars JJ, Bannister MT, et al. 2017, AJ, 154, 50
Trujillo CA, & Sheppard SS 2014, Nature, 507, 471 [PubMed: 24670765]

Figure 1.

Idealized, “two-orbit” toy model for the inclination instability. The top right of each panel shows the location of the orbit in the disk from a face-on perspective. (*top*): Orbit 1 experiences a net upward force f . This force produces a torque along the $\hat{b}$ axis, rotating the orbital plane such that $i_a < 0$. (*bottom*): Diametrically opposed orbit 2 feels a force due to the rotation of orbit 1; this force produces a torque along the $\hat{a}$ direction of orbit 2, rotating its orbital plane such that $i_b > 0$.

Figure 2.

(*top*): Median values of angles i_a and i_b for all particles in a disk as a function of time in units of orbital period. Shaded regions indicate one-sigma quantile contours. Both i_a and i_b rise exponentially at early times. As predicted by the two-orbit toy model, i_a and i_b increase with opposite sign. (*bottom*): Median value of orbital eccentricity for the same particles. As inclinations increase, eccentricities decrease.

Figure 3.

Squared growth rate γ^2 for two orbits gravitationally torquing each other. This is estimated from the linearized equations of motion for the system (equation 9), as a function of orbital eccentricity and angular separation between orbits in the plane, θ . Only the unstable region with $\gamma^2 > 0$ is colored. The white, uncolored region of the plot with dashed contours is stable. The thick, gray band near the $\gamma^2 = 0$ boundary marks the region in which the linear approximation in equation 5 breaks down, causing an order-unity uncertainty in our estimate of the timescale. This region also depends on the softening length we adopt.

Figure 4.

(top): Exponential growth of the inclination angle i_b for different numbers of disk particles (N) and fixed total mass $M_{\text{disk}} = 10^{-2}M$. Time is in units of orbital period at semimajor axis $a = 1$. Note that the slopes (growth rate) of the lines show a strong dependence on N .

(bottom): Instability growth rate (γ) for different values of N and M_{disk} . The expected secular growth rate (τ_{sec}^{-1}) is plotted as dashed gray lines (corresponding from top to bottom to disk masses of $M_{\text{disk}} = [10^{-2}, 10^{-3}, 10^{-4}]$). Note that the growth rate's dependence on N weakens as M_{disk} is reduced, and ultimately converges to the secular growth rate.